NBER WORKING PAPER SERIES

AN INTANGIBLES-ADJUSTED PROFITABILITY FACTOR

Ravi Jagannathan
Robert Korajczyk
Kai Wang

Working Paper 31068
http://www.nber.org/papers/w31068

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2023

Ravi Jagannathan & Robert Korajczyk : We are consultants to DSC Quantitative Group, which has created a portfolio of benchmark and investable indices seeking to track the gross performance of both the US private equity and venture capital industries. In our opinion, DSC Quantitative Group's policy positions, goals, or financial interests do not relate to this article. However, as a matter of abundant caution, we are disclosing this relationship. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

An Intangibles-Adjusted Profitability Factor
Ravi Jagannathan, Robert Korajczyk, and Kai Wang
NBER Working Paper No. 31068
March 2023
JEL No. G0,G1,G12

ABSTRACT

We treat expenditures that create intangible assets as investments and instead of expensing them, we add them back to earnings when measuring the return on equity of firms while constructing the profitability factor in the Fama and French (2015) five factor model. The profitability factor we construct has significant alpha relative to many extant multi-factor asset-pricing models, including the standard Fama-French five factor model. When the profitability factor in the Fama and French (2015) five factor model is replaced with our intangibles adjusted profitability factor, the model performs better in explaining the cross section of stock returns, and several anomalies documented in the literature. Portfolios that exploit price momentum, earnings momentum, and operating leverage no longer have significant alphas. The improvement is consistent with the dividend discount model for equity valuation. Adjusted earnings constructed by treating expenditures that create intangible assets as investments help forecast the cross section of future cash dividends and operating cash flows on stocks better, especially at longer horizons. Adopting our adjustment when constructing the monthly rebalanced profitability factor in the Hou et al. (2015) four factor model improves its performance as well. Our intangible adjusted profitability factor has smaller left tail risk and co-tail risk with the market when compared to the standard profitability factor.

Ravi Jagannathan
Kellogg Graduate School of Management
Northwestern University
2211 Campus Drive
Evanston, IL 60208
and NBER
rjaganna@kellogg.northwestern.edu

Robert Korajczyk
Department of Finance
Kellogg School of Management
Northwestern University
2211 Campus Drive, Room 4357
Evanston, IL 60208-0898
r-korajczyk@northwestern.edu

Kai Wang
Kellogg School of Management
Northwestern University
2211 Campus Dr
Evanston, IL 60208
kai.wang2@kellogg.northwestern.edu

1 Introduction

Linear factor models are widely used in the asset pricing literature to represent economy-wide pervasive risk.[1] Connor (1995) categorizes factors models into three types, statistical, macroeconomic, and fundamental, based on the method of choosing the asset pricing factors (as well as hybrid versions that combine statistical, macroeconomic, or fundamental factors). The logic of the factor model approach is that priced risks should be common, pervasive risks across assets, in the sense that they cannot be diversified away in portfolios with many assets. Statistical factor models (starting with Roll and Ross (1980), Connor and Korajczyk (1986, 1988), Lehmann and Modest (1988)) exploit the covariance structure of asset returns to estimate the important common factors. The advantage of statistical models is that they extract factors based on systematic co-movement. The disadvantages are that the factors are not explicitly linked to economic aggregates and that they have a rotational indeterminacy that makes them non-comparable across periods, although they can be rotated to match macroeconomic or fundamental factors (Connor and Korajczyk (1991)).

Fundamental financial factor models specify factor loadings (betas) of stock returns as functions of observable characteristics and estimate latent factors from the relation between asset returns and betas (for example, Rosenberg and McKibben (1973), Connor et al. (2012), Kelly et al. (2019), Kim et al. (2021), and Ge et al. (2022)). These fundamental factor models are economically interpretable to the extent that the characteristics explaining factor loadings are interpretable. This approach has been successful in addressing time variations in the elements of the covariance matrices associated with stock returns.

Macroeconomic factor models specify, in advance, the nature of the systematic factors. They have the advantage that the factors are interpretable and, ideally, should be correlated with macroeconomic variables that are linked to the marginal utility of investors. Examples are Chen et al. (1986), Breeden et al. (1989), Cochrane (1991), Braun (1993), Jagannathan and Wang (1996), Papanikolaou (2011), Kogan and Papanikolaou (2014), among others. While they have had some success in explaining time variation in stock returns at business cycle frequencies, they have several well-recognized disadvantages: macroeconomic time series often have a low correlation with asset returns, possibly due to measurement error, temporal aggregation, aggregation across heterogeneous investors, infrequent decision making due to set up costs (Jagannathan and Wang (2007), Abel et al. (2013)), model misspecification, or asset returns anticipating future changes in macroeconomic factors.

[1] Linear factor models take the market risk premium and the risk-free rate as given when explaining the cross section of returns on securities. In contrast, consumption-based asset pricing models focus on explaining the market risk premium and the risk free rate and their dynamics (see Hansen and Singleton (1982, 1983), Bansal and Yaron (2004), and Parker and Julliard (2005) among others).

Another approach in the literature is to first identify cross-sectional variation in expected asset returns based on introspection and theory. Then in a second step, construct zero investment portfolios that are long stocks with higher expected returns and short stocks with lower expected returns. To the extent the market is informationally efficient, the returns on such long-short portfolios will capture economy-wide pervasive risks.

Fama and French (1993) and Fama and French (2015) use this approach to construct a three-factor model (FF3) and a five-factor model (FF5) respectively. They use the present value relation between asset valuation and future payouts (the dividend discount model) to identify cross-sectional variation in expected returns. The FF5 model augments their original three-factor model in Fama and French (1993) (FF3) with factors based on profitability (robust-minus-weak (RMW)) and asset growth (aggressive-minus-conservative (CMA)).

We propose a modification of the FF5 profitability factor which takes into account the fact that profitability based on U.S. generally accepted accounting principles (GAAP) includes currently expensed costs that should rightly be considered investments in intangible assets (such as research and development (R&D) and portions of sales general and Administrative (SG&A) expenses). That is, GAAP profitability reflects investments in physical capital by capitalizing and depreciating those expenditures over time. However, GAAP immediately expenses R&D and SG&A expenditures.

We find that the intangible-adjusted profitability factor has positive alpha against the original FF5 profitability factor alone and the full FF5 model. Substituting the original FF profitability with the intangibles-adjusted profitability factor leads to a model that effectively subsumes some of the well-known asset pricing anomalies, such as momentum (Jegadeesh and Titman (1993)), operating leverage, and monthly-rebalanced ROE (Hou et al. (2015)).

2 The Dividend Discount Model and Implied Cost of Capital

Fama and French (2006, 2015) use the dividend discount model to motivate the relation between the market-to-book equity ratio, profitability, asset growth, and the internal rate of return (or implied cost of capital (ICC, Gebhardt et al. (2001))) on an asset's cash flows.

2.1 The Dividend Discount Model

The dividend discount model equates the market price of a share of a firm's stock to its present value of expected dividends (per share):

$$P_t = \sum_{\tau=1}^{\infty} \frac{\mathbb{E}(D_{t+\tau})}{(1+r)^{\tau}}, \tag{1}$$

where P_t is the share price, $\mathbb{E}(D_{t+\tau})$ is the expected dividend at time $t+\tau$, and r is the ICC. The dividend at any future time $t+\tau$ can be written as the earnings minus retained earnings:

$$D_{t+\tau} = Y_{t+\tau} - (Y_{t+\tau} - D_{t+\tau})$$

where the $Y_{t+\tau}$ is the earnings. When $Y_{t+\tau}$ is after-tax earnings, "clean surplus accounting" implies that the net change of book equity is equal to the retained earnings:

$$dB_{t+\tau} = Y_{t+\tau} - D_{t+\tau} \tag{2}$$

where $dB_{t+\tau} \equiv B_{t+\tau} - B_{t+\tau-1}$ is the net change of book equity in year $t+\tau$.

Replacing $D_{t+\tau}$ with $Y_{t+\tau} - dB_{t+\tau}$ and dividing both sides of (1) by the current book equity B_t, the dividend discount model becomes

$$\frac{P_t}{B_t} = \sum_{\tau=1}^{\infty} \mathbb{E}\left(\frac{Y_{t+\tau}}{B_t} - \frac{dB_{t+\tau}}{B_t}\right) \Big/ (1+r)^{\tau}. \tag{3}$$

There are four variables in the above equation in addition to the expected return: the market price, P_t, the book value, B_t, (and the market-to-book ratio P_t/B_t), expected profitability $\mathbb{E}(Y_{t+\tau})/B_t$, and expected equity growth $\mathbb{E}(dB_{t+\tau})/B_t$. The present value relation suggests that holding the other variables constant, firms with 1) higher market-to-book ratios should have higher expected returns; 2) higher expected profitability should have higher expected returns; 3) higher expected growth in book value should have lower expected returns; and 4) higher market value (holding book value constant) should have lower expected returns.

Fama and French are careful to point out that the present value relation is an accounting identity so that the ability of market-to-book, profitability, asset growth, and size do not necessarily constitute asset pricing anomalies. However, properly controlling for these variables, the ability of other variables (such as momentum) to explain returns constitutes evidence of pricing anomalies.

Any variables that help explain expected profitability and expected asset growth might be candidates for the study of (3) and constructing candidate pricing factors, including lagged values of proxies for profitability and asset growth. These include lagged values of profitability and asset growth, which form the basis of some FF5 factors. The factors in FF5 model are the excess return on the market portfolio (clearly motivated by the Capital Asset Pricing Model (CAPM)) and four long-short portfolios, all of which can be motivated by the present value relation, (3). SMB (Small Minus Big) is a portfolio that is long small capitalization stocks and short large capitalization stocks. HML (High Minus

Low) is a portfolio that is long high book-to-market stocks and short low book-to-market stocks. RMW (Robust Minus Weak) is a portfolio that is long high expected profitability stocks and short low expected profitability stocks. Finally, CMA (Conservative Minus Aggressive) is a portfolio that is long low expected asset growth stocks and short high expected asset growth stocks. Holding, B_t, $\mathbb{E}(Y_{t+\tau})/B_t$, and $\mathbb{E}(dB_{t+\tau})/B_t$, the present value relation suggests that SMB should have a positive expected return, as argued by Berk (1995). By similar logic, HML, RMW, and CMA should have positive average returns.

2.2 Proxies for Profitability

Any study of equation (3) and implementation of the FF5 model requires empirical measures of expected profitability and expected asset growth (size and book-to-market are observable). Our focus is on the profitability factor. Fama and French (2006) use net income divided by the book equity (which we call bottom-line earnings hereafter) as a proxy for expected profitability. They find a significant relation between expected returns and current profitability. However, they find no significant relation between expected returns and expected profitability measured by fitted values from cross-sectional regression of future profitability on current profitability and additional variables. They offer two possible explanations for the puzzling results: (1) the measurement error in the first-stage cross-section profitability regressions and (2) the collinearity problem since the book-to-market ratio enters both the first-stage and the second-stage regressions.

Measurement error in profitability has been discussed extensively in the literature. Several alternative measures of profitability have been suggested as alternatives to bottom-line earnings since they may be less susceptible to manipulation by management. Using Compustat definitions, we move from the top-line of the income statement (revenue) to bottom-line earnings, as follows:[2]

$$
\begin{aligned}
&\text{Revenue (REVT)}\\
-&\text{Cost of goods sold (COGS)} = \text{Gross profit (GP)}\\
-&\text{Selling, general and administrative expenses (XSGA)} = \text{Operating income before depreciation (OIBDP)}\\
-&\text{Depreciation and amortization (DP)} = \text{Operating income after depreciation (OIADP)}\\
-&\text{Interest and related expense (XINT)}\\
+&\text{Special items (SPI)}\\
+&\text{Non-operating income (NOPI)} = \text{Pretax income (PI)}\\
-&\text{Income taxes (TXT)}\\
-&\text{Minority interest (MII)} = \text{Income before extraordinary items (IB)}
\end{aligned}
$$

To avoid the largely non-recurring items, Li and Mohanram (2014) measure earnings as income before extraordinary items minus special items, IB-SPI. Novy-Marx (2013) uses

[2] All the items have quarterly counterparts.

gross profit (GP) as the proxy for profitability and argues that *"The farther down the income statement one goes, the more polluted profitability measures become, and the less related they are to true economic profitability."* To measure earnings in their profitability factor, Fama and French (2015) subtract the selling, general, and administrative expenses and interest expenses from the gross profit (EBTDA = GP-XSGA-XINT).

We assess the ability of various earnings measures to predict future cash flows (years $t+1$ through $t+10$) by regressing future cash flows (measured by the Compustat Cash Flow from Operations, OANCF) on current earnings metrics (Table 1). The cross-sectional regressions are for 50 portfolios sorted by 5 industry groups and the Fama-French measure of profitability. Of the measures that do not adjust for intangibles, we use (moving from bottom-line to top-line measures) net income (Fama and French (2006)), earnings before extraordinary items (Li and Mohanram (2014)), earnings before taxes (EBT), earnings before taxes, depreciation, and amortization (EBTDA, Fama and French (2015)), gross profit, (Novy-Marx (2013)), and revenue. Since the Compustat variable for cash flows from operations (OANCF) is available starting in 1988, our predictive regressions are for the period 1988-2021. For all horizons, the average R^2 values of the regression increase from net income through the Fama and French (2015) EBTDA measure and then decline as we move to more top-line measures. The results for a subset of these measures, net income; earnings before taxes, depreciation, and amortization (EBTDA); and gross profit, are reported in Table 1. Across all horizons, the gross profit measure of profitability has higher average values of R^2 than net income. Additionally, the FF5 measure of profitability has higher values of R^2 than gross profitability Because of this result, we follow Fama and French (2015) and measure unadjusted profitability using EBTDA instead of the more bottom-line measure used in Fama and French (2006) or the pore top-line measure of gross profitability in Novy-Marx (2013).

2.3 FF5 Factors and Investments in Intangibles

Analysts have long recognized that the accounting treatment of investments in physical assets and intangible assets has been asymmetric (for example, Bloomberg (1934), Graham et al. (1962), Graham (1973), Greenwald et al. (2001)).[3] Investments in physical capital are capitalized and depreciated over time while investments in intangible capital,

[3] In the 2010 CEO's letter to Amazon shareholders, Jeffrey Bezos emphases that Amazon's continuing investment in intangible assets has played a determinant role in advancing the company and the ultimate payoff to shareholders: *"The advances in data management developed by Amazon engineers have been the starting point for the architectures underneath the cloud storage and data management services offered by Amazon Web Services (AWS). ... in my opinion, these techniques are not idly pursued – they lead directly to free cash flow."*

In the 1997 annual report Bezos illustrates his principle of investing: *"We will make bold rather than timid investment decisions where we see a sufficient probability of gaining market leadership advantages. ... When forced to choose between optimizing the appearance of our GAAP accounting and maximizing the present value of future cash flows, we'll take the cash flows."*

which build intellectual property, brands, and organizational capital (which show up in R&D and SG&A), are expensed immediately (unless they are purchased intangibles). Treating investments in intangible capital similarly to physical capital has implications for the measurement of earnings and book value and influences three of the variables in the present value relation (3), P_t/B_t, $dB_{t+\tau}/B_t$, and $Y_{t+\tau}/B_t$.

Similarly, incorporating investment in intangible capital might affect the way we define three of the factors in the FF5 model, RMW, HML, and CMA. Eisfeldt et al. (2022) develop an intangibles-adjusted version of HML and argue it has significant advantages, in terms of pricing the cross-section, relative to the original HML. Wang (2023) proposes an intangibles-adjusted asset growth factor, similar to CMA, and argues it outperforms the original CMA factor.

Fundamental analysts often adjust earnings for investment in intangibles, e.g., Greenwald et al. (2001) add 25% of both the R&D and SG&A expenses back to Intel's earnings to project its future earnings (Table 7.7 in the chapter "Inside Intel" in Greenwald et al. (2001)). Even earlier, Graham (1973) highlights the danger of earnings "smoothing" and the choice between fully charging off R&D costs or amortizing them.[4]

The academic literature has advocated several approaches to account for investment in intangibles when explaining equity valuation. For example, Lev and Sougiannis (1996) capitalize R&D expense and adjust bottom-line earnings for R&D expense. Peters and Taylor (2017) and Crouzet and Eberly (2022) consider all R&D expense as investment in knowledge capital and 30% of the SG&A expense net of R&D expense as the investment in organizational capital.[5] Belo et al. (2022) treat R&D and advertising expenses and the change in the number of employees as separate components of investment in intangibles. Eisfeldt and Papanikolaou (2013) capitalize the sum of R&D and SG&A expenses as measure of firms' organizational capital. Eisfeldt et al. (2022) suggest: *"Without better estimates of the fraction of SG&A spending that is investment in intangible assets, ... it is best to use 100% of SG&A ... to avoid introducing noise."*

We wish to stay as close to the FF5 model as possible. Since the FF5 RMW factor corresponds to earnings before taxes, depreciation, and amortization (EBTDA), our corresponding adjustment to earnings is the equivalent of EBTDA with intangibles investment. That is, we add back the expense of intangibles investment to the FF5 version of earnings. We focus on three alternative measures of intangibles investment a) one that adds to earnings R&D expenses only as in Lev and Sougiannis (1996) (noted with a superscript RD below), b) one that adds to earnings R&D expenses and 30% of SG&A

[4]In chapter 12 "Things to consider about per-share earnings": *"Still another factor, important at times, is the choice between charging off research and development costs in the year they are incurred or amortizing them over a period of years".*

[5]The SG&A expense has to be net of R&D because it is the practice of Compustat to include the R&D expenses in the SG&A expenses and record it as the data item XSGA unless it is included in the cost of goods sold by the company.

expenses as in Peters and Taylor (2017) and Crouzet and Eberly (2022) (noted with a superscript RD+.3SG below), and c) one that adds to earnings R&D expenses and 100% of SG&A expenses as in Eisfeldt et al. (2022) (noted with a superscript RD+SG below). The first of these measures was included in early versions of Fama and French (2018) but not in the published version.

We first study the ability of intangibles-adjusted profitability measures to predict future cash flows, a requirement for them to be proxies for expected cash flows in (3). There are fewer individual firms available at longer horizons due to firms entering and dropping out of the sample (the number of firms with cash flow data at a horizon of ten years is 60% of those with data at a horizon of one year). Because of this, we run regressions using 50 portfolios sorted by the FF5 profitability measure ($Y(\mathrm{FF})/B_t$) and Fama-French industry classification (details on portfolio formation are given below). Portfolios are rebalanced as firms exit the sample. In the profitability regressions (Table 1), $Y(\mathrm{FF})/B_t$ is the single best predictor of profitability among the unadjusted profitability measures. A subset of these is reported in the table. The gross profitability measure of Novy-Marx (2013) has a higher average R^2, relative to net income (IB), for all horizons (Panels A and B of the table). The difference in average R^2 values is significant for horizons 3 and 10 (t-statistics for the difference are reported in Panel B). The EBITDA measure of Fama and French (2015) has a higher average R^2, relative to gross profitability, for all horizons (Panels A and B of the table) and the differences are significant at horizons of 1 through 9 years.

When we include the intangibles adjustments (Panels C and D) $Y(\mathrm{RD})/B_t$ has less predictive power (measured by R_2) than $Y(\mathrm{FF})/B_t$ at horizons of 1 to 2 years and greater at horizons of 3 to 10 years, with horizons 5 to 10 being significantly higher. This makes intuitive sense since R&D expenditures may take some time to impact cash flows. $Y(\mathrm{RD}+.3\mathrm{SG})/B_t$ has lower predictive power (measured by R_2) than $Y(\mathrm{RD})/B_t$ at horizons of one year through four years and higher average R^2s for longer horizons (significant for years eight through 10). $Y(\mathrm{RD}+\mathrm{SG})/B_t$ has lower predictive power (measured by average R_2) than $Y(\mathrm{RD} + .3\,\mathrm{SG})/B_t$ at all horizons.

Figure 1 plots the time series of the difference in R^2 between the profitability measures $Y(\mathrm{NM})/B$ versus $Y(\mathrm{IB})/B$, $Y(\mathrm{FF})/B$ versus $Y(\mathrm{NM})/B$, and $Y(\mathrm{RD})/B$ versus $Y(\mathrm{FF})/B$ across three different forecasting horizons $\tau = 1, 5, 10$. Though $Y(\mathrm{NM})/B$ does not always outperform $Y(\mathrm{IB})/B$ in forecasting future profitability, $Y(\mathrm{FF})/B$ clearly performs better than $Y(\mathrm{NM})/B$. $Y(\mathrm{RD})/B$ forecasts future profitability better than $Y(\mathrm{FF})/B$ in the longer-horizon (5- and 10-year ahead) but worse in the short-horizon (1-year ahead).

Among measures that do not adjust for intangibles, the results in Table 1 provide empirical support for the ability of the earnings measure used in Fama and French (2015) to predict future cash flows (Panels A and B). The results also provide support for increased predictive power at longer horizons when using intangibles-adjusted profitability

(Panels C and D).

Table 2 performs a similar portfolio-level cross-sectional regression in which the FF5 profitability measure, RD_t/B_t, and SGA_t/B_t are included as explanatory variables together. The Fama-French profitability measure is significant at all horizons. The R&D adjustment is significantly negative at the one-year horizons, but positive at all others (with its slope coefficients and t-statistics generally increasing with horizon). The SG&A adjustment is positive at all horizons (also with its slope coefficients and t-statistics generally increasing with horizon). The results in Table 2 confirm the intuition from Table 1 that R&D takes some time to influence future cash flows. SG&A kicks in by year 1 and also grows in importance at longer horizons, consistent with portions of SG&A being investments in intangible capital that payoff in future cash flows. Thus the Fama and French (2015) EBTDA profitability measure as well as measures of EBTDA that adjust for investment in intangibles seem to pass the test of being able to predict future cash flows.

We report the predictive regression results for the cash dividends in Tables 3 and 4 and Figure 2. In these tables, the dependent variables are future cash dividends in year $t + \tau$ divided by the book equity in year t and the independent variables remain the same as in the operating cash flows regressions. The portfolio-level regressions largely overcome the problem that many firms do not pay dividends. Differences between the results of forecasting cash dividends and the results of forecasting operating cash flows are that dividends are generally less forecastable than operating cash flows and $Y(\mathrm{NM})/B$ does not appear better than $Y(\mathrm{IB})/B$ when forecasting cash dividends while $Y(\mathrm{RD})/B$ outperforms $Y(\mathrm{FF})/B$ even in the short horizon (1-year ahead). The importance of the intangibles adjustment increases (both in magnitude and t-statistics of the slope coefficients) as the forecasting horizon grows.

Now that the profitability measures seem to pass the minimum requirement for being useful for being tied to future expected returns (or the ICC), that is having the ability to forecast future cash flows, we can move on to asking whether factors derived from them are useful in pricing the cross-section of asset returns. Before we do that, we provide a more detailed discussion of the underlying data.

3 R&D Disclosure and Compustat Data

The Financial Accounting Standards Board Statement number 2 (FASBS2, FASB (1974)) mandated all material R&D costs encompassed by the statement to be charged to expense when incurred and the statement became effective for fiscal years beginning on or after January 1, 1975. The accounting treatment of R&D costs varies across firms before 1975. Moreover, FASB (1974) required *"... research and development costs be charged to expense when incurred shall be applied retroactively by prior period adjustment, ...*

they shall be restated to reflect the prior period adjustment. The prior period adjustment shall recognize any related income tax effect." Therefore, the recorded R&D expense in restated financial statements prior to the fiscal year of 1975 may not have been available to the public at the time of original publication. Therefore, the performance of portfolios constructed using those restated data may contain look-ahead biases. For this reason, we restrict our analysis to the sample period from 1975 to 2021 when using any accounting information.

Moreover, even post-FASB2, not all firms report (positive) R&D expenses. In the calendar year 1975 40% of firms report R&D expenses, experiencing a slight decline (but always above 35%) before 2000, and then gradually increasing to 52% in 2021. Koh and Reeb (2015) state that the materiality threshold is typically defined as amounts exceeding one percent of sales. Since R&D expense is an item of mandatory disclosure, blank or empty R&D fields should imply zero or immaterial R&D expense. However, Koh and Reeb (2015) find that 10.5% of the missing R&D firms file and receive patents, which is 14 times greater than zero R&D firms. Koh and Reeb (2015), therefore, argue that a blank R&D field could also represent a firm's conscious decision not to separate R&D expense from other reported expenses, rather than implying zero R&D expense. Without a better approach to impute missing R&D expense, we set it as zero if missing in this paper. The consequence is that our R&D measure is subject to an errors in variables (EIV) problem. The EIV issue should weaken our results and lead to a bias toward failing to find a significant effect of the intangibles adjustment.

For the firms reporting R&D expense, the expense can be shown either as (1) a separate line item on the income statement or (2) separately stated in notes to their financial statements.[6] It is Compustat's standard to collect R&D expense (Compustat data item XRD) from either line items or notes to financial statements. In the Appendix, we illustrate this practice using Apple and Illinois Tool Works as examples.

Since we construct both annual- and monthly-rebalanced profitability factors using annual and quarterly filings data respectively, we also examine the R&D data recorded in quarterly Computat. First, we find that no firms report R&D in the quarterly filings (quarterly Compustat data item XRDQ) before 1989 and only on average 41% of the firms have XRDQ record in the quarterly Compustat data between 1989 and 2021.[7] Moreover, among all firms that have XRDQ records, there are about 25% of the firms report only once a year and the amount is the same as in the XRD recorded in annual Compustat data. We verify that firms reporting their R&D expense as a separate line item in the income statement typically have multiple XRDQ records in a year while the firms integrating R&D expense into other line items and only reporting in the notes to

[6] `https://www.irs.gov/businesses/corporations/faqs-irc-41-qres-and-asc-730-lbi-directive`.

[7] 98% of the firms reporting annual R&D expense report quarterly R&D expense.

financial statement typically have only one XRDQ record each year.[8] From 1989 and 2021, the aggregate R&D expense for all firms with more than one XRDQ record is 8.8% of their total sales which is a much larger fraction relative to the 1.8% for firms with only one XRDQ record.

As an example, consider General Electric (GE). GE reports R&D expenses only in the notes to financial statements in the quarterly and annual filings before the fiscal year ending in December 31, 2019. In their notes to financial statements in the 2019 10-K, GE explains "R&D expenses are classified in cost of goods and services sold in our consolidated Statement of Earnings (Loss)." After 2019, GE changes this practice and reports R&D expense as a line item in both quarterly and annual filings. This change is also reflected in the quarterly Compustat data.

GE integrates R&D expense in the cost of goods sold (Compustat data item COGS), though most firms include R&D expense in the selling, general and administrative expense (Compustat data item XSGA). Compustat does not provide any information to identify with which item the firm integrates R&D expense. This practice has significant implications for our analysis as we need to separate SG&A and R&D expenses.

We follow the procedure in Peters and Taylor (2017) and separate the SG&A and R&D expenses. We set SG&A equal to XSGA minus XRD minus RDIP[9] with an additional screen: when XRD exceeds XSGA but is less than COGS, or when XSGA is missing, SG&A is set equal to XSGA with no further adjustments or zero if XSGA is missing. We recognize that this approach is not flawless as it does not work for firms with higher SG&A than R&D but integrating R&D in the cost of goods sold. For example, GE in 2019 10-K reports SG&A expense of $13,949 M and R&D expense of $3,115 M (modified to $3,118 as a line item in 2020 10-K, but remained as $3,115 in Compustat) with zero acquired R&D integrated into the cost of goods and services sold. We denote firms' ex-R&D values of SG&A as SGA and use XSGA to refer to the cum-R&D Compustat data item.

Panel A and B in Figure 3 report R&D expense as fractions of the total sales and book equity, respectively, for all firms for five industries (using the Fama and French industry definitions). R&D expense relative to sales and book equity of all firms are stable over the entire sample but exhibits heterogeneity across industries with HighTech and Health industries being the most R&D intensive industries and experiencing rapid growth over the sample period.

[8] 64% of the firms reporting quarterly R&D expense report four times a year.

[9] RDIP represents the portion of R&D considered to be "purchased" and written off immediately upon acquisition if the R&D items are deemed not to have an alternative use.

4 Data

Monthly stock returns, dividends, stock prices, and numbers of common shares outstanding are from the Center for Research in Security Prices (CRSP), and annual and quarterly accounting information is from Compustat. To avoid look-ahead bias in using R&D data, the accounting information is from January 1975 to December 2021. To guard against issues with backfilling we remove the first two years of Compustat observations for each firm. We include all firms with share codes 10 and 11 and traded on NYSE, Amex, and Nasdaq. Table 15 contains a comprehensive list of variable definitions.

The profitability of firms is measured as the earnings divided by total book equity. We define annual total book equity as Compustat data item CEQ[10] and quarterly book equity following Hou et al. (2015).[11] As mentioned above, we perform analyses using several measures of earnings but only report a subset of results to conserve space. Our earnings measures include (1) total revenue, $Y(\text{REVT})$, (2) gross profit, $Y(\text{NM})$ following Novy-Marx (2013), (3) gross profit minus interest expense, $Y(\text{RD}+\text{SG})$ following definition in Eisfeldt et al. (2022) for investment in intangibles, (4) gross profit minus interest expense minus 70% of SG&A expense, $Y(\text{RD} + .3\,\text{SG})$ following the definition in Peters and Taylor (2017) for investment in intangibles, (5) gross profit minus interest expense minus 100% of SG&A expense, $Y(\text{RD})$, (6) gross profit minus interest expense minus 100% of SG&A expense minus 100% of R&D expense, $Y(\text{FF})$ following definition in Fama and French (2015) for profitability, (7) gross profit minus interest expense minus 100% of SG&A expense minus 100% of R&D expense minus 100% depreciation, $Y(\text{EBT})$, (8) income before extraordinary items minus special items, $Y(\text{SPI})$, and (9) income before extraordinary items, $Y(\text{IB})$.

Due to the quarterly R&D data issues described in the earlier section, we focus on the annual-rebalanced profitability factors using annual Compustat data in most of our analyses. To be included in the profitability forecasting regression for $t+\tau$ where $\tau = 1, 2, ..., 10$ in Tables 1,2, 3, and 4, a firm must have Compustat data for calendar year t on book equity and revenue or income before extraordinary items. For the firms reporting accounting information more than once in a given calendar year, potentially due to the change of the fiscal year-end month, we use the most recently released information. In addition, to ensure that the accounting variables are known at the month of portfolio formation and return regressions, we follow Fama and French (2006) and include a firm

[10]The book equity computation in Fama and French (2018) has broader coverage than the Compustat data item CEQ before 1968. Since our sample period starts in 1975, using CEQ or computed book equity does not alter our results.

[11]Quarterly book equity is shareholders' equity, plus balance-sheet deferred taxes and investment tax credit (item TXDITCQ) if available, minus the book value of the preferred stock. Depending on availability, we use stockholders' equity (item SEQQ), or common equity (item CEQQ) plus the carrying value of the preferred stock (item PSTKQ), or total assets (item ATQ) minus total liabilities (item LTQ) in that order as shareholders' equity. We use redemption value (item PSTKRQ) if available, or carrying value for the book value of the preferred stock.

in the portfolio and return regression for July of t+1 to June of t+2 if it has available book equity, revenue or income before extraordinary items in the calendar year t. This procedure guarantees an at least six-month period for the information to be released.[12] Figure 4 depicts the timeline for the empirical implementation.

We exclude firms with negative book equity, missing share price, shares outstanding when they are used in the explanatory variables (that is, at time t). Following Fama and French (2006), we remove firms with total assets less than \$25 million or book equity less than \$12.5 million in year t to avoid the influence of small firms.

Other data, including various competing asset pricing factors, are taken from the authors' websites. All of the series are available throughout our sample period with the exception of the factors from Stambaugh and Yuan (2017), which are available through December 2016. In particular, the factors in the Fama-French five factor model (FF5) and the momentum (UMD and UMD(d)) are taken from Ken French's data library (`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`, last visited August 10, 2022). The same data library is also the source of the five industry portfolios that we use in the ROE prediction regressions. The factors in the Hou-Xue-Zhang four factor model (HXZ4) and their expected growth factor (R_EG) are from `https://global-q.org/factors.html`, last visited August 10, 2022. The factors in the Daniel et al. (2020) behavioral three factor model (DHS3) are from `https://sites.google.com/view/linsunhom`, last visited August 10, 2022. The factors in the Stambaugh and Yuan (2017) four factor model (SY4) are from `https://finance.wharton.upenn.edu/~stambaug/`, last visited August 10, 2022. The intangibles-adjusted HML factor, HML(EKP) is from Eisfeldt et al. (2022) and available from `https://github.com/edwardtkim/intangiblevalue,`, last visited August 10, 2022. The monthly rebalanced version of HML, HML(Devil) of Asness and Frazzini (2013) is from AQR.[13]

Since firm-level cross-section predictive regressions with horizons of one to ten years are subject to severe survivorship biases, we forecast the future operating cash flows of fifty managed portfolios. Table 16 in the Appendix reports the number of firms that remain available after τ years. 11%, 21%, and 41% of firms drop out of the sample after 3, 5, and 10 years respectively. To deal with the severe survivorship bias in the firm-level profitability regressions, we assign firms into 5-by-10 industry and $Y(\mathrm{FF})/B$ groups. The portfolios are rebalanced at the end of each month in the next ten years to ensure that the portfolio weights on each firm are proportional to their respective market size. The portfolio-level profitability regressions take the portfolio-level profitability (earnings in

[12]This conservative implementation could include some stale accounting information. For example, for a firm with a fiscal year-end in February 2021 with announced earnings in April 2021, the accounting information is already fourteen months ago when we first use it in the portfolio and return regression analysis in July 2022. This will tend to bias against finding significant results.

[13]`https://www.aqr.com/Insights/Datasets/The-Devil-in-HMLs-Details-Factors-Monthly`, last visited March 1, 2023.

year $t+\tau$ divided by book equity in year t) in year $t+\tau$ as the dependent variables and the current profitability, and other variables, in year t as the independent variables.

In the regressions reported in Tables 1, 2, 3, and 4) the earnings of portfolio k in the year t after its formation as its net worth times the value-weighted earnings-to-market ratio, i.e.,

$$\mathrm{Y}_{k,t+\tau} = \underbrace{w_{k,t+\tau}}_{\text{portfolio net worth}} \sum_{i \in k} \underbrace{\frac{M_{i,t+\tau}}{\sum_{j \in k} M_{j,t+\tau}}}_{\text{weight on firm i}} \frac{\mathrm{Y}_{i,t+\tau}}{M_{i,t+\tau}} \qquad \text{where } \tau = \{1, 2, \cdots, 10\} \tag{4}$$

where $w_{k,t+\tau} = \prod_{s=\{1,2,\cdots,\tau\times 12\}} \left(1 + \sum_{i \in k} \frac{M_{i,s-1}}{\sum_{j \in k} M_{j,s-1}} r_{i,s}\right)$ is one plus the cumulative return of the portfolio from its formation to year t where $r_{i,s}$ is the monthly return of firm i in month s after the portfolio formation.

Following Fama and French (2015), we leave at least six months for the firms to announce earnings. The portfolios are formed by the June-end in year t based on the accounting information in calendar year t-1 from Compustat. All the firms are sorted independently into 5-by-10 industry and $Y(\mathrm{FF})/B$ groups and we keep tracking the portfolios' returns, earnings, and cash flows for the next five years after their formations based on the rebalancing scheme described earlier.

5 Profitability Factors with an R&D Adjustment

Though FF5 contains only annually rebalanced factors, many other factor models are based on monthly-rebalanced portfolios. Therefore, we also construct monthly-rebalanced profitability factors using quarterly accounting data. While our focus is on annual rebalancing, we also report results with monthly rebalancing

5.1 The performance of Profitability in Pricing Assets

Table 5 reports mean returns and (t-statistics) of portfolios ranked independently on market capitalization (based on the median market equity for NYSE firms) and various measures of profitability (three groups based on the 30^{th} and 70^{th} NYSE percentiles of profitability). The table also provides the difference in average returns in high- versus low-profitability firms for each size category. The last two columns show the mean profitability factor return spread and its alpha, relative to a one-factor model using the FF RMW factor. The return spreads are generally significant for all unadjusted profitability measures with t-statistics increasing from bottom-line profitability to top-line, except for revenue, which is less significant than gross profitability.

However, the largest average return spreads and t-statistics are for the three intangibles-adjusted measures, followed by the gross profitability measure of Novy-Marx (2013). Similarly, the only measures with significantly positive alpha versus RMW are the intangibles-adjusted measures and gross profitability.

Panel A and B in Figure 5 present the cumulative return of three selected annual- and monthly-rebalanced profitability factors when investing \$1 at June-end in 1976. For both the annual- and monthly-rebalanced factors, the R&D-adjusted profitability factors outperform the other two, especially during the post-2000 subsample. However, we show, in the tests below, that the performance of intangibles-adjusted profitability factors is not solely a post-2000 phenomenon.

It is common to test alternative pricing models by choosing some test assets and asking whether one model provides lower estimated abnormal returns (alphas) than a competing model. The alphas are typically measured as the intercept in a time series regression of the excess return on each test asset on the factor returns (or factor-mimicking portfolios). Barillas and Shanken (2017) and Barillas et al. (2020) show that, when factors are traded portfolios, alternative models can be compared (in terms of which one leads to higher squared Sharpe ratios) by asking if each model can price the factors in the alternative model. The left-hand side "test assets" drop out in this case (they also show that looking at test asset alphas may be misleading).

We study the mispricing (or alphas) of the profitability factors studied in Table 5 by regressing their excess returns on eight important multi-factors asset pricing models in the literature. The profitability alphas are reported in Table 6. The models that we use are (1) the FF5 model (Fama and French (2015)); (2) the FF6 model which adds the momentum factor, UMD, from Carhart (1997) to FF5; (3) FF6/Devil which is a version of the FF6 model in which the value factor, HML is replaced with a version that is updated on a monthly basis (Asness and Frazzini (2013)); (4) FF6/EKP which is a version of the FF6 model in which the value factor, HML is replaced with the intangibles-adjusted value factor from Eisfeldt et al. (2022); 5) the four-factor model of Hou et al. (2015); (6) the five-factor model from Hou et al. (2021); (7) the long- and short-horizon-mispricing thee-factor model of Daniel et al. (2020); and (8) the mispricing four-factor model of Stambaugh and Yuan (2017).

Our main results are the estimated values of α, $\hat{\alpha}$, for the array of profitability factors, and their associated t-statistics, relative to the eight competing models, which are reported in Table 6. For each measure of profitability, we do 2-by-3 sorts on size and profitability. For the annually rebalanced factors, the size sort breakpoint is the median size for NYSE firms. The profitability-sort breakpoints ate the 30^{th} and 70^{th} NYSE percentiles of profitability. Using the Fama-French terminology "Small Robust" is the value-weighted portfolio of small-cap firms with high profitability and "Small Weak" is the portfolio of small-cap firms with low profitability (with equivalents of "Big Robust" and "Big Weak" for large-cap firms). Our factors are constructed as 1/2 (Small Robust + Big Robust) - 1/2 (Small Weak + Big Weak).

The first three columns of results show the profitability factors moving from bottom-line to top-line measures: net income (IB), income before special items (SPI), earnings before taxes (EBT), and earnings before taxes depreciation and amortization (EBTDA), used in Fama and French (2015) (FF), have insignificant values of $\hat{\alpha}$. This is not surprising since our predictive regressions for operating cash show that IB, SPI, and EBT all have less predictive power than (EBTDA). The results for FF in column 4 are also not surprising since the RMW is one of the factors in the first 4 models. The fact that $\hat{\alpha}$ for our version of the FF5 profitability factor is not identically zero is just an indication that our construction of RMW differs slightly from the version of RMW on Ken French's data library, which is the source of RMW on the LHS of the regression.

The next three columns include versions of our intangibles-adjusted profitability factors. The factor $r_{\text{ROE}}^{\text{RD}}$ uses the FF5 measure of profitability plus R&D expenditures. The factor $r_{\text{ROE}}^{\text{RD+.3SG}}$ also adds 30% of SG&A and $r_{\text{ROE}}^{\text{RD+SG}}$ also adds 100% of SG&A. For $r_{\text{ROE}}^{\text{RD}}$ and $r_{\text{ROE}}^{\text{RD+.3SG}}$ the values of $\hat{\alpha}$ are positive and the largest across all profitability measures and statistically are significant across all eight benchmark models. That is, adding these factors to the benchmark factors would significantly increase the squared Sharpe ratio of the portfolio space spanned by the factors. The t-statistics range from 2.26 to 7.25 and exceed 3.26 for all models except the HMXZ5 model. The factor which adds R&D and all of SG&A, $r_{\text{ROE}}^{\text{RD+SG}}$, and the gross profitability measure, $r_{\text{ROE}}^{\text{NM}}$, have slightly smaller alphas and have significantly positive $\hat{\alpha}$ for all models except HMXZ5 and SY4. The six significant alphas for $r_{\text{ROE}}^{\text{RD+SG}}$ have t-statistics ranging from 2.77 to 4.40 while the significant alphas for $r_{\text{ROE}}^{\text{NM}}$ have t-statistics ranging from 1.97 to 3.51. The results for $r_{\text{ROE}}^{\text{RD+SG}}$ and $r_{\text{ROE}}^{\text{NM}}$ are close since profitability, when adding back both R&D and SG&A, differs from gross profitability only by interest expense.

Our main focus is on factor risk premia and the results of Table 6 are consistent with a significant premium to the intangibles-adjusted profitability factors. For some factors, it can be the case that abnormal returns come mainly from the short side of the factor portfolio. A practitioner attempting to implement a smart beta portfolio based on a given factor might have concerns that there are large asset borrowing costs on the short side. We study which side is the source of abnormal returns by looking at the values of $\hat{\alpha}$ for the long and short sides separately.

Table 7 shows our results. Panel A of the table shows the long side for annually-rebalanced factors. $r_{\text{ROE}}^{\text{RD}}$ and $r_{\text{ROE}}^{\text{RD+.3SG}}$ have significant long-side alphas for each benchmark model, with one exception. $r_{\text{ROE}}^{\text{RD+.3SG}}$ has an insignificant long-side alpha relative to SY4. Panel C shows the short side for annually-rebalanced factors. The values of $\hat{\alpha}$ for $r_{\text{ROE}}^{\text{RD}}$ and $r_{\text{ROE}}^{\text{RD+.3SG}}$ are statistically significant for all models, but two (HMXZ5 and DHS3). In particular, while the long-side alphas were not very strong against SY4, the short-side alphas are. Similarly, while the short-side alphas were not very strong against HMXZ5 and DHS3, the long-side alphas are.

The takeaway here is that the intangibles-adjusted profitability measures add explanatory power for the cross-section of assets beyond that captured by the models studied here. Against the nine benchmark models $r^{\mathrm{RD}}_{\mathrm{ROE}}$ and $r^{\mathrm{RD+.3SG}}_{\mathrm{ROE}}$ have abnormal returns coming from both sides of the position for five models, from the long side for two models, and from the short side for one model. Therefore, the significant intangibles adjusted factor alphas are not solely coming from the short side.

5.2 What Fraction of SG&A?

The evidence so far suggests that adjusting earnings for R&D expense is important, however, it remains unclear how much of SG&A should be added back to earnings. The results in Table 5 suggest that adding 30% of SG&A can improve the explanatory power of the intangibles-adjusted factors. Alpha increases in size, versus only adjusting for R&D, for seven of the eight models, although the t-statistics only increase for two of the eight models. Adding 100% of SG&A always reduces both the magnitude of the estimated αs and the t-statistics. Table 8 repeats the type of analysis of Table 5, for profitability factors that adjust the fraction of SG&A added back from 0% to 100% in increments of 10%. The highest levels of α are always for adding a fraction of 20% (with two ties to 30% and one tie to 0%-10%-30%). The t-statistics for the estimated α is the largest at 20% for seven models and at 10% for one model. This suggests that R&D expense plays a crucial role in strengthening the relationship between profitability measure and expected returns while a fraction of SG&A expense is important. This is consistent with our predictive regressions in Tables 1 and 3.

We suspect that this may be due to industry variation in the fraction of SG&A that represents the investment in intangibles. We leave this question for future study and will focus on the R&D-adjusted factor for the remainder of the paper.

5.3 Subsample Performance

The cumulative returns shown in Figure 5, Panel A, suggest a potential regime shift in the returns to both the FF5 factor, RMW, and the intangibles-adjusted profitability factors, around the year 2000. We wish to determine if the significance of these factors is a post-2000 phenomenon. We first report the time-series average monthly percent returns and t-statistics of several earnings-related factors/anomalies for the full sample and two subperiods (using December 1999 as the breakpoint) in Table 9. All the factors considered here are significant in the full sample and in the first subsample. All of the factor t-statistics fall in the second subsample, except for intangibles-adjusted profitability. The quarterly ROE factor from Hou et al. (2015) (R_ROE) and both versions of the momentum factor become insignificant in the second subsample. The two R&D-adjusted profitability factors, $r^{\mathrm{RD}}_{\mathrm{ROE}}$ and $r^{\mathrm{RD}}_{\mathrm{ROEQ}}$, have the highest t-statistics in the second subsample. Importantly, the two R&D-adjusted profitability factors are not solely post-2000 phenomena, with both being statistically significant in each subperiod.

5.4 Relations with Profitability-Related Factors

MAny existing factor models include a profitability or earnings factor, e.g., RMW in FF5, R_ROE in HXZ4, PEAD in DHS3, and PERF in SY4. The (price) momentum factor is often related to the earnings momentum phenomenon. For example, Chan et al. (1996) find that price momentum is largely due to the market's underreaction to earnings-related information. Even though these factors are all connected to the earnings or post-earnings market reactions, they manifest very different properties of some asset price anomalies. For example, the monthly-rebalanced PEAD and R_ROE factors can subsume the (price) momentum factor while Fama and French (2016) show that a momentum factor is required to take away the significance of the momentum phenomenon though they have an annual-rebalanced profitability factor, RMW, in FF5.

We examine the relation of the R&D-adjusted profitability factor with profitability factors in existing models and three other potentially related factors, gross profitability ($r^{\text{NM}}_{\text{ROA}}$), gross profit divided by total assets, by Novy-Marx (2013), operating leverage ($r^{\text{NM}}_{\text{OL}}$) by Novy-Marx (2011), and an intangibles-adjusted value factor (HML(EKP)) by Eisfeldt et al. (2022). Panel A in Table 10 reports the univariate regression results of the R&D-adjusted profitability factor on the profitability-related factors. The alphas of $r^{\text{RD}}_{\text{ROE}}$ are significant against all included factors. As independent variables, RMW, R_ROE, and PERF have the strongest explanatory power for $r^{\text{RD}}_{\text{ROE}}$. Two factors, PEAD and HML(EKP) are unrelated with $r^{\text{RD}}_{\text{ROE}}$ with insignificant (or negative) slope coeffiсiMany and very low R^2s.

Reverse regressions of profitability-related factors on the $r^{\text{RD}}_{\text{ROE}}$ are reported in Panel B of the table. The alphas of most related factors turn insignificant except for PEAD and HML(EKP). Interestingly, the annually-rebalanced $r^{\text{RD}}_{\text{ROE}}$ lowers the t-statistics of the alphas of monthly-rebalanced factors, R_ROE, UMD, and UMD(d) below the critical 2.0 level. In panel C, we run similar univariate regressions but use RMW as the indpendent variable. We do not find that RMW has a similar explanatory power to $r^{\text{RD}}_{\text{ROE}}$ on the related factors. Comparing the results on panels B and C, it is clear that the R&D adjustment is important for the profitability factor.

5.5 Within Industry Performance

Figure 3 shows that firms within HighTech and Health industries are highly R&D intensive and their R&D expenses as a fraction of total sales have been growing through the entire sample period. The R&D adjustment might, therefore, have a stronger impact on the firms within these two industries. We examine the performance of the profitability factors within each of the Fama-French five industries and report them in Table 11.

Specifically, we assign firms into two size groups based on median within-industry NYSE market equity and three profitability groups based on 30^{th} and 70^{th} percentile within-industry NYSE profitability at the June end of year t. The firms are held until the June end of the year $t+1$. The within-industry profitability factors are 1/2(Small Robust+Big Robust) - 1/2(Small Weak+Big Weak) within each industry. As before, we select the earnings measures from the bottom-line earnings to the top-line, including IB, SPI, EBT, FF, RD, RD+.3SG, NM, and REVT.

Panel A shows that the HighTech and Health industries have the highest fraction of firms reporting positive R&D expenses (column RD>0) and the highest R&D expenses as factions of both book equity (column RD/B) and total sales (column RD/REVT). Panel B reports the mean excess return of the profitability factors and the associated t-statistics. Relative to the bottom-line earnings, FF earnings measure captures the profitability return spread in the Consumer and Manufacturing firms. The R&D adjustment improves the FF earnings measure by capturing of profitability return spread in the two R&D intensive industries.

6 Model Comparison Tests

In this section, perform additional tests of the ability of the intangibles-adjusted factor to expand the portfolio space, relative to other asset pricing models. We rely on the test of Barillas et al. (2020) which tests whether there is a statistically significant improvement in the squared Sharpe ratio of one model versus another. Our model comparisons take two forms.

The first, which is only applicable to models which include profitability-related factors (all except DHS3), we replace the original profitability factor in the benchmark model with $r^{\text{RD}}_{\text{ROE}}$. Thus the comparison is the model with its profitability factor replaced with $r^{\text{RD}}_{\text{ROE}}$. We denote this as MODEL/$r^{\text{RD}}_{\text{ROE}}$ in Table 12.

The second type of test tests whether adding the intangibles-adjusted profitability factor to the benchmark model improves the squared Sharpe ratio. We denote this as MODEL+$r^{\text{RD}}_{\text{ROE}}$ in the table. For example, this tests, when applied to the FF5 model, asks whether a six-factor model consisting of the FF5 factors, plus $r^{\text{RD}}_{\text{ROE}}$, has a significantly higher R^2 than the FF5 model.

In Table 12 we report the results of the Barillas et al. (2020) test for the full period and two subperiods. the tests are against the same eight benchmark models used above. The first row for each model is the change in the squared Sharpe ratio from either swapping profitability factors or adding the intangibles-adjusted profitability factor as an extra factor (that is the improvement of the model listed in the column over the model listed in the row). The second row for each comparison is the p-value for the null hypothesis that the change in the Sharpe ratio is zero.

For the augmented models, where we add $r^{\text{RD}}_{\text{ROE}}$ as a factor, all of the tests reject the null of no improvement. The largest p-value across the full sample and two subsamples is 0.009. Thus the intangibles-adjusted factor improves the portfolio opportunity set.

For the models that replace an existing profitability factor with our intangibles-adjusted profitability, the improvement in R^2 is not significant for the HMXZ5 (for the full sample and both subsamples) nor for FF6 using the intangibles-adjusted HML factor of Eisfeldt et al. (2022) (in the first subperiod only).

The overall picture from the Barillas et al. (2020) tests is that the intangibles-adjusted profitability factor is important for explaining the cross-section of asset returns.

7 Factor or Anomaly

Whenever a new pricing factor is proposed there is always the concern about whether the identified factor is a true proxy for systematic risk or an anomaly. If the latter, is it true mispricing, which is likely to be arbitraged away over time, or is it due to data mining? The data realities of R&D reporting do not allow us the type of out-of-sample analysis in Mclean and PontiffF (2016) or Linnainmaa and Roberts (2016), so we report some suggestive statistics.

The profitability factors across various models are linked to expected returns through the present value relation in Equation 3. However, as noted by Fama and French (2006, 2015), Equation 3 is unable to distinguish between risk factors and anomalies. The fact that the pricing results are stable across time is of some comfort, but certainly not definitive.

A true factor should be risky. We look at the risk of profitability factors in two ways. We first ask what is the maximal drawdown for each factor. The results are reported in Table 13. All of the profitability factors have non-trivial downside risk although it is clear that the intangibles-adjusted profitability factors have the smallest realized downside.

Table 14 looks at the behavior of the profitability factors during market downturns. One might expect from their construction (and even the name "Robust Minus Weak") that these factors might provide some downside protection since weak firms might be expected to suffer more from economic downturns. That is born out in the tale. The factors have higher returns when the excess return on the market is negative than when it is positive. This is true for all of the profitability factors, not just those adjusted for intangibles.

The results suggest that, if anything, these factors are hedging against market risk. However, the pricing results and R^2 improvement results reported above, control the inherent beta risk of the profitability factors.

8 Conclusions

Fama and French (2015) develop a clever method for constructing an economy-wide pervasive risk factor that captures shocks to firms' profitability which they denote as RMW (robust-minus-weak). Constructing this factor requires measuring firms' profitability. Fama and French (2015) use earnings before taxes and depreciation reported in firms' financial statements prepared according to Generally Accepted Accounting Principles (GAAP). Even though R&D and some components of SG&A expenditures should be treated as investments, GAAP requires full expensing of such expenditures over our sample period. We adjust the Fama-French earnings measure by adding the R&D expenditures (or R&D and a fraction of SG&A) which we call intangibles-adjusted earnings.

Intangibles-adjusted earnings forecast the cross section of future operating cash flows and cash dividends of firms better than unadjusted earnings at longer horizons. The profitability factor that we construct based on this adjustment has a significant alpha against an array of existing factor models. When the profitability factor in the FF5 model is replaced by this R&D-adjusted profitability factor, it is able to subsume various anomalies like momentum and operating leverage.

References

Abel, Andrew B., Janice C. Eberly, Stavros Panageas. 2013. Optimal inattention to the stock market with information costs and transactions costs. *Econometrica* **81**(4) 1455–1481. URL `http://www.jstor.org/stable/23524182`.

Asness, Clifford, Andrea Frazzini. 2013. The devil in hml's details. *The Journal of Portfolio Management* **39**(4) 49–68. URL `https://doi.org/10.3905/jpm.2013.39.4.049`.

Bansal, Ravi, Amir Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance* **59**(4) 1481–1509.

Barillas, Francisco, Raymond Kan, Cesare Robotti, Jay Shanken. 2020. Model comparison with sharpe ratios. *Journal of Financial and Quantitative Analysis* **55**(6) 1840–1874.

Barillas, Francisco, Jay Shanken. 2017. Which alpha? *The Review of Financial Studies* **30**(4) 1316–1338.

Belo, Frederico, Vito D Gala, Juliana Salomao, Maria Ana Vitorino. 2022. Decomposing firm value. *Journal of Financial Economics* **143**(2) 619–639.

Berk, Jonathan B. 1995. A critique of size-related anomalies. *The Review of Financial Studies* **8**(2) 275–286. URL `http://www.jstor.org/stable/2962273`.

Bloomberg, Lawrence N. 1934. The investment value of goodwill. Ph.D. thesis, Johns Hopkins University, Baltimore, MD.

Braun, Phillip A. 1993. Asset pricing and capital investment. Ph.D. thesis, University of Chicago, Chicago, IL.

Breeden, Douglas T., Michael R. Gibbons, Robert H. Litzenberger. 1989. Empirical test of the consumption-oriented capm. *The Journal of Finance* **44**(2) 231–262. URL `http://www.jstor.org/stable/2328589`.

Carhart, Mark M. 1997. On persistence in mutual fund performance. *The Journal of Finance* **52**(1) 57–82. URL `http://www.jstor.org/stable/2329556`.

Chan, Louis KC, Narasimhan Jegadeesh, Josef Lakonishok. 1996. Momentum strategies. *The Journal of Finance* **51**(5) 1681–1713.

Chen, Nai-Fu, Richard Roll, Stephen A. Ross. 1986. Economic forces and the stock market. *The Journal of Business* **59**(3) 383–403. URL `http://www.jstor.org/stable/2352710`.

Cochrane, John H. 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. *The Journal of Finance* **46**(1) 209–237. URL `http://www.jstor.org/stable/2328694`.

Connor, Gregory. 1995. The three types of factor models: A comparison of their explanatory power. *Financial Analysts Journal* **51**(3) 42–46. URL `http://www.jstor.org/stable/4479845`.

Connor, Gregory, Matthias Hagmann, Oliver Linton. 2012. Efficient semiparametric estimation of the Fama–French model and extensions. *Econometrica* **80**(2) 713–754.

Connor, Gregory, Robert A. Korajczyk. 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics* **15**(3) 373–394. doi:https://doi.org/10.1016/0304-405X(86)90027-9. URL `https://www.sciencedirect.com/science/article/pii/0304405X86900279`.

Connor, Gregory, Robert A. Korajczyk. 1988. Risk and return in an equilibrium apt: Application of a new test methodology. *Journal of Financial Economics* **21**(2) 255–289. doi:https://doi.org/10.1016/0304-405X(88)90062-1. URL `https://www.sciencedirect.com/science/article/pii/0304405X88900621`.

Connor, Gregory, Robert A. Korajczyk. 1991. The attributes, behavior, and performance of u.s. mutual funds. *Review of Quantitative Finance and Accounting* **1**(1) 5–26. URL `https://doi.org/10.1007/BF02408404`.

Crouzet, Nicolas, Janice C Eberly. 2022. Rents and intangible capital: A q+ framework. *Journal of Finance* **Forthcoming**. URL `https://ssrn.com/abstract=3880221`.

Daniel, Kent, David Hirshleifer, Lin Sun. 2020. Short-and long-horizon behavioral factors. *The Review of Financial Studies* **33**(4) 1673–1736.

Eisfeldt, Andrea L., Edward T. Kim, Dimitris Papanikolaou. 2022. Intangible value. *Critical Finance Review* **11**(2) 299–332.

Eisfeldt, Andrea L, Dimitris Papanikolaou. 2013. Organization capital and the cross-section of expected returns. *The Journal of Finance* **68**(4) 1365–1406.

Fama, Eugene F, Kenneth R French. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* **33**(1) 3–56.

Fama, Eugene F, Kenneth R French. 2006. Profitability, investment and average returns. *Journal of Financial Economics* **82**(3) 491–518.

Fama, Eugene F, Kenneth R French. 2015. A five-factor asset pricing model. *Journal of Financial Economics* **116**(1) 1–22.

Fama, Eugene F, Kenneth R French. 2016. Dissecting anomalies with a five-factor model. *The Review of Financial Studies* **29**(1) 69–103.

Fama, Eugene F, Kenneth R French. 2018. Choosing factors. *Journal of Financial Economics* **128**(2) 234–252.

FASB. 1974. Statement of financial accounting concepts no. 2, accounting for research and development costs. *Financial Accounting Standards Board, Norwalk, CT* .

Ge, Shuyi, Shaoran Li, Oliver Linton. 2022. Dynamic peer groups of arbitrage characteristics. *Journal of Business & Economic Statistics* (Forthcoming) 1–24. doi:10.1080/07350015.2021.2011736. URL https://doi.org/10.1080/07350015.2021.2011736.

Gebhardt, William R, Charles MC Lee, Bhaskaran Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* **39**(1) 135–176.

Graham, Benjamin. 1973. The intelligent investor, a book of practical counsel, 4th rev. ed. *New York: Harper & Brothers Publishers* .

Graham, Benjamin, David L. Dodd, Sidney Cottle. 1962. *Security Analysis: Principles and Technique*. McGraw-Hill Book Company, New York, NY.

Greenwald, Bruce C, Judd Kahn, Paul Sonkin, Michael van Biema. 2001. Value investing: from graham to buffett and beyond. *John Wiley & Sons* 107–137.

Hansen, Lars Peter, Kenneth J Singleton. 1982. Generalized instrumental variables estimation of nonlinear rational expectations models. *Econometrica: Journal of the Econometric Society* 1269–1286.

Hansen, Lars Peter, Kenneth J Singleton. 1983. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *Journal of political economy* **91**(2) 249–265.

Hou, Kewei, Haitao Mo, Chen Xue, Lu Zhang. 2021. An augmented q-factor model with expected growth. *Review of Finance* **25**(1) 1–41.

Hou, Kewei, Chen Xue, Lu Zhang. 2015. Digesting anomalies: An investment approach. *The Review of Financial Studies* **28**(3) 650–705.

Jagannathan, Ravi, Yong Wang. 2007. Lazy investors, discretionary consumption, and the cross-section of stock returns. *The Journal of Finance* **62**(4) 1623–1661. URL http://www.jstor.org/stable/4622313.

Jagannathan, Ravi, Zhenyu Wang. 1996. The conditional capm and the cross-section of expected returns. *The Journal of Finance* **51**(1) 3–53.

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

Kelly, Bryan, Seth Pruitt, Yinan Su. 2019. Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics* **134** 501–534.

Kim, Soohun, Robert A. Korajczyk, Andreas Neuhierl. 2021. Arbitrage Portfolios. *The Review of Financial Studies* **34**(6) 2813–2856. doi:10.1093/rfs/hhaa102. URL `https://doi.org/10.1093/rfs/hhaa102`.

Kogan, Leonid, Dimitris Papanikolaou. 2014. Growth opportunities, technology shocks, and asset prices. *The Journal of Finance* **69**(2) 675–718.

Koh, Ping-Sheng, David M Reeb. 2015. Missing r&d. *Journal of Accounting and Economics* **60**(1) 73–94.

Lehmann, Bruce N., David M. Modest. 1988. The empirical foundations of the arbitrage pricing theory. *Journal of Financial Economics* **21**(2) 213–254. doi:https://doi.org/10.1016/0304-405X(88)90061-X. URL `https://www.sciencedirect.com/science/article/pii/0304405X8890061X`.

Lev, Baruch, Theodore Sougiannis. 1996. The capitalization, amortization, and value-relevance of r&d. *Journal of Accounting and Economics* **21**(1) 107–138.

Li, Kevin K, Partha Mohanram. 2014. Evaluating cross-sectional forecasting models for implied cost of capital. *Review of Accounting Studies* **19**(3) 1152–1185.

Linnainmaa, Juhani T., Michael R. Roberts. 2016. The history of the cross section of stock returns. URL `http://dx.doi.org/10.2139/ssrn.2897719`.

Mclean, R. David, Jeffrey PontiffF. 2016. Does academic research destroy stock return predictability? *The Journal of Finance* **71**(1) 5–32. doi:https://doi.org/10.1111/jofi.12365. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/jofi.12365`.

Novy-Marx, Robert. 2011. Operating leverage. *Review of Finance* **15**(1) 103–134.

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

Papanikolaou, Dimitris. 2011. Investment shocks and asset prices. *Journal of Political Economy* **119**(4) 639–685.

Parker, Jonathan A, Christian Julliard. 2005. Consumption risk and the cross section of expected returns. *Journal of Political Economy* **113**(1) 185–222.

Peters, Ryan H, Lucian A Taylor. 2017. Intangible capital and the investment-q relation. *Journal of Financial Economics* **123**(2) 251–272.

Roll, Richard, Stephen A. Ross. 1980. An empirical investigation of the arbitrage pricing theory. *The Journal of Finance* **35**(5) 1073–1103. URL `http://www.jstor.org/stable/2327087`.

Rosenberg, Barr, Walt McKibben. 1973. The prediction of systematic and specific risk in common stocks. *Journal of Financial and Quantitative Analysis* **8**(2) 317–333. URL `http://www.jstor.org/stable/2330027`.

Stambaugh, Robert F, Yu Yuan. 2017. Mispricing factors. *The Review of Financial Studies* **30**(4) 1270–1315.

Wang, Kai. 2023. Organizational capital, tobin's q, and the cross section of stock returns. *Tobin's Q, and the Cross Section of Stock Returns (February 19, 2023)* .

Tables and Figures

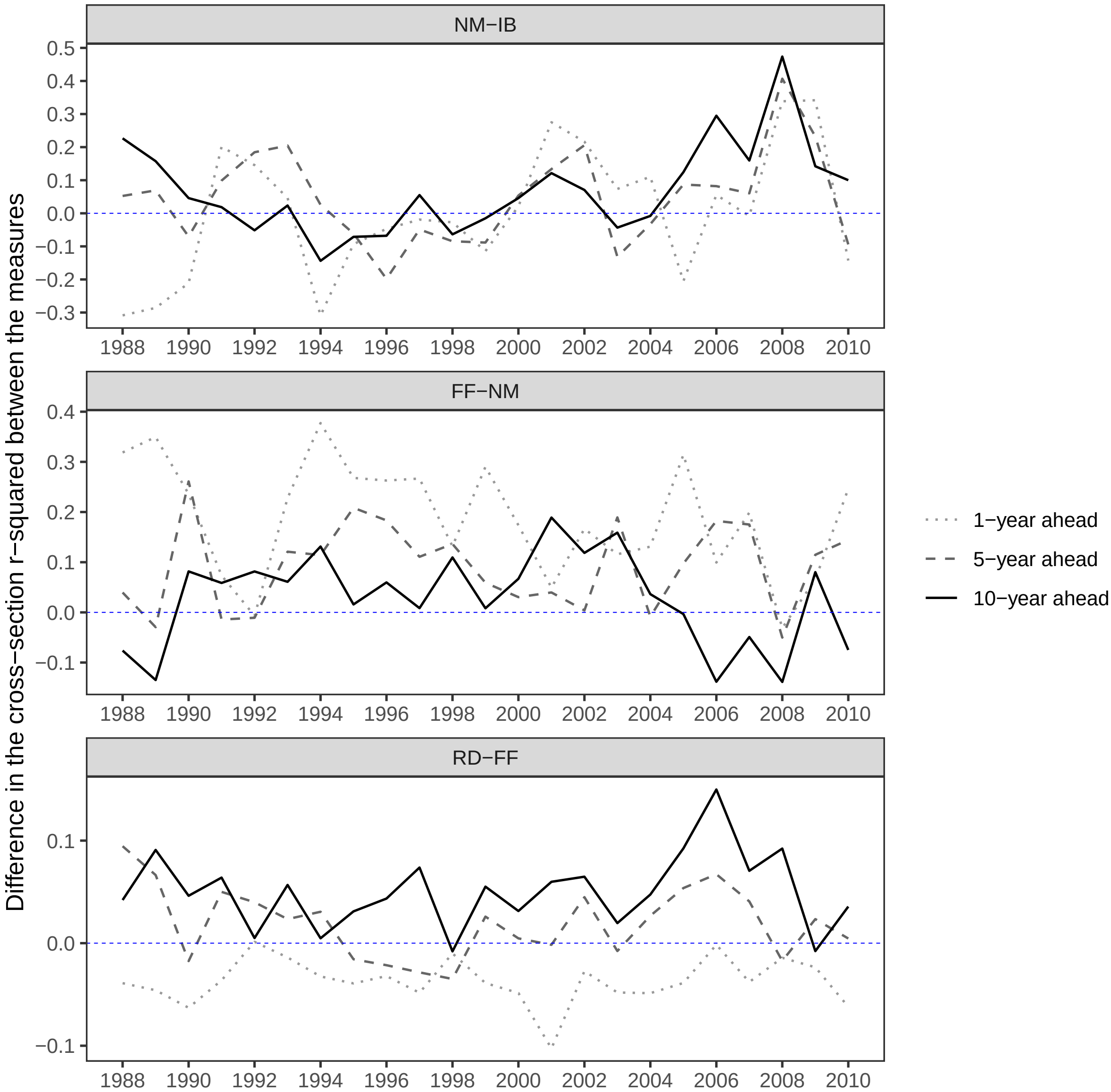

Figure 1: Difference in the cross-section cash profitability regression r-squared between measures at 1-, 5-, 10-year horizon.

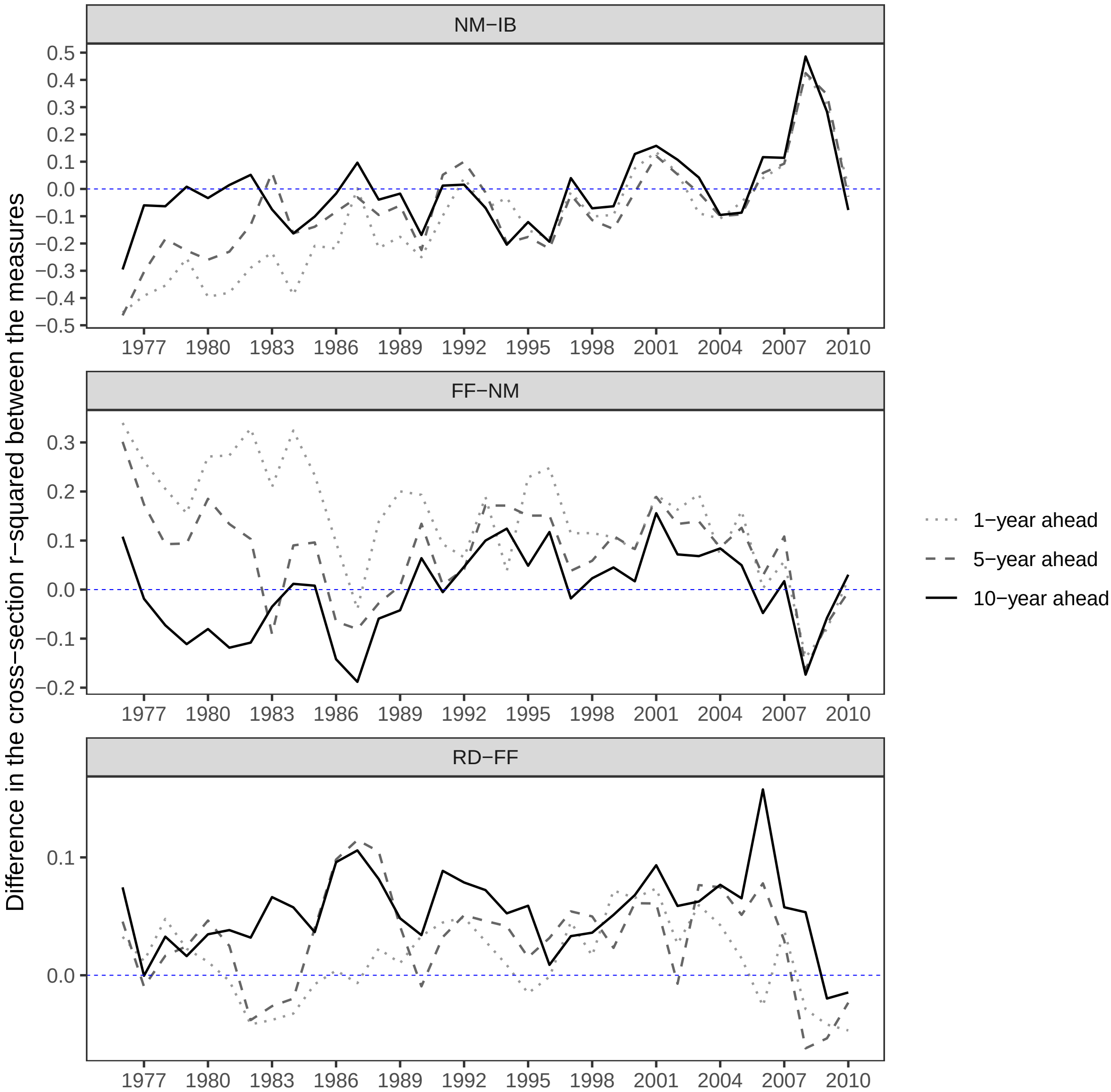

Figure 2: Difference in the cross-section cash dividend regression r-squared between measures at 1-, 5-, 10-year horizon.

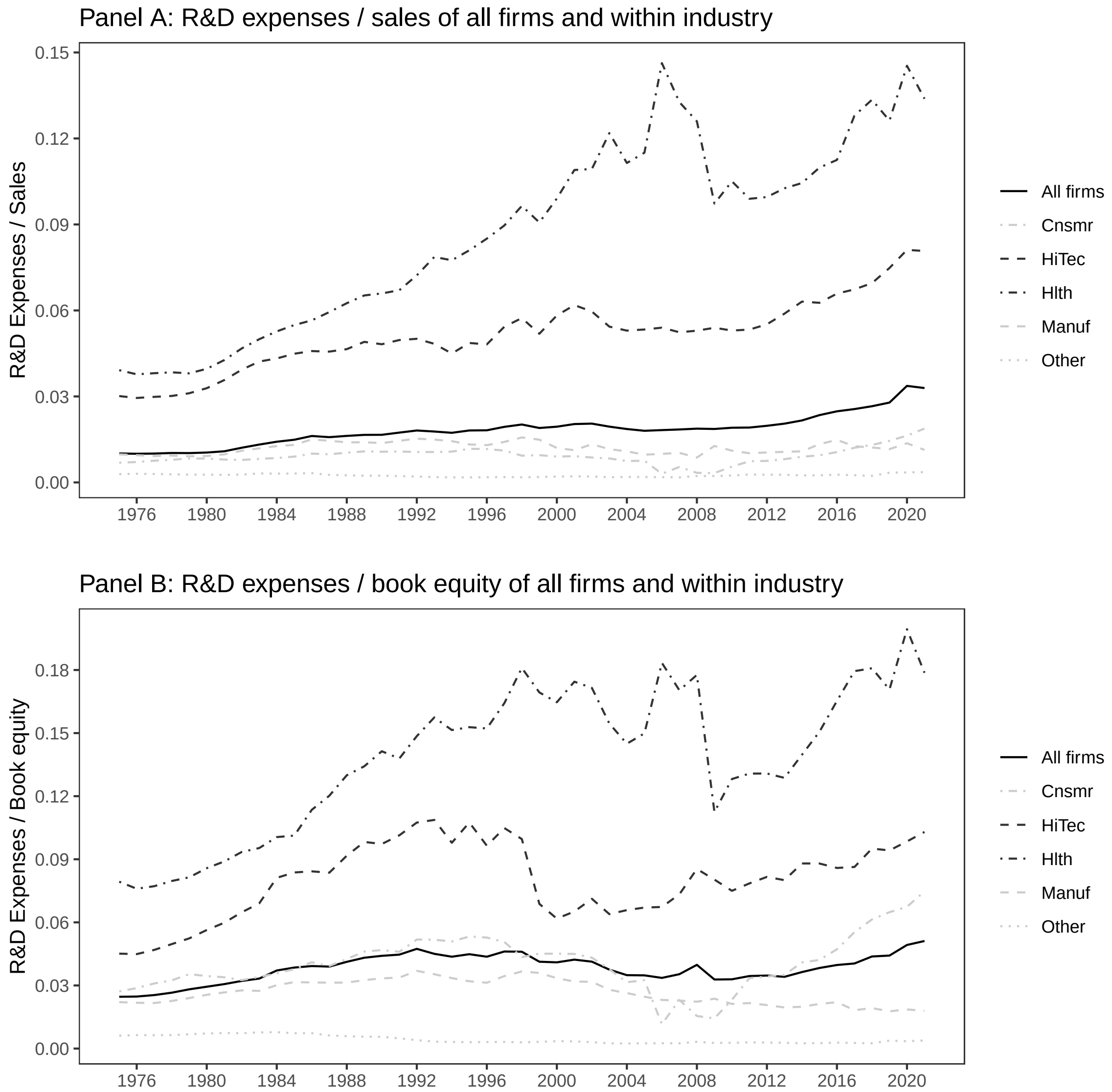

Figure 3: R&D expenses as fractions of sales and book equity, July 1976 to December 2021.

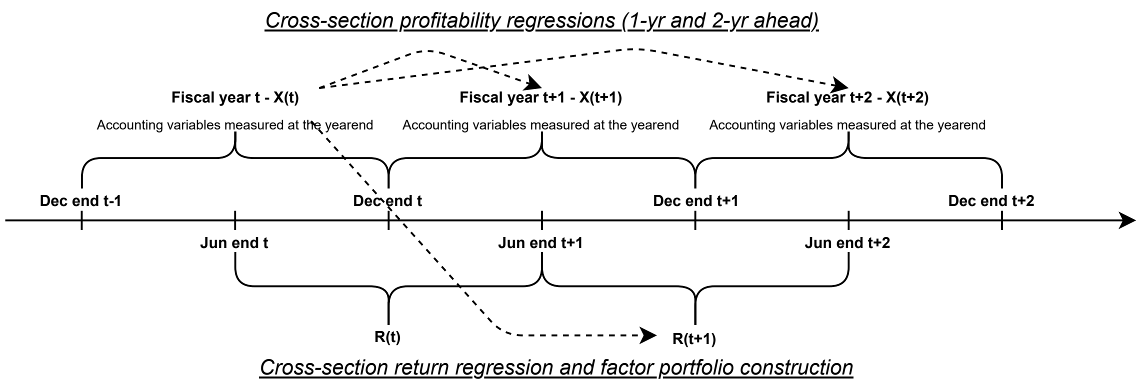

Figure 4: Time line for portfolio formation and cross-section profitability regressions.

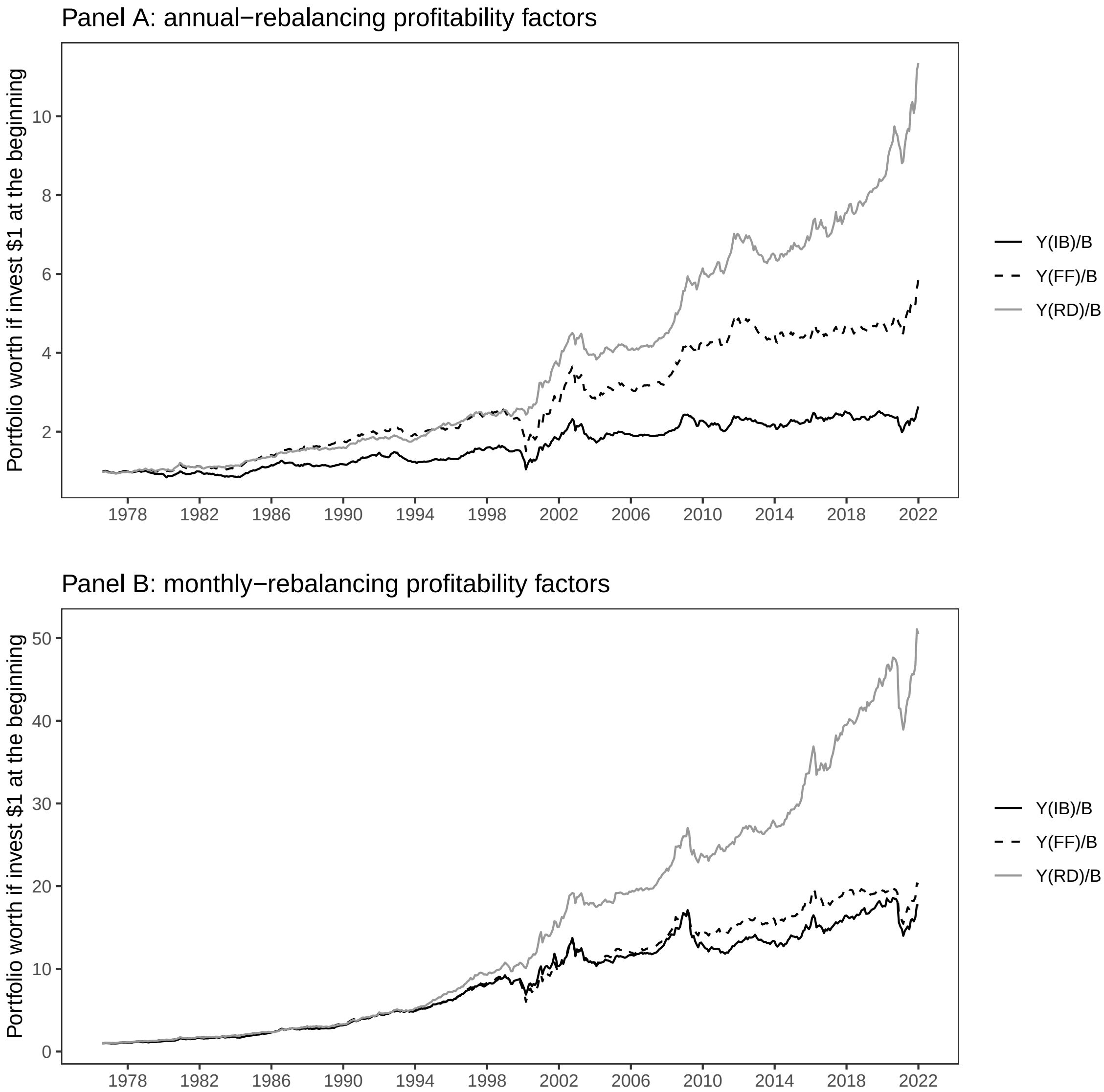

Figure 5: Cumulative return of annual- and monthly-rebalanced profitability factors, July 1976 to December 2021.

Table 1: Cross-section regression of future cash profitability at the portfolio level (1988 - 2021).

This table reports the time-series averages of the slope coefficients, t-statistics, and adjusted r-squared for the cross-section regressions of future cash profitability at the portfolio level. Firms are assigned into 5-by-10 industry and ROE(FF) groups. The operating cash flow is the Compustat data item OANCF. For notation convenience, we denote $Z_{t+\tau}^{X}$ as $Y_{t+\tau}^{X}/B_t$ where X includes IB, NM, FF, RD, RD+.3 SG, and RD+SG, and we denote the future cash profitability as $Z_{t+\tau}^{\mathrm{OCF}}$ where $\tau = 1, 2, ..., 10$. We also test the null hypothesis that the difference of the r-squared time-series between $Z_{t+\tau}^{\mathrm{NM}}$, $Z_{t+\tau}^{\mathrm{FF}}$, $Z_{t+\tau}^{\mathrm{RD}}$, $Z_{t+\tau}^{\mathrm{RD+.3\,SG}}$, $Z_{t+\tau}^{\mathrm{RD+SG}}$ and $Z_{t+\tau}^{\mathrm{IB}}$, $Z_{t+\tau}^{\mathrm{NM}}$, $Z_{t+\tau}^{\mathrm{FF}}$, $Z_{t+\tau}^{\mathrm{FF}}$, $Z_{t+\tau}^{\mathrm{FF}}$ is zero and report the t-statistics in the panel B and D. We remove firms with total assets less than \$12.5 million or book equity less than \$25 million in year t to avoid the influence of small firms. The sample period is from 1988 to 2021 as items from the statement of the cash flows become available only after 1988.

Panel A: Average slopes for Z_t^{IB}, Z_t^{NM}, and Z_t^{FF}

	R^2	Intc	Z_t^{IB}		R^2	Intc	Z_t^{NM}		R^2	Intc	Z_t^{FF}
Z_{t+1}^{OCF}	0.51	0.14	1.04	Z_{t+1}^{OCF}	0.51	0.04	0.29	Z_{t+1}^{OCF}	0.70	0.05	0.70
Z_{t+2}^{OCF}	0.46	0.16	1.08	Z_{t+2}^{OCF}	0.49	0.04	0.31	Z_{t+2}^{OCF}	0.64	0.06	0.74
Z_{t+3}^{OCF}	0.43	0.18	1.09	Z_{t+3}^{OCF}	0.50	0.05	0.33	Z_{t+3}^{OCF}	0.60	0.08	0.75
Z_{t+4}^{OCF}	0.43	0.20	1.17	Z_{t+4}^{OCF}	0.48	0.05	0.35	Z_{t+4}^{OCF}	0.58	0.09	0.79
Z_{t+5}^{OCF}	0.39	0.22	1.20	Z_{t+5}^{OCF}	0.44	0.08	0.36	Z_{t+5}^{OCF}	0.53	0.12	0.80
Z_{t+6}^{OCF}	0.37	0.25	1.23	Z_{t+6}^{OCF}	0.42	0.10	0.37	Z_{t+6}^{OCF}	0.51	0.14	0.84
Z_{t+7}^{OCF}	0.35	0.27	1.28	Z_{t+7}^{OCF}	0.38	0.12	0.38	Z_{t+7}^{OCF}	0.46	0.16	0.87
Z_{t+8}^{OCF}	0.30	0.30	1.30	Z_{t+8}^{OCF}	0.34	0.15	0.39	Z_{t+8}^{OCF}	0.40	0.19	0.88
Z_{t+9}^{OCF}	0.30	0.34	1.43	Z_{t+9}^{OCF}	0.35	0.15	0.45	Z_{t+9}^{OCF}	0.40	0.22	0.96
Z_{t+10}^{OCF}	0.27	0.37	1.47	Z_{t+10}^{OCF}	0.34	0.18	0.46	Z_{t+10}^{OCF}	0.37	0.25	0.98

Panel B: t-statistics for Panel A

		Intc	Z_t^{IB}		ΔR^2 (NM-IB)	Intc	Z_t^{NM}		ΔR^2 (FF-NM)	Intc	Z_t^{FF}
1		10.27	11.93	1	0.06	3.22	20.30	1	7.93	2.95	21.14
2		12.87	12.76	2	0.70	3.92	19.39	2	6.59	4.66	22.98
3		12.89	10.36	3	2.06	3.94	27.01	3	5.28	5.71	25.31
4		13.68	10.77	4	1.81	4.66	31.48	4	4.81	6.59	25.05
5		17.91	9.97	5	1.61	5.84	18.80	5	5.02	10.47	20.77
6		22.28	10.42	6	1.58	5.79	14.40	6	4.59	12.00	15.80
7		27.09	17.45	7	0.77	4.37	9.96	7	3.99	12.19	14.00
8		22.07	17.59	8	1.32	3.90	7.60	8	3.21	11.03	12.56
9		19.65	20.91	9	1.78	5.32	10.84	9	2.45	14.20	19.52
10		20.31	16.31	10	2.42	5.04	10.28	10	1.45	15.51	23.65

Panel C: Average slopes for Z_t^{RD}, $Z_t^{\mathrm{RD+.3\,SG}}$, and $Z_t^{\mathrm{RD+SG}}$

	R^2	Intc	Z_t^{RD}		R^2	Intc	$Z_t^{\mathrm{RD+.3\,SG}}$		R^2	Intc	$Z_t^{\mathrm{RD+SG}}$
Z_{t+1}^{OCF}	0.66	0.04	0.62	Z_{t+1}^{OCF}	0.62	0.03	0.50	Z_{t+1}^{OCF}	0.48	0.06	0.29
Z_{t+2}^{OCF}	0.63	0.05	0.66	Z_{t+2}^{OCF}	0.60	0.03	0.54	Z_{t+2}^{OCF}	0.47	0.06	0.32
Z_{t+3}^{OCF}	0.60	0.07	0.68	Z_{t+3}^{OCF}	0.59	0.05	0.56	Z_{t+3}^{OCF}	0.47	0.07	0.33
Z_{t+4}^{OCF}	0.59	0.07	0.72	Z_{t+4}^{OCF}	0.57	0.06	0.59	Z_{t+4}^{OCF}	0.46	0.08	0.35
Z_{t+5}^{OCF}	0.55	0.09	0.74	Z_{t+5}^{OCF}	0.54	0.07	0.61	Z_{t+5}^{OCF}	0.43	0.10	0.37
Z_{t+6}^{OCF}	0.54	0.11	0.79	Z_{t+6}^{OCF}	0.53	0.09	0.65	Z_{t+6}^{OCF}	0.42	0.12	0.38
Z_{t+7}^{OCF}	0.50	0.13	0.82	Z_{t+7}^{OCF}	0.49	0.11	0.67	Z_{t+7}^{OCF}	0.38	0.14	0.40
Z_{t+8}^{OCF}	0.45	0.15	0.85	Z_{t+8}^{OCF}	0.44	0.13	0.70	Z_{t+8}^{OCF}	0.35	0.16	0.42
Z_{t+9}^{OCF}	0.45	0.17	0.93	Z_{t+9}^{OCF}	0.44	0.14	0.77	Z_{t+9}^{OCF}	0.36	0.17	0.46
Z_{t+10}^{OCF}	0.42	0.20	0.96	Z_{t+10}^{OCF}	0.42	0.17	0.80	Z_{t+10}^{OCF}	0.36	0.19	0.49

Panel D: t-statistics for Panel D

	ΔR^2 (RD-FF)	Intc	Z_t^{RD}		ΔR^2 (RD+.3 SG-FF)	Intc	$Z_t^{\mathrm{RD+.3\,SG}}$		ΔR^2 (RD+SG-FF)	Intc	$Z_t^{\mathrm{RD+SG}}$
1	-7.92	2.91	24.87	1	-7.11	2.12	23.37	1	-9.98	5.68	18.46
2	-1.84	4.10	24.29	2	-3.29	2.80	22.83	2	-7.72	6.54	17.58

3	0.21	5.49	34.74	3	-1.05	3.73	31.84	3	-6.79	7.96	25.69
4	1.25	6.30	31.35	4	-0.26	4.20	29.81	4	-6.35	8.60	27.08
5	2.72	9.72	24.27	5	0.36	7.13	24.88	5	-4.47	13.86	24.06
6	3.87	7.82	14.82	6	1.50	6.36	16.71	6	-4.71	9.68	17.15
7	3.99	6.62	11.95	7	1.54	5.22	12.58	7	-3.68	6.48	11.91
8	4.61	5.71	9.95	8	2.64	4.32	9.73	8	-2.50	4.69	8.39
9	5.54	8.40	15.54	9	3.40	6.50	15.45	9	-1.72	6.93	13.09
10	6.61	9.92	17.83	10	4.01	7.21	17.40	10	-0.64	6.43	12.89

Table 2: Cross-section regression of future cash profitability at portfolio level with RD_t/B_t and SGA_t/B_t as additional independent variables (1988 - 2021).

This table reports the time-series averages of the slope coefficients, t-statistics, and adjusted r-squared for the cross-section regressions of future cash profitability. The operating cash flows is the Compustat data item OANCF. For notation convenience, we denote $Z^X_{t+\tau}$ as $Y^X_{t+\tau}/B_t$ where X includes IB, NM, FF, and RD and we denote the future cash profitability as $Z^{\mathrm{OCF}}_{t+\tau}$ where $\tau = 1, 2, ..., 10$. We remove firms with total assets less than \$12.5 million or book equity less than \$25 million in year t to avoid the influence of small firms. The sample period is from 1988 to 2021 as items from the statement of the cash flows become available only after 1988.

Panel A: Average slopes					
	Adj. R^2	Intc	Z^{FF}_t	RD_t/B_t	SGA_t/B_t
Z^{OCF}_{t+1}	0.37	0.04	0.63	-0.04	0.04
Z^{OCF}_{t+2}	0.31	0.06	0.63	0.03	0.05
Z^{OCF}_{t+3}	0.27	0.07	0.65	0.06	0.06
Z^{OCF}_{t+4}	0.24	0.08	0.70	0.14	0.07
Z^{OCF}_{t+5}	0.21	0.10	0.73	0.22	0.09
Z^{OCF}_{t+6}	0.19	0.11	0.78	0.30	0.10
Z^{OCF}_{t+7}	0.17	0.12	0.85	0.38	0.12
Z^{OCF}_{t+8}	0.15	0.14	0.93	0.49	0.14
Z^{OCF}_{t+9}	0.14	0.16	1.01	0.52	0.16
Z^{OCF}_{t+10}	0.12	0.18	1.09	0.63	0.19
Panel B: t-statistics					
		Intc	Z^{FF}_t	RD_t/B_t	SGA_t/B_t
1		4.06	29.34	-2.13	2.41
2		7.56	28.55	0.71	2.37
3		12.93	19.87	0.97	2.15
4		10.99	16.54	1.68	2.56
5		9.93	14.09	2.40	2.78
6		5.87	10.33	2.54	3.11
7		5.33	10.74	2.86	4.22
8		3.71	6.38	2.69	4.18
9		5.03	6.07	2.26	4.99
10		3.90	4.81	3.44	3.57

Table 3: Cross-section regression of future dividend at portfolio level (1976 - 2021).

This table reports the time-series averages of the slope coefficients, t-statistics, and adjusted r-squared for the cross-section regressions of future cash profitability at portfolio level. Firms are assigned into 5-by-10 industry and ROE(FF) groups. The dividend is the Compustat data item DVC. For notation convenience, we denote $Z^X_{t+\tau}$ as $Y^X_{t+\tau}/B_t$ where X includes IB, NM, FF, RD, RD+.3 SG, RD+SG, and we denote the future cash profitability as $Z^{\text{DVC}}_{t+\tau}$ where $\tau = 1, 2, ..., 10$. We also test the null hypothesis that the difference of the r-squared time-series between $Z^{\text{NM}}_{t+\tau}$, $Z^{\text{FF}}_{t+\tau}$, $Z^{\text{RD}}_{t+\tau}$, $Z^{\text{RD+.3 SG}}_{t+\tau}$, $Z^{\text{RD+SG}}_{t+\tau}$ and $Z^{\text{IB}}_{t+\tau}$, $Z^{\text{NM}}_{t+\tau}$, $Z^{\text{FF}}_{t+\tau}$, $Z^{\text{FF}}_{t+\tau}$, $Z^{\text{FF}}_{t+\tau}$ is zero and report the t-statistics in the panel B and D. We remove firms with total assets less than \$12.5 million or book equity less than \$25 million in year t to avoid the influence of small firms. The sample period is from 1976 to 2021 as R&D expense data is only valid after January 1975.

Panel A: Average slopes for Z^{IB}_t, Z^{NM}_t, and Z^{FF}_t

	R^2	Intc	Z^{IB}_t		R^2	Intc	Z^{NM}_t		R^2	Intc	Z^{FF}_t
Z^{DVC}_{t+1}	0.47	0.02	0.25	Z^{DVC}_{t+1}	0.36	0.01	0.06	Z^{DVC}_{t+1}	0.50	0.01	0.14
Z^{DVC}_{t+2}	0.50	0.02	0.29	Z^{DVC}_{t+2}	0.38	0.00	0.07	Z^{DVC}_{t+2}	0.52	0.01	0.16
Z^{DVC}_{t+3}	0.48	0.02	0.32	Z^{DVC}_{t+3}	0.37	0.00	0.07	Z^{DVC}_{t+3}	0.49	0.01	0.17
Z^{DVC}_{t+4}	0.47	0.02	0.35	Z^{DVC}_{t+4}	0.39	0.00	0.08	Z^{DVC}_{t+4}	0.48	0.01	0.19
Z^{DVC}_{t+5}	0.47	0.03	0.39	Z^{DVC}_{t+5}	0.40	0.00	0.09	Z^{DVC}_{t+5}	0.47	0.01	0.21
Z^{DVC}_{t+6}	0.46	0.03	0.43	Z^{DVC}_{t+6}	0.41	-0.01	0.11	Z^{DVC}_{t+6}	0.46	0.01	0.23
Z^{DVC}_{t+7}	0.45	0.03	0.47	Z^{DVC}_{t+7}	0.40	-0.01	0.12	Z^{DVC}_{t+7}	0.45	0.01	0.25
Z^{DVC}_{t+8}	0.44	0.03	0.52	Z^{DVC}_{t+8}	0.40	-0.01	0.13	Z^{DVC}_{t+8}	0.43	0.01	0.27
Z^{DVC}_{t+9}	0.42	0.03	0.57	Z^{DVC}_{t+9}	0.39	-0.01	0.14	Z^{DVC}_{t+9}	0.40	0.02	0.29
Z^{DVC}_{t+10}	0.40	0.04	0.60	Z^{DVC}_{t+10}	0.39	-0.01	0.15	Z^{DVC}_{t+10}	0.39	0.02	0.31

Panel B: t-statistics for Panel A

		Intc	Z^{IB}_t		ΔR^2 (NM-IB)	Intc	Z^{NM}_t		ΔR^2 (FF-NM)	Intc	Z^{FF}_t
1		5.80	14.54	1	-3.42	0.74	6.11	1	7.61	1.40	13.24
2		5.19	14.05	2	-3.70	0.37	6.33	2	7.68	0.44	9.87
3		5.45	13.08	3	-3.33	0.45	7.55	3	6.26	1.33	16.34
4		5.20	13.48	4	-3.04	0.05	7.84	4	5.77	0.80	14.77
5		4.53	10.08	5	-2.39	-0.38	14.26	5	4.74	1.12	16.15
6		4.65	10.36	6	-1.93	-0.80	9.67	6	3.62	1.06	18.59
7		4.53	9.84	7	-1.49	-1.41	12.55	7	3.00	1.18	16.11
8		4.69	10.49	8	-1.41	-1.58	11.17	8	1.58	1.16	11.09
9		5.50	8.54	9	-0.78	-1.88	9.46	9	0.58	1.48	7.39
10		6.57	10.18	10	-0.40	-1.43	10.76	10	-0.17	1.97	9.32

Panel C: Average slopes for Z^{RD}_t, $Z^{\text{RD+.3 SG}}_t$, and $Z^{\text{RD+SG}}_t$

	R^2	Intc	Z^{RD}_t		R^2	Intc	$Z^{\text{RD+.3 SG}}_t$		R^2	Intc	$Z^{\text{RD+SG}}_t$
Z^{DVC}_{t+1}	0.52	0.01	0.13	Z^{DVC}_{t+1}	0.49	0.00	0.11	Z^{DVC}_{t+1}	0.36	0.01	0.06
Z^{DVC}_{t+2}	0.54	0.00	0.15	Z^{DVC}_{t+2}	0.51	0.00	0.12	Z^{DVC}_{t+2}	0.38	0.01	0.07
Z^{DVC}_{t+3}	0.52	0.01	0.16	Z^{DVC}_{t+3}	0.50	0.00	0.13	Z^{DVC}_{t+3}	0.38	0.01	0.07
Z^{DVC}_{t+4}	0.51	0.00	0.18	Z^{DVC}_{t+4}	0.50	0.00	0.15	Z^{DVC}_{t+4}	0.39	0.01	0.09
Z^{DVC}_{t+5}	0.50	0.00	0.20	Z^{DVC}_{t+5}	0.50	0.00	0.16	Z^{DVC}_{t+5}	0.40	0.00	0.10
Z^{DVC}_{t+6}	0.50	0.00	0.22	Z^{DVC}_{t+6}	0.51	-0.01	0.18	Z^{DVC}_{t+6}	0.41	0.00	0.11
Z^{DVC}_{t+7}	0.49	0.00	0.24	Z^{DVC}_{t+7}	0.50	-0.01	0.20	Z^{DVC}_{t+7}	0.41	0.00	0.12
Z^{DVC}_{t+8}	0.47	0.00	0.26	Z^{DVC}_{t+8}	0.49	-0.01	0.22	Z^{DVC}_{t+8}	0.41	0.00	0.13
Z^{DVC}_{t+9}	0.45	0.00	0.28	Z^{DVC}_{t+9}	0.47	-0.01	0.24	Z^{DVC}_{t+9}	0.41	0.00	0.15
Z^{DVC}_{t+10}	0.44	0.01	0.30	Z^{DVC}_{t+10}	0.47	-0.01	0.26	Z^{DVC}_{t+10}	0.41	0.00	0.16

Panel D: t-statistics for Panel C

	ΔR^2 (RD-FF)	Intc	Z^{RD}_t		ΔR^2 (RD+.3 SG-FF)	Intc	$Z^{\text{RD+.3 SG}}_t$		ΔR^2 (RD+SG-FF)	Intc	$Z^{\text{RD+SG}}_t$
1	2.44	0.92	16.22	1	-1.94	0.44	12.67	1	-8.31	1.03	5.76
2	2.68	0.36	18.53	2	-1.34	0.04	12.96	2	-8.52	0.52	4.91
3	3.28	0.77	24.56	3	0.40	0.08	17.34	3	-6.21	1.00	6.93

4	3.74	0.48	24.17	4	2.27	-0.26	20.79	4	-5.47	0.50	7.36
5	4.28	0.49	24.88	5	3.57	-0.75	25.06	5	-4.24	0.67	14.78
6	5.46	0.37	24.61	6	5.31	-1.11	21.69	6	-2.95	0.25	11.42
7	7.58	0.30	20.93	7	6.27	-1.64	19.67	7	-2.34	0.02	12.81
8	8.09	0.33	17.07	8	7.88	-1.67	18.21	8	-0.80	-0.26	13.31
9	7.65	0.40	11.78	9	7.65	-1.82	12.81	9	0.08	-0.71	10.71
10	9.24	0.57	14.34	10	9.08	-1.68	15.47	10	1.10	-0.44	13.07

Table 4: Cross-section regression of future dividend at portfolio level with RD_t/B_t and SGA_t/B_t as additional independent variables (1988 - 2021).

This table reports the time-series averages of the slope coefficients, t-statistics, and adjusted r-squared for the cross-section regressions of the future cash dividends. The cash dividend is the Compustat data item DVC. For notation convenience, we denote $Z^X_{t+\tau}$ as $Y^X_{t+\tau}/B_t$ where X includes IB, NM, FF, and RD and we denote the future cash profitability as $Z^{\mathrm{DVC}}_{t+\tau}$ where $\tau = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$. We remove firms with total assets less than $12.5 million or book equity less than $25 million in year t to avoid the influence of small firms. The sample period is from 1976 to 2021.

Panel A: Average slopes					
	Adj. R^2	Intc	Z_t^{FF}	RD_t/B_t	SGA_t/B_t
Z^{DVC}_{t+1}	0.52	0.01	0.13	0.09	0.01
Z^{DVC}_{t+2}	0.55	0.00	0.15	0.10	0.01
Z^{DVC}_{t+3}	0.52	0.00	0.16	0.12	0.01
Z^{DVC}_{t+4}	0.52	0.00	0.17	0.14	0.02
Z^{DVC}_{t+5}	0.52	0.00	0.19	0.16	0.03
Z^{DVC}_{t+6}	0.52	-0.01	0.20	0.19	0.04
Z^{DVC}_{t+7}	0.51	-0.01	0.21	0.22	0.04
Z^{DVC}_{t+8}	0.50	-0.01	0.23	0.26	0.05
Z^{DVC}_{t+9}	0.48	-0.01	0.25	0.29	0.06
Z^{DVC}_{t+10}	0.47	-0.01	0.26	0.36	0.07

Panel B: t-statistics					
		Intc	Z_t^{FF}	RD_t/B_t	SGA_t/B_t
1		0.81	19.85	3.20	0.88
2		0.31	17.26	4.64	0.89
3		0.50	24.83	4.12	1.16
4		0.07	18.71	4.30	1.54
5		-0.47	23.18	4.20	2.30
6		-1.02	23.61	5.06	2.12
7		-1.59	18.86	5.93	2.23
8		-2.00	16.25	5.96	2.61
9		-2.45	12.48	4.36	2.37
10		-2.07	8.96	8.98	3.22

Table 5: Summary statistics for monthly 2-by-3 ME-ROE(X) portfolio and factor percent returns (July 1976 - December 2021).

In the June end of each year t, firms are assigned into two size (ME) groups based on median June end market equity of NYSE firms, and assigned independently into three profitability (Y/B) groups based on the 30 and 70 percentiles profitability of all NYSE firms reported annual earnings in calendar year t-1. Mean and t-statistics are the time-series average and the associated t -statistics are given in parentheses. t-statistics of the monthly value-weighted return of each portfolio minus the one-month Treasury bill rate. H-L is a portfolio that is long the high-ROE group and short the low-ROE group within each size group, and Mean(H-L) is the equal-weighted average of the two H-L portfolios. RMW(a) is the intercept and the associated t-statistics in the regression of Mean(H-L) on the profitability factor RMW of the Fama-French 5-factor model. All other statistics reported in this table are the time-series average of the monthly statistics of each portfolio. N(firms) is the number of firms in each portfolio. Overlap(FF) is the fraction of firms that are in both $Y(X)/B$ portfolio and the corresponding $Y(\mathrm{FF})/B$ portfolio. Book equity is measured using Compustat data item CEQ and earnings is measured using $Y(\mathrm{IB})$, $Y(\mathrm{SPI})$, $Y(\mathrm{EBT})$, $Y(\mathrm{FF})$, $Y(\mathrm{RD})$, $Y(\mathrm{RD}+.3\,\mathrm{SG})$, $Y(\mathrm{RD}+\mathrm{SG})$, $Y(\mathrm{NM})$, and $Y(\mathrm{REVT})$.

		Small				Big					
		Low	Med	High	H-L	Low	Med	High	H-L	Mean(H-L)	RMW(a)
$Y(\mathrm{IB})/B$	Mean	0.71	0.94	1.01	0.30	0.61	0.68	0.72	0.11	0.21	-0.12
	t-statistics	2.46	4.38	4.10	2.49	2.60	3.66	3.88	0.97	1.98	-2.53
	N(firms)	1754	1029	611		172	352	344			
$Y(\mathrm{SPI})/B$	Mean	0.68	0.96	1.04	0.36	0.56	0.68	0.73	0.16	0.26	-0.06
	t-statistics	2.39	4.47	4.11	3.21	2.37	3.66	3.93	1.35	2.61	-1.22
	N(firms)	1841	968	585		171	358	339			
$Y(\mathrm{EBT})/B$	Mean	0.67	0.94	1.04	0.38	0.54	0.65	0.76	0.22	0.30	-0.06
	t-statistics	2.28	4.23	4.42	2.72	2.29	3.45	4.11	1.84	2.62	-1.31
	N(firms)	1749	1007	630		176	343	346			
$Y(\mathrm{FF})/B$	Mean	0.64	0.97	1.07	0.43	0.49	0.67	0.77	0.28	0.36	-0.01
	t-statistics	2.30	4.40	4.32	3.36	2.14	3.54	4.29	2.45	3.41	-0.29
	N(firms)	1831	980	574		188	355	322			
$Y(\mathrm{RD})/B$	Mean	0.55	0.99	1.14	0.59	0.43	0.68	0.77	0.34	0.47	0.25
	t-statistics	2.07	4.45	4.34	7.00	1.93	3.64	4.13	3.23	6.11	5.38
	N(firms)	1711	1034	640		176	355	334			
	Overlap(FF)	0.90	0.83	0.87		0.87	0.84	0.89			
$Y(\mathrm{RD}+.3\,\mathrm{SG})/B$	Mean	0.54	0.97	1.16	0.62	0.46	0.66	0.80	0.34	0.48	0.30
	t-statistics	2.07	4.27	4.35	7.02	2.23	3.44	4.35	3.39	6.26	5.19
	N(firms)	1443	1146	796		217	354	294			
	Overlap(FF)	0.71	0.64	0.76		0.80	0.65	0.72			
$Y(\mathrm{RD}+\mathrm{SG})/B$	Mean	0.58	0.95	1.10	0.52	0.54	0.66	0.82	0.29	0.40	0.24
	t-statistics	2.26	4.02	4.17	4.90	2.73	3.37	4.48	2.88	4.84	3.42
	N(firms)	1117	1292	976		250	364	251			
$Y(\mathrm{NM})/B$	Mean	0.58	0.97	1.09	0.50	0.53	0.67	0.81	0.28	0.39	0.22
	t-statistics	2.35	4.09	4.10	4.86	2.72	3.48	4.32	2.81	4.76	3.28
	N(firms)	1205	1217	962		258	363	244			
$Y(\mathrm{REVT})/B$	Mean	0.69	0.98	1.05	0.36	0.63	0.73	0.79	0.16	0.26	0.07
	t-statistics	2.78	4.07	3.98	3.22	3.23	3.90	4.06	1.79	2.96	1.02
	N(firms)	1446	1019	920		335	333	197			

Table 6: Pricing tests on annual- and monthly-rebalanced profitability factors (July 1976 - December 2021).

To construct annual-rebalanced profitability factors, $r^{\mathrm{X}}_{\mathrm{ROE}}$, in the June end of each year t, firms are assigned into two size groups based on the median June end market equity of NYSE firms and assigned independently into three profitability groups based on the 30 and 70 percentiles profitability of all NYSE firms reported annual earnings in calendar year t-1. To construct monthly-rebalanced profitability factors, $r^{\mathrm{X}}_{\mathrm{ROEQ}}$, at the end of each month, firms are assigned into two size (ME) groups based on median June end market equity of NYSE firms, and assigned independently into three asset growth ($d\mathrm{AT}_t/\mathrm{AT}_{t-1}$) groups based on the 30 and 70 percentiles asset growth of all NYSE firms reported annual earnings in calendar year t-1, and assigned independently into three profitability groups based on the 30 and 70 percentiles (most recent available) profitability of all NYSE firms. The profitability factor return in any given month is the average return of the portfolios that are long the high profitability portfolios and short the low profitability portfolios within each size group. Firms' most recent available profitability is updated at the month end of the most recent quarterly earnings announcement date (quarterly Compustat data item RDQ, exclude if missing). To remove stale earnings, we exclude a firm if its most recent available quarter-end is 6 months away. The sample period is from July 1976 to December 2021 as the R&D expense data is only valid after January 1975. The quarterly R&D expenses (quarterly Compustat data item XRDQ) are only recorded after 1989 and some firms may report only once a year, we simply set missing XRDQ to zero with no further interpolation. The considered profitability measures include $Y(\mathrm{IB})/B$, $Y(\mathrm{SPI})/B$, $Y(\mathrm{EBT})/B$, $Y(\mathrm{FF})/B$, $Y(\mathrm{RD})/B$, $Y(\mathrm{RD}+.3\mathrm{SG})/B$, $Y(\mathrm{RD}+\mathrm{SG})/B$, $Y(\mathrm{NM})/B$, and $Y(\mathrm{REVT})/B$. We run time-series regressions of the constructed profitability factors on the existing factor models and report their intercept and associated t-statistics in the squared parenthesis. The factor models include Fama-French five factor model (FF5), FF5 plus momentum factor (FF6), FF6 with the value factor, HML, replaced by the monthly rebalanced value factor with most recent market equity, HML(Devil), from Asness and Frazzini (2013), FF6 with the value factor, HML, replaced by the intangible-adjusted value factor, HML(EKP), from Eisfeldt et al. (2022), Hou-Xue-Zhang four factor model (HXZ4) and Hou-Mo-Xue-Zhang five factor model (HMXZ5), Daniel-Hirshleifer-Sun behavioral three factor model (DHS3), and Stambaugh-Yuan four factor model (SY4). For monthly-rebalanced factors, we additionally report their time-series mean and t-statistics and the intercept and the associated t-statistics in the regression on monthly ROE factor (R_ROE) in HXZ4. The data of the factor models is taken from the authors' websites. Except for SY4 which is available only until December 2016, data on all factor models is available until December 2021.

Moving from bottom-line to top-line profitability: left to right

Panel A: Annual-rebalanced factors

	$r^{\mathrm{IB}}_{\mathrm{ROE}}$	$r^{\mathrm{SPI}}_{\mathrm{ROE}}$	$r^{\mathrm{EBT}}_{\mathrm{ROE}}$	$r^{\mathrm{FF}}_{\mathrm{ROE}}$	$r^{\mathrm{RD}}_{\mathrm{ROE}}$	$r^{\mathrm{RD+.3\,SG}}_{\mathrm{ROE}}$	$r^{\mathrm{RD+SG}}_{\mathrm{ROE}}$	$r^{\mathrm{NM}}_{\mathrm{ROE}}$	$r^{\mathrm{REVT}}_{\mathrm{ROE}}$
FF5	0.03	0.06	0.07	0.02	0.25	0.26	0.17	0.12	-0.16
	[0.77]	[1.28]	[1.52]	[0.80]	[6.44]	[5.55]	[2.77]	[1.97]	[-2.41]
FF6	0.04	0.06	0.07	0.02	0.24	0.26	0.20	0.15	-0.13
	[0.84]	[1.27]	[1.49]	[0.85]	[6.05]	[5.65]	[3.25]	[2.44]	[-1.99]
FF6/Devil	0.04	0.07	0.06	0.00	0.30	0.35	0.29	0.22	-0.13
	[0.91]	[1.55]	[1.27]	[0.15]	[7.25]	[7.19]	[4.40]	[3.51]	[-1.99]
FF6/EKP	0.04	0.06	0.04	0.02	0.28	0.29	0.21	0.15	-0.16
	[0.81]	[1.26]	[0.89]	[0.91]	[6.18]	[5.14]	[2.95]	[2.24]	[-2.55]
HXZ4	-0.01	0.03	0.04	0.07	0.31	0.32	0.24	0.18	-0.10
	[-0.18]	[0.48]	[0.45]	[0.93]	[4.90]	[4.67]	[3.02]	[2.31]	[-1.22]
HMXZ5	-0.04	-0.01	0.00	0.04	0.15	0.17	0.14	0.08	-0.05
	[-0.51]	[-0.10]	[-0.05]	[0.49]	[2.26]	[2.36]	[1.62]	[0.99]	[-0.57]
DHS3	-0.01	0.04	0.03	0.10	0.30	0.33	0.27	0.21	-0.01
	[-0.14]	[0.44]	[0.30]	[1.21]	[3.82]	[4.01]	[2.92]	[2.42]	[-0.09]
SY4	0.12	0.15	0.15	0.14	0.26	0.24	0.15	0.10	-0.09
	[1.25]	[1.56]	[1.40]	[1.41]	[3.91]	[3.26]	[1.63]	[1.18]	[-0.95]

Panel B: Monthly-rebalancing factors									
	$r^{\text{IB}}_{\text{ROEQ}}$	$r^{\text{SPI}}_{\text{ROEQ}}$	$r^{\text{EBT}}_{\text{ROEQ}}$	$r^{\text{FF}}_{\text{ROEQ}}$	$r^{\text{RD}}_{\text{ROEQ}}$	$r^{\text{RD+.3 SG}}_{\text{ROEQ}}$	$r^{\text{RD+SG}}_{\text{ROEQ}}$	$r^{\text{NM}}_{\text{ROEQ}}$	$r^{\text{REVT}}_{\text{ROEQ}}$
Mean	0.57	0.60	0.57	0.59	0.75	0.73	0.59	0.58	0.36
	[5.05]	[5.51]	[4.87]	[5.52]	[8.64]	[8.43]	[6.41]	[6.47]	[3.63]
FF5	0.39	0.40	0.36	0.32	0.60	0.52	0.36	0.34	-0.04
	[4.91]	[5.43]	[4.78]	[4.80]	[8.69]	[7.85]	[4.78]	[4.59]	[-0.59]
FF6	0.24	0.28	0.24	0.21	0.48	0.44	0.34	0.31	-0.03
	[3.69]	[4.36]	[3.60]	[3.64]	[8.11]	[7.09]	[4.48]	[4.27]	[-0.43]
FF6/Devil	0.30	0.33	0.28	0.22	0.54	0.51	0.41	0.37	-0.06
	[4.65]	[5.20]	[4.29]	[3.78]	[9.35]	[8.56]	[5.42]	[5.03]	[-0.86]
FF6/EKP	0.26	0.29	0.24	0.21	0.49	0.44	0.32	0.29	-0.11
	[3.93]	[4.49]	[3.66]	[3.52]	[8.14]	[6.79]	[4.10]	[3.83]	[-1.49]
HXZ4	0.04	0.09	0.06	0.09	0.37	0.34	0.26	0.24	-0.09
	[1.89]	[2.93]	[1.53]	[1.76]	[8.12]	[6.47]	[3.33]	[3.17]	[-1.06]
HMXZ5	0.03	0.06	0.03	0.06	0.23	0.22	0.16	0.15	-0.03
	[1.64]	[1.68]	[0.80]	[1.19]	[4.97]	[4.03]	[1.90]	[1.89]	[-0.30]
DHS3	0.21	0.26	0.21	0.23	0.49	0.44	0.34	0.31	0.04
	[1.90]	[2.45]	[1.88]	[2.31]	[5.45]	[4.88]	[3.45]	[3.35]	[0.48]
SY4	0.30	0.35	0.32	0.33	0.48	0.41	0.25	0.23	-0.01
	[3.21]	[3.81]	[3.23]	[3.46]	[6.82]	[5.14]	[2.54]	[2.44]	[-0.13]

Table 7: Pricing tests on the long- and short-side of the annual- and monthly-rebalanced profitability factors (July 1976 - December 2021).

To construct annual-rebalancing profitability factors, $r^{\mathrm{X}}_{\mathrm{ROE}}$, in the June end of each year t, firms are assigned into two size groups based on median June end market equity of NYSE firms, and assigned independently into three profitability groups based on the 30 and 70 percentiles profitability of all NYSE firms reported annual earnings in calendar year t-1. To construct monthly-rebalancing profitability factors, $r^{\mathrm{X}}_{\mathrm{ROEQ}}$, in the end of each month, firms are assigned into two size (ME) groups based on median June end market equity of NYSE firms, and assigned independently into three asset growth ($d\mathrm{AT}_t/\mathrm{AT}_{t-1}$) groups based on the 30 and 70 percentiles asset growth of all NYSE firms reported annual earnings in calendar year t-1, and assigned independently into three profitability groups based on the 30 and 70 percentiles (most recent available) profitability of all NYSE firms. The profitability factor return in any given month is the average return of the portfolios that are long the high profitability portfolios and short the low profitability portfolios within each size group. Firms' most recent available profitability is updated at the month end of the most recent quarterly earnings announcement date (quarterly Compustat data item RDQ, exclude if missing). To remove stale earnings, we exclude a firm if its most recent available quarter-end is 6 months away. The sample period is from July 1976 to December 2021 as the R&D expense data is only valid after January 1975. The quarterly R&D expenses (quarterly Compustat data item XRDQ) are only recorded after 1989 and some firms may report only once a year, we simply set missing XRDQ to zero with no further interpolation. The considered profitability measures include $Y(\mathrm{IB})/B$, $Y(\mathrm{SPI})/B$, $Y(\mathrm{EBT})/B$, $Y(\mathrm{FF})/B$, $Y(\mathrm{RD})/B$, $Y(\mathrm{RD}+.3\mathrm{SG})/B$, $Y(\mathrm{RD}+\mathrm{SG})/B$, $Y(\mathrm{NM})/B$, and $Y(\mathrm{REVT})/B$. Long- and short-side of the factors is the equal-weighted average excess return of the high- and low-profitability portfolios in each month. We run time-series regressions of the long- and short-side of the profitability factors on the existing factor models and report their intercept and associated t-statistics in the squared parenthesis. The factor models include Fama-French five factor model (FF5), FF5 plus momentum factor (FF6), FF6 with the value factor, HML, replaced by the monthly rebalancing value factor with most recent market equity, HML(Devil), from Asness and Frazzini (2013), FF6 with the value factor, HML, replaced by the intangible-adjusted value factor, HML(EKP), from Eisfeldt et al. (2022), Hou-Xue-Zhang four factor model (HXZ4) and Hou-Mo-Xue-Zhang five factor model (HMXZ5), Daniel-Hirshleifer-Sun behavioral three factor model (DHS3), and Stambaugh-Yuan four factor model (SY4). For monthly-rebalancing factors, we additionally report their time-series mean and t-statistics and the intercept and the associated t-statistics in the regression on monthly ROE factor (R_ROE) in HXZ4. The data of the factor models is taken from the authors' websites. Except for SY4 which is available only until December 2016, data on all factor models is available until December 2021.

Moving from bottom-line to top-line profitability: left to right

Panel A: Long side of the annual-rebalancing factors

	$r^{\mathrm{IB}}_{\mathrm{ROE}}$	$r^{\mathrm{SPI}}_{\mathrm{ROE}}$	$r^{\mathrm{EBT}}_{\mathrm{ROE}}$	$r^{\mathrm{FF}}_{\mathrm{ROE}}$	$r^{\mathrm{RD}}_{\mathrm{ROE}}$	$r^{\mathrm{RD}+.3\,\mathrm{SG}}_{\mathrm{ROE}}$	$r^{\mathrm{RD+SG}}_{\mathrm{ROE}}$	$r^{\mathrm{NM}}_{\mathrm{ROE}}$	$r^{\mathrm{REVT}}_{\mathrm{ROE}}$
Mean	0.87	0.88	0.90	0.92	0.96	0.98	0.96	0.95	0.92
	[4.19]	[4.22]	[4.49]	[4.51]	[4.46]	[4.57]	[4.51]	[4.40]	[4.17]
FF5	-0.01	-0.02	-0.01	-0.03	0.09	0.09	0.05	0.02	-0.09
	[-0.69]	[-0.78]	[-0.26]	[-1.33]	[3.77]	[3.27]	[1.33]	[0.47]	[-2.26]
FF6	0.00	0.00	0.01	-0.02	0.09	0.09	0.06	0.03	-0.07
	[0.16]	[0.04]	[0.58]	[-0.76]	[3.78]	[3.52]	[1.72]	[0.93]	[-1.74]
FF6/Devil	0.00	0.00	0.00	-0.04	0.11	0.12	0.09	0.05	-0.09
	[0.22]	[0.22]	[-0.12]	[-1.97]	[4.39]	[4.53]	[2.37]	[1.48]	[-2.17]
FF6/EKP	0.00	0.00	-0.01	-0.03	0.10	0.09	0.05	0.02	-0.11
	[0.04]	[0.09]	[-0.27]	[-1.37]	[3.90]	[3.22]	[1.43]	[0.59]	[-2.87]
HXZ4	0.01	0.02	0.02	0.03	0.14	0.13	0.09	0.06	-0.05
	[0.37]	[0.50]	[0.53]	[0.58]	[4.35]	[3.72]	[2.08]	[1.31]	[-0.96]
HMXZ5	0.03	0.04	0.04	0.05	0.10	0.10	0.09	0.07	0.02
	[0.82]	[0.91]	[1.01]	[1.11]	[3.04]	[2.55]	[1.95]	[1.48]	[0.28]

DHS3	0.14	0.15	0.15	0.16	0.25	0.27	0.24	0.21	0.13
	[2.51]	[2.70]	[2.69]	[2.79]	[4.30]	[4.20]	[3.42]	[3.05]	[1.60]
SY4	0.01	0.02	0.02	0.02	0.07	0.05	0.01	-0.01	-0.09
	[0.25]	[0.42]	[0.48]	[0.46]	[2.24]	[1.35]	[0.18]	[-0.27]	[-1.50]

Panel B: Long side of the monthly-rebalancing factors

	$r^{\text{IB}}_{\text{ROEQ}}$	$r^{\text{SPI}}_{\text{ROEQ}}$	$r^{\text{EBT}}_{\text{ROEQ}}$	$r^{\text{FF}}_{\text{ROEQ}}$	$r^{\text{RD}}_{\text{ROEQ}}$	$r^{\text{RD+.3 SG}}_{\text{ROEQ}}$	$r^{\text{RD+SG}}_{\text{ROEQ}}$	$r^{\text{NM}}_{\text{ROEQ}}$	$r^{\text{REVT}}_{\text{ROEQ}}$
Mean	1.09	1.11	1.10	1.10	1.14	1.15	1.08	1.08	0.99
	[5.17]	[5.20]	[5.30]	[5.23]	[5.28]	[5.27]	[4.96]	[4.94]	[4.4]
FF5	0.18	0.18	0.18	0.15	0.26	0.25	0.17	0.16	-0.02
	[5.06]	[5.41]	[5.35]	[4.43]	[7.70]	[7.50]	[4.29]	[4.05]	[-0.46]
FF6	0.14	0.15	0.15	0.13	0.23	0.23	0.17	0.16	0.00
	[4.10]	[4.60]	[4.56]	[3.68]	[6.97]	[6.90]	[4.25]	[4.05]	[-0.07]
FF6/Devil	0.16	0.17	0.17	0.13	0.25	0.26	0.20	0.19	-0.02
	[4.97]	[5.23]	[5.17]	[3.74]	[7.58]	[7.58]	[4.76]	[4.47]	[-0.44]
FF6/EKP	0.15	0.16	0.16	0.12	0.24	0.23	0.16	0.15	-0.05
	[4.23]	[4.68]	[4.70]	[3.59]	[6.85]	[6.47]	[3.84]	[3.64]	[-0.97]
HXZ4	0.10	0.13	0.13	0.12	0.23	0.23	0.18	0.18	0.01
	[3.15]	[3.80]	[3.80]	[3.17]	[6.64]	[6.53]	[4.15]	[4.03]	[0.13]
HMXZ5	0.11	0.12	0.13	0.13	0.19	0.20	0.18	0.18	0.06
	[3.22]	[3.40]	[3.62]	[3.21]	[5.09]	[5.19]	[3.71]	[3.72]	[0.98]
DHS3	0.31	0.33	0.32	0.31	0.41	0.40	0.35	0.35	0.21
	[4.61]	[4.86]	[4.79]	[4.62]	[6.18]	[5.82]	[4.87]	[4.75]	[2.46]
SY4	0.13	0.15	0.15	0.15	0.20	0.18	0.10	0.10	-0.03
	[3.42]	[3.86]	[3.71]	[3.23]	[5.16]	[4.70]	[2.20]	[2.14]	[-0.42]

Panel C: Short side of the annual-rebalancing factors

	$r^{\text{IB}}_{\text{ROE}}$	$r^{\text{SPI}}_{\text{ROE}}$	$r^{\text{EBT}}_{\text{ROE}}$	$r^{\text{FF}}_{\text{ROE}}$	$r^{\text{RD}}_{\text{ROE}}$	$r^{\text{RD+.3 SG}}_{\text{ROE}}$	$r^{\text{RD+SG}}_{\text{ROE}}$	$r^{\text{NM}}_{\text{ROE}}$	$r^{\text{REVT}}_{\text{ROE}}$
Mean	0.66	0.62	0.60	0.57	0.49	0.50	0.56	0.56	0.66
	[2.60]	[2.47]	[2.37]	[2.32]	[2.09]	[2.24]	[2.59]	[2.63]	[3.11]
FF5	-0.05	-0.07	-0.07	-0.04	-0.16	-0.17	-0.13	-0.10	0.06
	[-1.27]	[-1.89]	[-1.92]	[-2.01]	[-5.76]	[-5.72]	[-3.35]	[-2.95]	[2.00]
FF6	-0.03	-0.06	-0.06	-0.03	-0.15	-0.17	-0.14	-0.11	0.06
	[-0.90]	[-1.44]	[-1.48]	[-1.52]	[-5.24]	[-5.66]	[-3.76]	[-3.28]	[1.84]
FF6/Devil	-0.04	-0.06	-0.06	-0.04	-0.19	-0.22	-0.20	-0.17	0.04
	[-0.94]	[-1.67]	[-1.62]	[-1.99]	[-6.55]	[-7.27]	[-5.09]	[-4.54]	[1.23]
FF6/EKP	-0.04	-0.05	-0.05	-0.05	-0.18	-0.19	-0.15	-0.13	0.05
	[-0.93]	[-1.4]	[-1.22]	[-2.23]	[-5.68]	[-5.45]	[-3.54]	[-3.19]	[1.51]
HXZ4	0.03	-0.02	-0.01	-0.05	-0.17	-0.19	-0.15	-0.12	0.05
	[0.51]	[-0.32]	[-0.28]	[-1.06]	[-4.04]	[-4.47]	[-3.25]	[-2.79]	[1.32]
HMXZ5	0.07	0.04	0.05	0.01	-0.05	-0.07	-0.05	-0.01	0.07
	[1.32]	[0.79]	[0.83]	[0.24]	[-1.04]	[-1.68]	[-0.94]	[-0.27]	[1.60]
DHS3	0.15	0.11	0.12	0.06	-0.05	-0.06	-0.03	0.00	0.14
	[1.63]	[1.26]	[1.34]	[0.73]	[-0.54]	[-0.80]	[-0.35]	[0.00]	[2.13]
SY4	-0.11	-0.13	-0.12	-0.12	-0.18	-0.19	-0.14	-0.12	0.00
	[-1.58]	[-1.88]	[-1.73]	[-1.84]	[-3.54]	[-3.66]	[-2.33]	[-2.11]	[0.08]

Panel D: Short side of the monthly-rebalancing factors									
	$r^{\mathrm{IB}}_{\mathrm{ROEQ}}$	$r^{\mathrm{SPI}}_{\mathrm{ROEQ}}$	$r^{\mathrm{EBT}}_{\mathrm{ROEQ}}$	$r^{\mathrm{FF}}_{\mathrm{ROEQ}}$	$r^{\mathrm{RD}}_{\mathrm{ROEQ}}$	$r^{\mathrm{RD+.3\,SG}}_{\mathrm{ROEQ}}$	$r^{\mathrm{RD+SG}}_{\mathrm{ROEQ}}$	$r^{\mathrm{NM}}_{\mathrm{ROEQ}}$	$r^{\mathrm{REVT}}_{\mathrm{ROEQ}}$
Mean	0.52	0.51	0.54	0.51	0.39	0.42	0.49	0.50	0.63
	[2.05]	[2.02]	[2.10]	[2.06]	[1.65]	[1.82]	[2.21]	[2.29]	[2.92]
FF5	-0.21	-0.22	-0.18	-0.16	-0.33	-0.27	-0.19	-0.17	0.02
	[-3.51]	[-3.77]	[-3.05]	[-3.22]	[-6.41]	[-5.37]	[-3.61]	[-3.42]	[0.5]
FF6	-0.11	-0.13	-0.09	-0.09	-0.25	-0.21	-0.17	-0.15	0.03
	[-2.10]	[-2.53]	[-1.68]	[-1.91]	[-5.49]	[-4.47]	[-3.21]	[-2.96]	[0.65]
FF6/Devil	-0.13	-0.15	-0.11	-0.09	-0.28	-0.25	-0.21	-0.18	0.04
	[-2.55]	[-2.98]	[-2.04]	[-2.01]	[-6.27]	[-5.51]	[-4.07]	[-3.65]	[0.96]
FF6/EKP	-0.12	-0.13	-0.09	-0.08	-0.26	-0.21	-0.16	-0.13	0.06
	[-2.21]	[-2.54]	[-1.61]	[-1.80]	[-5.54]	[-4.33]	[-2.97]	[-2.63]	[1.43]
HXZ4	0.06	0.03	0.07	0.04	-0.14	-0.11	-0.07	-0.06	0.10
	[2.00]	[0.95]	[1.74]	[0.92]	[-3.77]	[-2.61]	[-1.43]	[-1.22]	[2.00]
HMXZ5	0.08	0.06	0.10	0.07	-0.04	-0.02	0.02	0.03	0.09
	[2.23]	[1.75]	[2.35]	[1.70]	[-1.06]	[-0.53]	[0.39]	[0.5]	[1.6]
DHS3	0.10	0.07	0.11	0.08	-0.08	-0.04	0.02	0.03	0.16
	[0.98]	[0.68]	[1.00]	[0.78]	[-0.85]	[-0.45]	[0.19]	[0.38]	[2.13]
SY4	-0.17	-0.20	-0.17	-0.18	-0.28	-0.22	-0.15	-0.13	-0.01
	[-2.12]	[-2.54]	[-2.08]	[-2.42]	[-4.28]	[-3.21]	[-1.93]	[-1.78]	[-0.17]

Table 8: Pricing tests on annual-rebalanced profitability factors varying the fraction of SGA (July 1976 - December 2021).

To construct annual-rebalanced profitability factors, in the June end of each year t, firms are assigned into two size groups based on median June end market equity of NYSE firms, and assigned independently into three profitability groups based on the 30 and 70 percentiles profitability of all NYSE firms reported annual earnings in calendar year t-1. The profitability factor return in any given month is the average return of the portfolios that are long the high profitability portfolios and short the low profitability portfolios within each size group. The sample period is from July 1976 to December 2021 as the R&D expense data is only valid after January 1975. We consider the profitability measures of $Y(\gamma)/B$ and $Y(\gamma) = (\text{REVT} - \text{COGS} - \text{XSGA} - \text{XINT} + \frac{\gamma}{10}\text{SGA} + \text{XRD})$ where $\gamma = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$. We run time-series regressions of the constructed profitability factors on the existing factor models and report their intercept and associated t-statistics in the squared parenthesis. The factor models include Fama-French five factor model (FF5), FF5 plus momentum factor (FF6), FF6 with the value factor, HML, replaced by the monthly rebalancing value factor with most recent market equity, HML(Devil), from Asness and Frazzini (2013), FF6 with the value factor, HML, replaced by the intangible-adjusted value factor, HML(EKP), from Eisfeldt et al. (2022), Hou-Xue-Zhang four factor model (HXZ4) and Hou-Mo-Xue-Zhang five factor model (HMXZ5), Daniel-Hirshleifer-Sun behavioral three factor model (DHS3), and Stambaugh-Yuan four factor model (SY4). For monthly-rebalancing factors, we additionally report their time-series mean and t-statistics and the intercept and the associated t-statistics in the regression on monthly ROE factor (R_ROE) in HXZ4. The data of the factor models is taken from the authors' websites. Except for SY4 which is available only until December 2016, data on all factor models is available until December 2021.

	0% ($r^{\text{RD}}_{\text{ROE}}$)	10%	20%	30% ($r^{\text{RD+.3 SG}}_{\text{ROE}}$)	40%	50%	60%	70%	80%	90%	100% ($\approx r^{\text{RD+SG}}_{\text{ROE}}$)
Mean	0.47	0.48	0.49	0.48	0.48	0.47	0.45	0.45	0.44	0.43	0.42
	[6.11]	[6.33]	[6.42]	[6.26]	[6.05]	[5.87]	[5.53]	[5.54]	[5.34]	[5.18]	[5.00]
FF5	0.25	0.26	0.27	0.26	0.24	0.24	0.21	0.22	0.21	0.20	0.19
	[6.44]	[6.41]	[6.31]	[5.55]	[4.94]	[4.52]	[3.92]	[3.86]	[3.56]	[3.27]	[3.03]
FF6	0.24	0.25	0.27	0.26	0.25	0.25	0.23	0.24	0.23	0.22	0.22
	[6.05]	[6.23]	[6.32]	[5.65]	[5.10]	[4.76]	[4.18]	[4.21]	[3.89]	[3.64]	[3.47]
FF6/Devil	0.30	0.32	0.35	0.35	0.33	0.33	0.31	0.33	0.32	0.31	0.30
	[7.25]	[7.66]	[7.85]	[7.19]	[6.57]	[6.21]	[5.60]	[5.52]	[5.18]	[4.93]	[4.69]
FF6/EKP	0.28	0.28	0.30	0.29	0.27	0.27	0.24	0.25	0.24	0.23	0.22
	[6.18]	[5.81]	[5.74]	[5.14]	[4.66]	[4.34]	[3.81]	[3.83]	[3.56]	[3.33]	[3.15]
HXZ4	0.31	0.32	0.33	0.32	0.31	0.30	0.27	0.28	0.27	0.26	0.26
	[4.90]	[4.90]	[5.02]	[4.67]	[4.26]	[4.07]	[3.67]	[3.68]	[3.50]	[3.30]	[3.21]
HMXZ5	0.15	0.16	0.18	0.17	0.17	0.17	0.15	0.16	0.16	0.15	0.15
	[2.26]	[2.34]	[2.60]	[2.36]	[2.22]	[2.17]	[1.88]	[1.98]	[1.87]	[1.74]	[1.71]
DHS3	0.30	0.32	0.33	0.33	0.32	0.31	0.30	0.31	0.29	0.29	0.28
	[3.82]	[3.98]	[4.12]	[4.01]	[3.77]	[3.64]	[3.39]	[3.46]	[3.26]	[3.17]	[3.08]
SY4	0.26	0.26	0.26	0.24	0.22	0.21	0.19	0.19	0.17	0.16	0.16
	[3.91]	[3.81]	[3.77]	[3.26]	[2.79]	[2.64]	[2.28]	[2.25]	[2.04]	[1.87]	[1.77]

Table 9: Mean and t-statistics of the monthly factor returns for the full sample and two subsamples (July 1976 - December 2021).

RMW, robust-minus-weak, is the annual-rebalanced profitability factor in Fama-French five factor model. R_ROE is the monthly-rebalnaced profitability factor in Hou-Xue-Zhang four factor model. $r^{\text{RD}}_{\text{ROE}}$ and $r^{\text{RD}}_{\text{ROEQ}}$ are self constructed annual- and monthly-rebalanced profitability factors based on the profitability measures $Y(\text{RD})/B$ where $Y(\text{RD}) = \text{REVT} - \text{COGS} - \text{XSGA} - \text{XINT} + \text{XRD}$ and their respective quarterly counterparts. PEAD is the post earnings announcement drift factor from Daniel-Hirshleifer-Sun behavioral three factor model. PERF is the performance factor from Stambaugh-Yuan four factor model (SY4). UMD is the momentum factor taken from Ken French's data library. UMD(d) is the decile momentum factor which is long the highest prior return decile portfolio and short the lowest prior return decile portfolio from Ken French's data library. The data of the factor models are taken from the authors' websites. Except for PERF, which is available only until December 2016, data on all factor models is available until December 2021.

	RMW	$r^{\text{RD}}_{\text{ROE}}$	R_ROE	$r^{\text{RD}}_{\text{ROEQ}}$	PEAD	PERF	UMD	UMD(d)
	Full Sample: July 1976 to December 2021							
Mean	0.35	0.47	0.55	0.75	0.58	0.68	0.59	1.04
t-stat	[3.54]	[6.11]	[4.95]	[8.64]	[7.12]	[3.79]	[3.17]	[3.22]
	July 1976 to December 1999							
Mean	0.27	0.34	0.77	0.85	0.79	0.64	1.00	1.71
t-stat	[2.92]	[3.93]	[6.62]	[8.30]	[8.51]	[3.67]	[5.07]	[5.37]
	January 2000 to December 2021							
Mean	0.45	0.60	0.33	0.64	0.35	0.74	0.16	0.31
t-stat	[2.45]	[4.70]	[1.68]	[4.51]	[2.59]	[2.08]	[0.50]	[0.55]

Table 10: Pricing tests of $r^{\mathrm{RD}}_{\mathrm{ROE}}$ on related factors and anomalies and the pricing tests of related factors and anomalies on $r^{\mathrm{RD}}_{\mathrm{ROE}}$ (July 1976 - December 2021).

RMW, robust-minus-weak, is the annual-rebalanced profitability factor in Fama-French five factor model. R_ROE is the monthly-rebalnacing profitability factor in Hou-Xue-Zhang four factor model. PEAD is the post earnings announcement drift factor from Daniel-Hirshleifer-Sun behavioral three factor model. PERF is the performance factor from Stambaugh-Yuan four factor model (SY4). UMD is the momentum factor taken from Ken French's data library. UMD(d) is the decile momentum factor which is long the highest prior return decile portfolio and short the lowest prior return decile portfolio from Ken French's data library. HML(EKP) is the intangible-adjusted value factor by Eisfeldt et al. (2022). $r^{\mathrm{NM}}_{\mathrm{ROA}}$ is self constructed annual-rebalanced gross profitability factor following the definition from Novy-Marx (2013) where the profitability is $Y(\mathrm{NM})/AT$ where $Y(\mathrm{NM}) = \mathrm{REVT} - \mathrm{COGS}$. $r^{\mathrm{RD}}_{\mathrm{ROE}}$ is self constructed annual-profitability factors based on the profitability measure $Y(\mathrm{RD})/B$ where $Y(\mathrm{RD}) = \mathrm{REVT} - \mathrm{COGS} - \mathrm{XSGA} - \mathrm{XINT} + \mathrm{XRD}$. $r^{\mathrm{NM}}_{\mathrm{OL}}$ is self constructed annual-rebalanced operating leverage factor following the definition from Novy-Marx (2011) where the operating leverage is $(\mathrm{COGS} + \mathrm{XSGA})/AT$. Except for $r^{\mathrm{NM}}_{\mathrm{ROA}}$, $r^{\mathrm{RD}}_{\mathrm{ROE}}$, and $r^{\mathrm{NM}}_{\mathrm{OL}}$, the data is taken from the authors' websites. Except for PERF which is available only until December 2016, data on all factors is available until December 2021. The table reports the intercept and slope coefficient with the associated t-statistics and r-squared of the univariate time-series regressions. $r^{\mathrm{RD}}_{\mathrm{ROE}}$ is the dependent and independent variables in the Panel A and B respectively.

Panel A: Explaining intangible-adjusted profitability factor using factor X, $r^{\mathrm{RD}}_{\mathrm{ROE}} = a + b\,X + \epsilon$

	RMW	$r^{\mathrm{NM}}_{\mathrm{ROA}}$	R_ROE	PEAD	PERF	UMD	UMD(d)	$r^{\mathrm{NM}}_{\mathrm{OL}}$	HML(EKP)
a	0.25	0.37	0.23	0.42	0.23	0.41	0.40	0.41	0.51
t(a)	[5.38]	[5.43]	[3.80]	[5.32]	[3.67]	[5.48]	[5.42]	[5.60]	[6.59]
b	0.61	0.37	0.42	0.07	0.28	0.09	0.06	0.27	-0.10
t(b)	[30.74]	[12.71]	[18.15]	[1.86]	[18.01]	[5.48]	[6.08]	[7.47]	[-2.70]
R^2	0.63	0.23	0.38	0.01	0.40	0.05	0.06	0.09	0.01

Panel B: Explaining factor X using intangible-adjusted profitability factor, $X = a + b\,r^{\mathrm{RD}}_{\mathrm{ROE}} + \epsilon$

	RMW	$r^{\mathrm{NM}}_{\mathrm{ROA}}$	R_ROE	PEAD	PERF	UMD	UMD(d)	$r^{\mathrm{NM}}_{\mathrm{OL}}$	HML(EKP)
a	-0.13	-0.02	0.13	0.54	0.08	0.33	0.54	0.05	0.49
t(a)	[-2.13]	[-0.21]	[1.46]	[6.43]	[0.55]	[1.75]	[1.68]	[0.58]	[5.73]
b	1.05	0.62	0.90	0.08	1.45	0.56	1.06	0.34	-0.13
t(b)	[30.74]	[12.71]	[18.15]	[1.86]	[18.01]	[5.48]	[6.08]	[7.47]	[-2.70]
R^2	0.63	0.23	0.38	0.01	0.40	0.05	0.06	0.09	0.01

Panel C: Explaining factor X using Fama-French profitability factor, $X = a + b\,\mathrm{RMW} + \epsilon$

	$r^{\mathrm{NM}}_{\mathrm{ROA}}$	R_ROE	PEAD	PERF	UMD	UMD(d)	$r^{\mathrm{NM}}_{\mathrm{OL}}$	HML(EKP)
a	0.22	0.28	0.60	0.43	0.54	0.89	0.17	0.40
t(a)	[3.28]	[3.43]	[7.28]	[2.65]	[2.85]	[2.77]	[1.97]	[4.75]
b	0.48	0.76	-0.05	0.71	0.15	0.40	0.11	0.10
t(b)	[16.6]	[21.65]	[-1.54]	[10.41]	[1.92]	[2.93]	[3.09]	[2.73]
R^2	0.34	0.46	0.00	0.18	0.01	0.02	0.02	0.01

Table 11: Summary statistics and performance of within-industry profitability factors (July 1976 - December 2021).

Firms are assigned into industry groups following Fama-French five industry definition. In Panel A, N(firms) is the time-series average of the number of firms, RD>0 is the time-series average fraction of firms that report positive R&D expense, RD/B is the time-series average total R&D expense divided by total book equity, RD/REVT is the time-series average total R&D expense divided by total revenue. In Panel B, we form within-industry profitability factors based on various profitability measures including Y(IB)/B, Y(SPI)/B, Y(EBT)/B, Y(FF)/B, Y(RD)/B, Y(RD + .3 SG)/B, Y(RD+SG)/B, Y(NM)/B, and Y(REVT)/B from the bottom-line to the top-line earnings. Specifically, in the June end of each year t, firms within each industry are assigned into two size (ME) groups based on median June end market equity of NYSE firms within the given industry, and assigned independently into three profitability groups based on the 30 and 70 percentiles profitability of all NYSE firms within the given industry reported annual earnings in calendar year t-1. The profitability factor return within each industry in any given month is the average return of the portfolios that are long the high profitability portfolios and short the low profitability portfolios within each size group.

Panel A: Summary statistics

	N(firms)	RD > 0	RD/B	RD/REVT
Cnsmr	804	0.21	0.03	0.01
Manuf	930	0.42	0.02	0.01
HiTec	739	0.70	0.07	0.04
Hlth	383	0.75	0.13	0.08
Other	1130	0.08	0.00	0.00

Panel B: Profitability factors formed on within-industry Y(X)/B

	Monthly Percentage Returns							
	IB	SPI	EBT	FF	RD	RD+.3 SG	NM	REVT
Cnsmr	0.15	0.06	0.23	0.37	0.46	0.47	0.39	0.30
Manuf	0.21	0.26	0.36	0.37	0.42	0.44	0.39	0.37
HiTec	0.11	0.07	0.01	0.15	0.29	0.26	0.22	0.21
Hlth	-0.06	0.07	0.04	0.11	0.36	0.47	0.42	0.15
Other	0.32	0.33	0.38	0.37	0.38	0.35	0.28	0.21
	t-statistics							
	IB	SPI	EBT	FF	RD	RD+.3 SG	NM	REVT
Cnsmr	1.15	0.47	1.77	3.53	4.68	5.21	3.85	3.19
Manuf	1.72	2.19	2.92	3.58	4.20	4.01	3.12	2.98
HiTec	0.74	0.51	0.09	0.95	2.38	2.26	2.16	1.74
Hlth	-0.29	0.36	0.20	0.53	2.21	2.78	2.23	0.82
Other	3.29	3.41	3.41	4.18	4.50	4.24	3.15	1.92

Table 12: Model comparison tests.

The R&D-adjusted profitability factors (annual-rebalanced $r^{\text{RD}}_{\text{ROE}}$ and monthly-rebalanced $r^{\text{RD}}_{\text{ROEQ}}$) are incorporated into the existing models in two ways. First, for models with profitability related factors, the original profitability factor in the model is replaced with the R&D-adjusted profitability factors. Second, for models without profitability related factors, the models are augmented with the R&D-adjusted profitability factors. The squared Sharpe ratio test by Barillas et al. (2020) is used to compare the original models and the models with the R&D-adjusted profitability factors. The first row of each panel is the difference in the squared Sharpe ratios between the column model and the row model, and the second row report the p-value of the test with the null hypothesis that the difference in the squared Sharpe ratio is zero. The following factor models are considered: Fama-French five factor model (FF5), FF5 plus momentum factor (FF6), FF6 with HML replaced by the monthly-rebalancing value factor using the monthly updated market equity HML(Devil) (FF6^1), FF6 with HML replaced by the intangible-adjusted value factor HML(EKP) (FF6^2), Hou-Xue-Zhang four factor model (HXZ4) and Hou-Mo-Xue-Zhang five factor model (HMXZ5), Daniel-Hirshleifer-Sun behavioral three factor model (DHS3).

July 1976 - December 2021			July 1976 - December 1999			January 2000 - December 2021		
	FF5/$r^{\text{RD}}_{\text{ROE}}$	FF5+$r^{\text{RD}}_{\text{ROE}}$		FF5/$r^{\text{RD}}_{\text{ROE}}$	FF5+$r^{\text{RD}}_{\text{ROE}}$		FF5/$r^{\text{RD}}_{\text{ROE}}$	FF5+$r^{\text{RD}}_{\text{ROE}}$
FF5	0.067	0.082	FF5	0.069	0.070	FF5	0.073	0.078
	0.000	*0.000*		*0.020*	*0.000*		*0.004*	*0.000*
	FF6/$r^{\text{RD}}_{\text{ROE}}$	FF6+$r^{\text{RD}}_{\text{ROE}}$		FF6/$r^{\text{RD}}_{\text{ROE}}$	FF6+$r^{\text{RD}}_{\text{ROE}}$		FF6/$r^{\text{RD}}_{\text{ROE}}$	FF6+$r^{\text{RD}}_{\text{ROE}}$
FF6	0.059	0.074	FF6	0.060	0.059	FF6	0.072	0.076
	0.000	*0.000*		*0.033*	*0.002*		*0.004*	*0.000*
	FF6^1/$r^{\text{RD}}_{\text{ROE}}$	FF6^1+$r^{\text{RD}}_{\text{ROE}}$		FF6^1/$r^{\text{RD}}_{\text{ROE}}$	FF6^1+$r^{\text{RD}}_{\text{ROE}}$		FF6^1/$r^{\text{RD}}_{\text{ROE}}$	FF6^1+$r^{\text{RD}}_{\text{ROE}}$
FF6^1	0.078	0.107	FF6^1	0.079	0.078	FF6^1	0.088	0.104
	0.000	*0.000*		*0.029*	*0.000*		*0.001*	*0.000*
	FF6^2/$r^{\text{RD}}_{\text{ROE}}$	FF6^2+$r^{\text{RD}}_{\text{ROE}}$		FF6^2/$r^{\text{RD}}_{\text{ROE}}$	FF6^2+$r^{\text{RD}}_{\text{ROE}}$		FF6^2/$r^{\text{RD}}_{\text{ROE}}$	FF6^2+$r^{\text{RD}}_{\text{ROE}}$
FF6^2	0.062	0.079	FF6^2	0.036	0.032	FF6^2	0.097	0.116
	0.000	*0.000*		*0.116*	*0.007*		*0.001*	*0.000*
	HXZ4/$r^{\text{RD}}_{\text{ROEQ}}$	HXZ4+$r^{\text{RD}}_{\text{ROEQ}}$		HXZ4/$r^{\text{RD}}_{\text{ROEQ}}$	HXZ4+$r^{\text{RD}}_{\text{ROEQ}}$		HXZ4/$r^{\text{RD}}_{\text{ROEQ}}$	HXZ4+$r^{\text{RD}}_{\text{ROEQ}}$
HXZ4	0.106	0.136	HXZ4	0.146	0.147	HXZ4	0.093	0.133
	0.000	*0.000*		*0.004*	*0.000*		*0.002*	*0.000*
	HMXZ5/$r^{\text{RD}}_{\text{ROEQ}}$	HMXZ5+$r^{\text{RD}}_{\text{ROEQ}}$		HMXZ5/$r^{\text{RD}}_{\text{ROEQ}}$	HMXZ5+$r^{\text{RD}}_{\text{ROEQ}}$		HMXZ5/$r^{\text{RD}}_{\text{ROEQ}}$	HMXZ5+$r^{\text{RD}}_{\text{ROEQ}}$
HMXZ5	0.028	0.057	HMXZ5	0.042	0.040	HMXZ5	0.035	0.068
	0.072	*0.000*		*0.131*	*0.009*		*0.110*	*0.000*
	DHS3+$r^{\text{RD}}_{\text{ROE}}$	DHS3+$r^{\text{RD}}_{\text{ROEQ}}$		DHS3+$r^{\text{RD}}_{\text{ROE}}$	DHS3+$r^{\text{RD}}_{\text{ROEQ}}$		DHS3+$r^{\text{RD}}_{\text{ROE}}$	DHS3+$r^{\text{RD}}_{\text{ROEQ}}$
DHS3	0.030	0.064	DHS3	0.048	0.155	DHS3	0.069	0.061
	0.000	*0.000*		*0.002*	*0.000*		*0.000*	*0.000*
July 1976 - December 2016			July 1976 - December 1999			January 2000 - December 2016		
	SY4/$r^{\text{RD}}_{\text{ROE}}$	SY4+$r^{\text{RD}}_{\text{ROE}}$		SY4/$r^{\text{RD}}_{\text{ROE}}$	SY4+$r^{\text{RD}}_{\text{ROE}}$		SY4/$r^{\text{RD}}_{\text{ROE}}$	SY4+$r^{\text{RD}}_{\text{ROE}}$
SY4	0.116	0.115	SY4	0.288	0.282	SY4	0.055	0.049
	0.001	*0.000*		*0.001*	*0.000*		*0.126*	*0.003*

Table 13: Maximum drawdown of the profitability factors (July 1976 - December 2021).

To compute the drawdowns for a factor, we keep tracking the cumulative returns of the factor portfolio and identify the episodes where the cumulative returns are below a historical peak. The episodes are identified independently across factors. The drawdowns are the cumulative percentage returns from the historical peak to the trough within each episode. The sample period is from July 1976 to December 2021. Panel A and B report the ten maximum drawdowns of each annual- and monthly-rebalanced profitability factors respectively.

Panel A: Annual-rebalanced profitability factors

Rank	$r_{\text{ROE}}^{\text{IB}}$	$r_{\text{ROE}}^{\text{SPI}}$	$r_{\text{ROE}}^{\text{EBT}}$	$r_{\text{ROE}}^{\text{FF}}$	$r_{\text{ROE}}^{\text{RD}}$	$r_{\text{ROE}}^{\text{RD+.3SG}}$	$r_{\text{ROE}}^{\text{RD+SG}}$	$r_{\text{ROE}}^{\text{NM}}$	$r_{\text{ROE}}^{\text{REVT}}$
1	-36.72	-34.35	-45.76	-42.25	-14.97	-12.35	-20.57	-21.36	-41.95
2	-25.76	-23.88	-26.92	-24.26	-12.05	-12.07	-20.16	-19.77	-14.27
3	-21.19	-18.87	-21.17	-12.97	-10.62	-10.70	-15.49	-12.86	-11.27
4	-19.06	-18.15	-18.54	-12.66	-9.64	-8.93	-10.71	-9.94	-9.69
5	-17.55	-17.39	-16.04	-12.34	-8.42	-7.77	-8.91	-9.92	-9.04
6	-17.52	-12.19	-13.87	-10.02	-6.24	-7.56	-7.75	-7.64	-8.28
7	-12.35	-10.57	-9.38	-6.67	-6.12	-7.11	-6.81	-7.13	-6.86
8	-8.76	-9.05	-9.15	-6.13	-5.99	-6.73	-6.46	-7.00	-5.59
9	-8.53	-7.06	-8.65	-5.67	-5.73	-5.72	-6.38	-6.87	-5.05
10	-8.20	-4.54	-7.53	-5.56	-5.65	-5.51	-6.28	-6.44	-4.77

Panel B: Monthly-rebalanced profitability factors

Rank	$r_{\text{ROEQ}}^{\text{IB}}$	$r_{\text{ROEQ}}^{\text{SPI}}$	$r_{\text{ROEQ}}^{\text{EBT}}$	$r_{\text{ROEQ}}^{\text{FF}}$	$r_{\text{ROEQ}}^{\text{RD}}$	$r_{\text{ROEQ}}^{\text{RD+.3SG}}$	$r_{\text{ROEQ}}^{\text{RD+SG}}$	$r_{\text{ROEQ}}^{\text{NM}}$	$r_{\text{ROEQ}}^{\text{REVT}}$
1	-30.89	-25.21	-34.75	-34.97	-18.30	-15.67	-25.96	-24.27	-43.28
2	-25.47	-24.09	-26.74	-22.31	-15.45	-11.80	-14.83	-12.57	-15.34
3	-24.80	-23.95	-25.36	-22.09	-9.87	-10.31	-10.23	-8.70	-14.48
4	-24.75	-21.87	-24.58	-16.87	-9.32	-9.95	-8.49	-8.37	-13.28
5	-12.73	-13.00	-13.10	-11.30	-8.99	-9.44	-8.26	-8.14	-9.59
6	-9.30	-12.14	-12.44	-9.28	-8.74	-6.49	-8.08	-7.68	-8.97
7	-7.38	-8.82	-10.23	-9.05	-8.13	-6.03	-7.68	-7.42	-8.43
8	-7.26	-8.11	-7.69	-8.13	-4.46	-5.93	-7.23	-7.09	-7.65
9	-6.42	-7.01	-6.57	-5.36	-4.17	-5.57	-6.44	-6.68	-6.66
10	-5.97	-6.49	-6.19	-5.28	-3.62	-5.44	-6.19	-6.56	-4.74

Table 14: Co-tail risk of factors with the market (July 1976 - December 2021).

Panel A reports the time-series average of the monthly percentage return of each factor and the associated t-statistics conditional on the market performance. r_{MKT} is the market excess return plus the risk-free rate in the Ken French's data library. N(months) is the total number of months that satisfy the given condition on market performance. The considered market conditions include months with market returns that are > 0, < 0, < -5, < -7, and < -10. The considered factors include the annual- and monthly-rebalanced profitability factors formed based on $Y(\text{IB})/B$, $Y(\text{FF})/B$, $Y(\text{RD})/B$, and $Y(\text{NM})/B$. Panel B reports the percentage return of each factor in each of the nine months in which the market return is lower than -10%.

Panel A: Subsamples

	N(months)	$r^{\text{IB}}_{\text{ROE}}$	$r^{\text{FF}}_{\text{ROE}}$	$r^{\text{RD}}_{\text{ROE}}$	$r^{\text{NM}}_{\text{ROE}}$	$r^{\text{IB}}_{\text{ROEQ}}$	$r^{\text{FF}}_{\text{ROEQ}}$	$r^{\text{RD}}_{\text{ROEQ}}$	$r^{\text{NM}}_{\text{ROEQ}}$
Full sample	546	0.19	0.34	0.45	0.37	0.57	0.59	0.75	0.58
		[1.88]	[3.24]	[5.87]	[4.50]	[5.05]	[5.52]	[8.64]	[6.47]
$r_{\text{MKT}} > 0$	352	-0.26	-0.09	0.28	0.42	0.27	0.34	0.68	0.65
		[-2.14]	[-0.76]	[3.29]	[4.18]	[1.94]	[2.53]	[6.21]	[5.86]
$r_{\text{MKT}} < 0$	194	1.02	1.12	0.75	0.28	1.11	1.05	0.87	0.45
		[5.64]	[6.08]	[5.11]	[1.95]	[6.08]	[5.97]	[6.17]	[2.97]
$r_{\text{MKT}} < -5$	46	2.42	2.50	1.75	0.52	2.65	2.18	1.70	0.58
		[5.85]	[5.41]	[5.39]	[1.53]	[5.74]	[4.54]	[4.66]	[1.58]
$r_{\text{MKT}} < -7$	27	2.40	2.49	2.00	0.66	2.67	1.94	1.74	0.38
		[4.05]	[3.68]	[4.35]	[1.38]	[4.02]	[2.88]	[3.27]	[0.77]
$r_{\text{MKT}} < -10$	9	4.04	3.40	2.99	0.86	3.42	2.19	2.13	0.24
		[3.62]	[2.25]	[3.23]	[0.72]	[2.72]	[1.73]	[2.14]	[0.21]

Panel B: Months with $r_{\text{MKT}} < -10$

Date	r_{MKT}	$r^{\text{IB}}_{\text{ROE}}$	$r^{\text{FF}}_{\text{ROE}}$	$r^{\text{RD}}_{\text{ROE}}$	$r^{\text{NM}}_{\text{ROE}}$	$r^{\text{IB}}_{\text{ROEQ}}$	$r^{\text{FF}}_{\text{ROEQ}}$	$r^{\text{RD}}_{\text{ROEQ}}$	$r^{\text{NM}}_{\text{ROEQ}}$
197810	-11.23	1.17	0.66	-0.29	-3.26	-2.27	-0.69	-0.69	-4.31
198003	-11.69	5.18	1.59	1.85	2.49	-0.85	-1.86	-1.86	0.91
198710	-22.64	3.89	4.33	1.91	-3.93	0.66	1.46	1.46	-3.67
199808	-15.65	2.87	3.47	2.18	0.19	3.26	1.22	0.24	-1.52
200011	-10.21	11.82	14.53	9.78	7.98	9.97	11.40	8.01	6.18
200209	-10.21	3.37	2.96	1.03	-0.08	4.45	2.63	0.98	0.76
200810	-17.15	5.72	3.27	3.99	-0.91	6.14	1.98	3.84	-0.93
200902	-10.09	0.54	1.40	3.54	4.05	3.94	2.99	3.91	4.22
202003	-13.26	0.91	-2.15	2.01	0.82	5.52	0.56	3.28	0.49

Table 15: Variable definitions.

Variable	Definition
REVT	Revenue (quarterly equivalent: REVTQ), Data source: Compustat
COGS	Cost of good sold (quarterly equivalent: COGSQ), Data source: Compustat
XSGA	Selling, general and administrative expense (quarterly equivalent: XSGAQ), Data source: Compustat
XRD	Research and development expense (quarterly equivalent: XRDQ), Data source: Compustat
DP	Depreciation (quarterly equivalent: DPQ), Data source: Compustat
XINT	Interest and related expense (quarterly equivalent: XINTQ), Data source: Compustat
SPI	Special items (quarterly equivalent: SPIQ), Data source: Compustat
IB	Income before extraordinary items (quarterly equivalent: IBQ), Data source: Compustat
CEQ	Common/ordinary equity, Data source: Compustat
AT	Total assets, Data source: Compustat
DVC	Common/ordinary dividends, Data source: Compustat
CSHO	Common shares outstanding, Data source: Compustat
RDQ	Report date of quarterly earnings, Data source: Compustat
PRC	Price, Data source: CRSP
SHROUT	Number of shares outstanding, Data source: CRSP
PRIMEXCH	Primary exchange, Data source: CRSP
SHRCD	Share code, Data source: CRSP
SGA	XSGA excluding research and development related expense following Peters and Taylor (2017)
RD	Equal to XRD
ME	Market equity: PRC*SHROUT. For factor construction and cross-section return regressions, we use June-end market equity; while for cross-section earnings forecasting regressions, we use December-end market equity
B	Book equity, equal to CEQ
BQ	Quarterly book equity computed following Hou et al. (2015)
B^{PT}	Intangible-adjusted book equity following Peters and Taylor (2017), data source: `https://wrds-www.wharton.upenn.edu/`, date visited: September 24, 2022
BM	Book-to-market ratio, B/ME
AG	Asset growth, dAT(t)/AT(t-1)
$Y(\mathrm{REVT})$	REVT
$Y(\mathrm{NM})$	REVT-COGS, following Novy-Marx (2013)
$Y(\mathrm{RD}+\mathrm{SG})$	REVT-COGS-XINT, following Eisfeldt et al. (2022) definition of investment in intangibles
$Y(\mathrm{RD}+.3\,SG)$	REVT-COGS-XSGA-XINT+0.3*SGA+XRD, following Peters and Taylor (2017) definition of investment in intangibles
$Y(\mathrm{RD})$	REVT-COGS-XSGA-XINT+XRD
$Y(\mathrm{FF})$	REVT-COGS-XSGA-XINT, following Fama and French (2015) instruction of RMW factor
$Y(\mathrm{EBT})$	REVT-COGS-XSGA-XINT-DP
$Y(\mathrm{SPI})$	IB-SPI
$Y(\mathrm{IB})$	IB, following Fama and French (2006) using annual earnings IB and following Hou et al. (2015) using quarterly earnings IBQ
FF5	Fama-French five factor model, Fama and French (2015), data source: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`, date visited: August 10, 2022
FF6	Fama-French five factor model plus momentum factor, Fama and French (2016)
HXZ4	Hou-Xue-Zhang four factor model, Hou et al. (2015), data source: `https://global-q.org/factors.html`, date visited: August 10, 2022
HMXZ5	Hou-Mo-Xue-Zhang five factor model, HXZ4 plus expected growth factor, Hou et al. (2021), data source: `https://global-q.org/factors.html`, date visited: August 10, 2022
DHS3	Daniel-Hirshleifer-Sun three factor model, Daniel et al. (2020), data source: `https://sites.google.com/view/linsunhome`, date visited: August 10, 2022
SY4	Stambaugh-Yuan four factor model Stambaugh and Yuan (2017), available up to the end of 2016, data source: `https://finance.wharton.upenn.edu/~stambaug/`, date visited: August 10, 2022
UMD	Momentum factor, data source: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`, date visited: August 10, 2022
UMD(d)	(10-1) decile momentum factor, data source: `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`, date visited: August 10, 2022
RMW	Robust-minus-weak factor in FF5
PEAD	Post-earnings announcement drift factor from DHS3
PERF	Performance factor from SY4
R_ROE	Quarterly ROE factor from HXZ4
HML(EKP)	Intangible-adjusted value factor, Eisfeldt et al. (2022), data source: `https://github.com/edwardtkim/intangiblevalue`, date visited: August 10, 2022
HML(Devil)	Monthly-rebalanced HML factor with most recent market equity, Asness and Frazzini (2013), data source: `https://www.aqr.com/Insights/Datasets/The-Devil-in-HMLs-Details-Factors-Monthly`, date visited: March 1, 2023
$r^{\mathrm{NM}}_{\mathrm{OL}}$	Operating leverage factor, Novy-Marx (2011), data are computed
$r^{\mathrm{NM}}_{\mathrm{ROA}}$	Gross profitability factor, Novy-Marx (2013), data are computed

Appendix

Table 16: Number of firms after τ years (1988 - 2021).

This table reports the time-series average of the number of firms that have available operating cash flows data after $\tau = 1, 2, ..., 10$ years. The operating cash flows is the Compustat data item OANCF. $(N_\tau - N_1)/N_1$ reports the fraction of firms that drop out of the sample after $\tau(> 1)$ years relative to the year-1 number of firms. We remove firms with total assets less than \$12.5 million or book equity less than \$25 million in year t to avoid the influence of small firms. The sample period is from 1988 to 2021 as items from the statement of the cash flows become available only after 1988.

	Number of firms									
τ	1	2	3	4	5	6	7	8	9	10
N(firms)	3469	3265	3073	2897	2732	2578	2433	2298	2173	2058
$(N_1 - N_\tau)/N_1$		0.06	0.11	0.16	0.21	0.26	0.30	0.34	0.37	0.41