The Profitability and Investment Premium: Pre-1963 Evidence*

Sunil Wahal
WP Carey School of Business
Arizona State University
Tempe, AZ 85287
Sunil.Wahal@asu.edu

January 2017

* I am indebted to Stephen Bergauer, Iskandar Pashayev and Carter Wendt for coordinating data collection and for painstakingly checking the veracity of data. I thank Greg Waymire for advice on historical accounting conventions and Hank Bessembinder, Gene Fama, Ken French, Juhani Linnainmaa, Robert Novy-Marx and Seth Pruitt for comments. Wahal is a consultant to Dimensional Fund Advisors. DFA provided no funding or data for this research.

The Profitability and Investment Premium: 1940-1963

Abstract

I investigate the profitability and investment premium in stock returns using hand-collected data from Moody's Manuals for 1940-1963. Three results emerge. First, the profitability premium in 1940-1963 is similar in magnitude to the post-1963 period. Second, I detect no reliable relation between investment and returns, regardless of whether investment is measured using growth in total assets or book equity. The lack of an investment premium extends back to 1926. Third, unlike in 1963-2013, HML is not redundant in the Fama and French (2015) five-factor model.

1. Introduction

Comparative statics on the Miller and Modigliani (1961) dividend discount model motivate Fama and French (2015) to construct a five-factor model to describe the cross-section of returns. The five-factor model augments the market (R_M-R_F), size (small-minus-big, SMB) and value (high-minus-low, HML) factors in the three-factor model (Fama and French (1993)), with profitability (robust-minus-weak, RMW) and investment (conservative-minus-aggressive, CMA) factors. These additions are enabled by Novy-Marx's (2013) use of gross profitability as a proxy for expected economic profitability, and by Aharoni, Grundy and Zeng's (2013) approach to measuring investment using asset growth at the firm rather than per-share level. Pricing errors associated with the five-factor model are lower than those of the three-factor model, representing an improvement in the description of returns (Fama and French (2016a)). The four-factor model of Hou, Xue and Zhang (2015) also employs profitability and investment factors, motivated using q-theory and constructed slightly differently from the Fama and French (2015) factors. They too report lower pricing errors than the three-factor model. Profitability and investment clearly play a key role in these improvements.

The accumulated evidence on the ability of factor models to capture the cross-section of returns is typically based on CRSP-Compustat samples that start in July 1963. In this paper, I employ hand-collected data from Moody's Manuals to examine the role of profitability and investment in stock returns from July 1940 to June 1963. Three major results emerge from the out-of-sample tests.

First, the profitability premium is alive and well in the 1940-1963 period. The average slope on gross profitability (measured as revenues minus cost-of-goods sold scaled by total assets) in cross-sectional Fama-MacBeth regressions is 0.70 with a t-statistic of 3.66. By way of comparison, Novy-Marx (2013) reports a slope of 0.75 with a t-statistic of 5.49 for the 1963-2010 period. Small stocks do not play an inordinate role – even in large stocks, the slope on profitability is reliably positive. The intercept in time series regressions of a high-minus-low profitability portfolio on the three-factor model is 0.41 percent per month (t-statistic=3.69) in 1940-1963. Novy-Marx (2013) reports an intercept of 0.52 percent per month (t-statistic=4.49) for 1963-2010. The slopes on HML in these regressions are positive for low profitability portfolios and negative for high profitability portfolios, consistent with the negative correlation between profitability and value in the post-1963 data. This negative correlation pushes the intercepts in the three factor model above the average returns in these portfolios. Indeed, double sorts on value and profitability show that holding value roughly constant, the spread in profitability quintiles is as much as 1 percent per month. Operating profitability, which excludes SG&A and interest expense, and scales by book equity, delivers similar results.

Second, I am unable to detect a reliable relation between investment and stock returns. In Fama-MacBeth return regressions, the average slope on (prior) annual growth in assets is statistically indistinguishable from zero. In time-series tests, intercepts in three-factor models for portfolios sorted on investment are also not reliably different from zero. Since investment in the Miller and Modigliani (1961) valuation model is the expected growth in book equity, not total assets, I also replicate the above tests using growth in book equity. The results, or lack thereof, are similar.

One possibility is that in this sample period, realized changes in assets or book equity are a poor proxy for expected future investment. In the spirit of Fama and French (2006), I estimate Fama-MacBeth regressions that ask (cross-sectionally) whether prior growth in assets is predictive of growth in assets one, two, or three years ahead. The average slopes on prior growth in assets in the 1940-1963 period are similar to those for 1964-2015, implying that the proxy is no better and no worse in the pre-1963 period. Another possibility is that the lack of a premium associated with investment is due to lack of power in the shorter 1940-1963 period. I also estimate Fama-MacBeth return regressions and time series portfolio tests for the 1926-1963 period.[1] Even in this longer time series, there is little evidence of a robust relation between investment and returns. Either the pre-1963 period is unusual and there really is a relation between investment and returns, or the results for the post-1963 period are fragile. Another possibility is that there is a structural break in the relation, but it is hard to see an economic mechanism that might cause such a break.

[1] As I explain in detail in Section 2, the quality and consistency of income statement data prior to the establishment of the Securities and Exchange Commission in 1934, and the Committee on Accounting Practices in 1939, is extremely poor. The evolution of standardization in reporting practices and conventions indicates that profitability can be reliably calculated starting in fiscal years ending in 1938. Balance sheet data such as total assets and book equity, however, appear to be fairly good going back to 1925. Since the main tests require both profitability and investment, and because measuring investment require total assets from fiscal years t-2 and t-1, I start my primary test sample in July 1940. For a subset of tests that focus solely on investment, I use data going back to 1926.

Regardless, the lack of a premium associated with investment raises issues for the investment factor in the Fama and French (2015) and Hou, Xue and Zhang (2015) factor models, as well as for the literature that views the asset growth-return relation as an anomaly due to mispricing (Titman, Wei, and Xie (2004), Cooper, Gulen and Schill (2008)).

Third, I investigate the efficacy of RMW and CMA, and the redundancy of HML, in capturing average returns in 1940-1963. I do so in two ways. I first examine the ability of the three-, four- and five-factor models to explain the returns of 5x5 portfolios constructed at the intersection of various combinations of size, value, profitability, and investment. For most test assets, the Gibbons, Ross and Shanken (1989) test rejects the null hypothesis that some linear combination of the factor portfolios is on the minimum variance boundary. More importantly, the addition of RMW to the three-factor model lowers the GRS test-statistic and the drops the average absolute intercept by between 2 to 3 basis points per month for various test portfolios. This improvement is similar in magnitude to that reported by Fama and French (2015) for 1963-2013. The addition of CMA to the three-factor model, however, adds nothing in terms of spanning the test portfolios. On the other hand, the addition of the HML factor lowers the average absolute intercept, indicating that HML is useful in explaining the time series of returns on test portfolios.

I estimate spanning regressions of each factor on the others as a direct test of redundancy. For a regression of HML on the remaining factors in the 1963-2013 period, Fama and French (2015) report an intercept of -0.04 (t-statistic=0.47), drawing the conclusion that HML does not improve the mean-variance efficient tangency portfolio produced by the remaining factors. At least part of the story is that the large average HML return is absorbed by the extremely large slope on CMA (1.04, t-statistic=23.03), which has a large premium in this period. In stark contrast, in the 1940-1963 period, a regression of HML on the remaining factors has an intercept of 0.35 percent per month with a t-statistic of 3.70. The explanation is in the covariance of RMW with HML, and the lack of a CMA premium. RMW loads negatively on HML, which pushes the intercept in the HML regression above its average return. CMA still has a positive slope in this period (0.30, t-statistic=3.20) but it is dramatically reduced. Since CMA is unable to spread returns, it leaves room for HML to do so. [2]

The 276 months in the 1940-1963 period is about 30 percent of the length of the 1963-2013 period used by Fama and French (2015). A shorter time series should reduce power, making it harder to reject the null hypothesis that HML is redundant. The fact that the data reject the null, even in a shorter time series, suggests that power is not the problem. Of course, one can maximizing power by estimating the above spanning regressions for the longest possible time series, 1940-2015. For this period, the intercept in the HML regressions rises to 0.11 percent per month but with a t-statistic of only 1.58. I also split the post-1963 sample into three subperiods of about 210 months each. The intercept in the HML regression is only negative in one of the three subperiods suggesting that the redundancy of HML is sample-period specific.[3]

There is some overlap in my results and Linnainmaa and Roberts (2016). While their main purpose is to go after data snooping biases in 36 accounting-based return anomalies, they also examine profitability and investment back to 1926. They report that RMW and CMA are statistically indistinguishable from zero for the 1926-1963 period. The consistency of their investment premium results with mine is reassuring. The difference in the profitability results likely stems from two sources. As I explain in detail below, the first likely issue is the historical inconsistency of reporting and accounting treatment of expenses including COGS and SG&A. To capture expected economic profitability, we must be able to reliably observe true (realized) costs for a large, if not full, cross-section of firms. That appears to be the case only after the standardization of financial statements following establishment of the SEC in 1934, and specific prescriptions regarding the content and format of financial reports established by the Committee on Accounting Practices in 1939, and Regulation S-X in 1940. The 15-year period prior to 1940 likely adds considerable noise to estimates of a profitability premium. Indeed, in their subsamples

[2] The weakness of CMA in the pre-1963 period is also reflected in factor regression in which CMA is the dependent variable. For the 1963-2013 period, Fama and French (2015) report an intercept of 0.28 percent per month (t-statistic=5.03). In the 1940-1963 period, the intercept is a paltry 0.05 percent per month and statistically indistinguishable from zero.

[3] There are at least two other reasons to believe that HML is unlikely to be redundant. First, Fama and French (2016b) report that when using the cash profitability measure of Ball et al. (2016), the intercept the HML spanning regression rises to 0.30 with a t-statistic of 3.49. Second, in out of sample tests using international data, Fama and French (2015b) find that HML is not redundant.

that start in 1938, their profitability factor has a three factor intercept of 0.30 percent per month. A second possibility is in the manner with which Linnainmaa and Roberts (2016) examine the profitability premium. The 2x3 sorts on size and profitability used to construct RMW ignore the middle 40 percent of firms. But Fama-MacBeth regressions, and portfolio sorts on profitability (either univariate or in conjunction with size, value and investment) use the full cross-section; in both types of tests, profitability reliably predicts returns.

The remainder of the paper is organized as follows. In section 2, I describe the data collection process that accounts for the vagaries introduced by historical accounting practices. Section 3 contains tests of profitability and investment. Section 4 presents standard asset pricing tests. Section 5 concludes.

2. Historical accounting practices, data collection and sample construction

2.1 Historical accounting conventions

An understanding of historical accounting conventions is critical to the data collection effort and its usefulness in asset pricing tests. Carey (1969) provides a comprehensive overview of the historical accounting practices between 1896 and 1936 (for an analysis of the impact of standardization in accounting, see Madsen (2011)). He describes the lack of rules and uniformity in accounting conventions in detail, and points to the discretion available to accountants in the early 1900s in deciding what to record and how to record it: "Our brethren of law have the Supreme Court to whose dictates they must conform…in our profession it is left to each individual to be a law unto himself and the result is a mass of conflicting options on many subjects, each one of which receives its values principally from the reputation of the person holding it, or the more or less convincing way in which he can express it" (pg. 76). This dismal assessment of discretion is shared by others, most notably Berle and Means (1932, pg. 310), who argue that "the integrity of the accountant and the soundness of his method are the greatest single safeguard to the public investor…But the rules of accounting are not as yet fully recognized rules of law…In fact, the failure of the law to recognize accounting standards is probably due to the lack of agreement among accountants".

The lack of uniformity is also obvious in the data. Carey (1969) describes situations in which assets and liabilities are sometimes stated as if proposed financing had actually become effective, so that "pro forma" financial statements are stated as "actual" financial statements. He discusses the workings of a joint committee of the Institute of Accountants and the SEC charged with accounting standardization. This committee entertained proposals that reveal the degree of disagreement among accountants about measurement, even with respect to seemingly innocuous items like revenues. For example, the above committee addressed the request from not insubstantial quarters that that disclosure of total sales not be required because doing so might attract competition detrimental to shareholders of the disclosing firm. In fact, many early income statements do not report revenues and start with an arbitrary measure of gross profit.[4]

The above is just one of many examples that make extracting useful economic information from early financial statements challenging. Most importantly for calculating profitability, there appears to be considerable discretion in recording and assigning expenses. Greer (1928) urges that "the distinction between cost of goods sold and operating expense be ignored, and that all outlays in connection with purchase and sale be considered one grand total cost." The implication is that reporting firms can conflate COGS and SG&A. In other cases, depreciation is either allocated to COGS, or SG&A, or both, in addition to being reported separately. Berle and Means (1932, pg. 310-311) describe eight common methods which were used to manipulated reported profits, but Patton (1932) provides a more authoritative account of the accounting variations and conventions of the time, including the rationale behind them. For example, he describes a variety of expense items that can be classified in COGS or SG&A, depending on the accounting convention used by the firm. And in describing the sequence and grouping of income statement items, Patton (1932, pg. 25) points out that "variations of this arrangement are many", citing a variety of authors who argue that the same expense should enter into different accounts within the income statement.[5] If unprofitable firms are less likely to report COGS and other expense items, or more likely to

[4] For instance, income statements for the Hamilton Watch Company contained in the 1931 edition of the Moody's Manual start with gross income and do not include cost of goods sold. The 1936 edition, however, contains income statements that start with Net Sales and include cost of goods sold for fiscal year 1935. For fiscal years 1933 and 1934 (which are in the same edition), the entry for Net Sales contains "Not stated", and cost of goods sold is left blank.

[5] Interestingly, Patton (1932) also points out that suggestions by the Federal Reserve Board regarding the calculation of net income appear to differ from recommendations by the U.S. Supreme Court.

manipulate expenses, this could bias tests of profitability. On the other hand, if the distribution of firms not reporting a set of expenses is random, this merely adds noise to the tests. Unfortunately, the noise is not inconsequential since it reduces power and hence the ability to detect a profitability premium.

A chronology of standardization is useful in understanding changes in accounting conventions. In 1932, the New York Stock Exchange announced that it would require audited financial statements from listed companies starting the following year; prior to that, firms could choose to report unaudited statements. In 1935, the SEC created the office of the Chief Accountant as a way to coordinate the standardization process. The office of the Chief Accountant started to provide opinions on financial statements via Accounting Release Statements (ASR), the first of which was released in 1937. ASR No. 4, released in 1938, specifically declared that disclosure was inadequate if there was no "authoritative support" or if it deviated from SEC suggestions (contained in other ASRs). Standardization started to take hold the following year with the formation of the Committee on Accounting Practices in 1939, the enactment of Regulation S-X, and the publication of Patton and Litton in 1940. Regulation S-X prescribed the content and format of financial reports, with Rule 503 specifically requiring disclosure of income statement items. Patton and Litton (1940) argued that accounting statements must allocate costs and revenues to periods, as opposed to merely valuing assets and liabilities.

Even later in the times series, some accounting conventions are industry specific so that improper treatment of them can lead to systematic data problems. For example, income statements for railroads are organized differently from other industrial firms to conform to the rules of the Interstate Railroad Commission (see, for example, Hooper (1916)). The commission required a six-account system which broke out expenses in ways that differ substantially from other firms. Similarly for railroads and utilities, maintenance and repairs, which represent the majority of costs, are broken out separately so that not including them would substantially overstate profits. Another example comes from mining and extraction firms. In such firms, discovery expense can be quite large and is sometimes lumped in "other expenses", even though one could view it as analogous to R&D because it is incurred in a current period to perhaps generate future revenue.

My reading of the historical literature suggests data from the income statement is adequate starting about 1938. High-level balance sheet information such as total assets, however, seems to be of high quality going back to 1926. This normative assessment is confirmed by the data itself, described below.

2.2 Sample construction and data collection process

The starting point for the data collection process is the historical book equity data file provided on Ken French's website.[6] That file includes the data used in Davis, Fama and French (2000) as well as additional data for non-industrial firms. The file includes a CRSP Permno, the first and last year of the edition of the Moody's Manual used to collect book equity data, and the book value of common equity in year *t*. I match each Permno from the Davis, Fama and French (2000) file to all possible company names reported in CRSP. This generates an exhaustive list of company names over time, which are then matched with firm names in the Mergent database (www.mergentarchives.com). Research assistants then download a PDF for each firm-year from the relevant Moody's Industrial, Bank & Finance, or Utilities Manual.

I employ a team of data assistants to hand-collect the following data items: revenue (defined as sales net of discounts, returns and allowances), cost of goods sold (COGS), selling, general and administrative expenses (SG&A), depreciation (including amortization and depletion), interest expense, and total assets. The data assistants first go through a selection mechanism in which a large group of assistants collects the same data for a subsample of firms with a variety of accounting complexities. The collected data are compared to (known) true values, so that the larger group can be culled. This leaves a smaller group of data assistants who have demonstrated an understanding of the data requirements.

The data are collected from the first Moody's date through the last year the company appeared in the Moody's Manuals according to Davis, Fama and French (2000), or else the first year the company was included in Compustat data. The data are recorded in a standardized spreadsheet. The quality and readability of the PDFs varies from those that are barely legible to

[6] This is a different starting point from Linnainmaa and Roberts (2016) who start with all CRSP firms going back to 1918.

extremely clear (generally, PDFs from early in the time series are hard to read). In cases where the data are not clear, data assistants leave the field blank.

2.3 Accounting issues and data cross-checks

Three research assistants with expertise in finance and accounting oversee the data collection effort to ensure that there is a mapping between the historical data and modern accounting standards.[7] Before doing so, they learn the accounting practices of the time by reading Patton (1932), and where they have questions, consult me. For instance, they ensure that maintenance and repairs for railroads and utilities are included in COGS to be consistent with industrial firms. Similarly, they reclassify discovery expense for mining and extraction firms as R&D so that it does not enter COGS or SG&A. In addition to the above tasks, these research assistant perform three checks to guard against data collection errors. First, they flag the 1st and 99th percentile for all variables in each annual cross-section and then, if necessary, correct the tails of the distribution for accounting or data recording errors. Second, for each firm, they flag extreme values in the time series distribution of each variable. If the extreme values are unreasonably small or large, they check the data against the original Moody's document. Third, they manually check each data point in the database, making sure that the data item collected conforms to modern accounting principles.

The hand-collected data are spliced with data from Compustat starting in 1950. There is some deliberate overlap between the two data sources that allows me to cross-check the accuracy of the data collection process. For instance, COGS collected from Moody's Manuals for a firm in 1952 should be exactly equal to that reported by Compustat. Reassuringly, disagreement between the two sources is minimal, largely due to rounding in Compustat.

The data collection process goes back to Moody's Manuals in 1925. For most firms, basic balance sheet data are readily available. Between 1926 and 1940, the average annual percentage of firms with valid book equity that are missing total assets is only 6.4 percent. The worst coverage is in 1928 when 17 percent of firms are missing total assets. By 1938, only 2 percent of firms do not have total assets. Consistent with the narrative on historical accounting conventions in Section

[7] These research assistants are distinct from the data collection team and therefore an independent data check.

2.1, income statement data are much more problematic. In the 1926-1940 period, over 27 percent of firms with valid book equity are missing COGS. Again, the worst coverage is in 1928 when over 40 percent of firms do not have valid COGS information. Data coverage improves starting in 1935. By 1939, COGS are available for over 95 percent of firms with valid book equity.[8]

Given the missing data issues and the above historical context, I make the judgment that data quality is adequate starting about 1938. Because calculating investment requires data from two fiscal years prior to returns (i.e. fiscal years in t-2 and t-1), I start my time series in July 1940. Throughout, I maintain the convention of at least a six-month time lag between the fiscal year end and the return data.

3. Cross-Sectional tests

3.1 Fama-MacBeth regressions

Table 1 shows average slopes (multiplied by 100) and t-statistics from monthly Fama-MacBeth regressions of stock returns. I use the two measures of profitability employed by Novy-Marx (2013) and Fama and French (2015): (a) gross profitability, defined as revenues minus COGS, scaled by total assets (GP/AT), and (b) operating profitability, defined as revenues minus COGS, minus SG&A, minus interest expense, scaled by book equity (OP/BE).[9] I also use two measures of investment: (a) the growth in total assets from fiscal year t-2 to t-1, referred to as dA_{t-1}/A_{t-2}, and (b) the growth in book equity from fiscal year t-2 to fiscal year t-1, termed dB_{t-1}/B_{t-2}. As is standard, the regressions control for size (log(ME)), book-to-market ratios (log(B/M)), and prior returns ($R_{1,0}$ and $R_{2,12}$). Independent variables are trimmed at the 1 and 99 percentile to mitigate the influence of outliers. The slopes on size, book-to-market, and prior returns are similar in magnitude and significance to those reported in many other studies so I do not dwell on them further.

The valuation equation that is at the center of the Fama and French (2015) five-factor model suggests a premium for expected profitability, holding book-to-market ratios and investment

[8] A data item can be missing either because it is was not reported in the financial statement, or because the quality of the PDF was so low that it could not be read.

[9] In separate unreported tests, I also perform the scaling decomposition in Ball et al. (2015). The results are largely similar.

constant (and similarly for different combinations of the three variables). With two measures of profitability and two measures of investment, controlling for different measures results in four regression specifications. These regressions are displayed in Panel A of Table 1 for the 1940-1963 period.

Across all stocks, the average slope on gross profitability is 0.70 (t-statistic=3.66) when controlling for dA_{t-1}/A_{t-2}, and 0.65 (t-statistic=3.55) when controlling for dB_{t-1}/B_{t-2}. By way of comparison, Novy-Marx (2013) reports a slope of 0.75 (t-statistic=5.49) for the July 1963 to December 2010 period. The slopes on operating profitability are similar: 0.91 (t-statistic=3.24) and 0.77 (t-statistic=3.01) with corresponding controls for investment. Separate regressions for large and small stocks (based on NYSE median market value cutoffs), yields similar slopes on the profitability measures.

The average slopes on investment are small with t-statistics well below 2.00. This is true, regardless of whether I measure investment using growth in assets or book equity. In large capitalization stocks, the average slopes are well within two standard errors. In small stocks, the slope on growth in book equity is positive (rather than negative as in the 1963-2015 period), with t-statistics of 1.99 and 2.26. Using growth in assets, the slopes are statistically indistinguishable from zero. Broadly, there seems to be little evidence that investment is negatively related to future returns in the 1940-1963 period. This is in stark contrast to the evidence in Cooper, Gulen and Schill (2008), Aharoni, Grundy and Zeng (2013), Titman, Wei and Xie (2004) and others.

It could be that deviations from clean surplus accounting make it difficult to detect a relation between investment and returns, particularly given accounting practices in the pre-1963 period. There are several reason why this is unlikely to be the case. First, problems with clean surplus should influence book-to-market and profitability just as significantly as investment; that is clearly not the case since book-to-market ratios and profitability are reliably related to returns. Second, clean surplus is only required if one interprets investment in the context of the Miller-Modigliani valuation formula. If one subscribes to the mispricing explanation based on an over-reaction to past firm growth rates (Titman, Wei and Xie (2004), Cooper, Gulen and Schill (2008)), then clean surplus is irrelevant. Third, Linnainmaa and Roberts (2016) conduct a test of clean surplus accounting and conclude that violations are no different in the pre- versus post-1963 period.

It is possible that the lack of a relation between investment and returns in the 1940-1963 period is because of low power. Extending the sample period beyond 1963 is problematic because the test loses its out of sample character. I can, however, extend the sample back to 1926. The cost of doing so is that I can no longer control for profitability. This matters if one is interested in the effect of investment within the confines of the Fama and French (2015) or Hou, Xue and Zhang (2015) factor models, both of which require holding profitability constant. [10] But if investment is of independent interest, or driven by mispricing, then the 1926-1963 period is just as informative. Panel B shows estimates of regressions for the 1926-1963 period using growth in assets and book equity. When the regression is estimated across all stocks, the average slope on dA_{t-1}/A_{t-2}, is -0.93 with a t-statistic of 1.97. In contrast, the slope on dB_{t-1}/B_{t-2} is statistically indistinguishable from zero. In large stocks, both variables have large standard errors. In small stocks, the slope on dA_{t-1}/A_{t-2}, is -2.23 with a t-statistic of 2.37. Since Fama-MacBeth regressions ascribe equal weights to all stocks, it is likely that the negative slope on dA_{t-1}/A_{t-2} for the full cross-section is driven by small stocks.

3.2 Univariate portfolio sorts

In this section, I examine the performance of value-weighted portfolios based on the two measures of profitability and investment. This addresses concerns that the Fama-MacBeth regressions may be sensitive to extremes or overly influenced by small stocks. I form five portfolios each June based on NYSE breakpoints and rebalance annually. I impose the same data restriction on the sample as the Fama-MacBeth regressions to ensure that the two sets of results are comparable.

---

[10] The scaled version of the Miller-Modigliani valuation model is $\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E(Y_{t+\tau} - dB_{t+\tau})/(1+r)^{\tau}}{B_t}$ in which firms with higher future growth in book equity have lower expected returns, controlling for book-to-market and expected profitability. And in the Hou, Xue and Zhang (2015) argument, the relation between investment and stock returns is also conditional on expected profitability because firms invest more when their marginal q is high; holding expected profitability constant, low discount rates imply high marginal q and high investment.

Table 2 shows excess returns, as well as intercepts and slopes from the three-factor model. Although not shown in the table, the portfolios are reasonably well diversified, on average containing 130 securities in each quintile. Panel A contains estimates for gross and operating profitability portfolios for the 1940-1963 period. The spreads in excess returns between low and high profitability portfolios are positive, between 0.10 and 0.15 percent per month, but with large standard errors. The slopes on HML are monotonic across quintiles. As in the post-1963 data, low profitability firms have sharply positive loadings on HML and the opposite is true for high profitability firms. These loadings on HML push the intercepts in the three-factor model above the average excess returns. The low gross profitability portfolio loses about 0.17 percent per month (t-statistic=2.14) while the high gross profitability portfolio has an intercept of 0.24 percent per month (t-statistic=3.31). The high-minus-low spread portfolio earns about 0.41 percent per month with a t-statistic of 3.41. The spread is comparable to that reported by Novy-Marx (2013) for the 1963-2010 period (0.52 percent per month with a t-statistic of 4.42). The spread using operating profitability is similar, 0.37 percent per month with a t-statistic of 3.58.

Panel B contains equivalent estimates from portfolios formed on the two measures of investment. There is no consistent pattern in excess returns across investment quintiles using either growth in assets or book equity. Exploiting the extreme portfolios, the high-minus-low spread portfolio has an excess return of -0.07 percent per month using asset growth and 0.14 percent per month using growth in book equity. Both are well within two standard errors. Using the three factor model, the high-minus-low spread portfolio constructed from asset growth has a paltry intercept of -0.02 percent per month with a t-statistic of only 0.16. Similarly, for portfolios formed on dB_{t-1}/B_{t-2}, the intercept is 0.14 percent per month with a t-statistic of 1.35.

As with the Fama-MacBeth regressions, I extend the sample back to 1926. The results from those portfolios are in Panel C. Here the influence of small stocks in the Fama-MacBeth regressions in Panel B of Table 1 is evident. In those regressions, the negative slope on asset growth for small stocks generated a marginally significant negative slope (t-statistic=1.97) for all stocks. In the portfolio tests in Panel C for Table 2, however, the intercepts are well within two standard errors, suggesting that the negative slope was due to equal-weighting small stocks in the Fama-MacBeth regressions.

3.3 Expected profitability and expected investment

The valuation model motivating the five-factor model says that expected returns are related to expected profitability and expected investment. It is possible that in this sample period, past profitability and investment are poor predictors of future profitability and investment. In other words, it could be that the noise in forming expectations is greater in 1940-1963 than in 1963-2015. This is a plausible scenario if accounting statements are especially noisy or not readily available. To determine if this is the case, I use the approach of Fama and French (2006) and estimate annual cross-sectional regressions of growth in assets and operating profitability one, two and three years ahead ($dA_{t+\tau}/A_t$ and $OP_{t+\tau}/BE_t$) on lagged growth in assets and lagged operating profitability (dA_t/A_t and OP_t/BE_t). Panel A of Table 3 shows univariate regressions. To allow for side-by-side comparisons, the first set of columns show regressions for 1964-2015 and the second set shows the 1940-1963 period. For the asset growth regressions, the slopes on dA_t/A_t are somewhat lower in the 1940-1963 but the shrinkage is quite small. For the operating profitability regressions, the slopes on OP_t/BE_t are slightly larger.

Panel B shows similar multivariate regressions that control for size and book-to-market ratios. In these regressions, the average slopes on dA_t/A_t and those on OP_t/BE_t in the 1940-1963 period are very similar to those in the 1964-2015 period. For instance, the average slopes for the dA_t/A_t in the one, two and three period ahead regressions for the 1964-2015 period are 0.16, 0.23 and 0.33. The equivalent slopes for the 1940-1963 period are 0.15, 0.25 and 0.27.[11] This suggests that there is very little difference in the information content of current period investment for future investment between the two sample periods. Nonetheless, I also estimate the Fama-MacBeth regressions in Table 1 replacing the investment and profitability variables with the fitted values from the cross-sectional regressions in Panel B of Table 2. The results are not reported to avoid repetition because they are quite similar to those in Table 1.

[11] The comparable slopes in Fama and French (2006) for the 1963-2004 period are 0.05, 0.10, and 0.11. There are, however, several differences. Fama and French's (2006) implementation of these regressions are at a per share level, and include other accounting variables related to earnings, dividends, and the probability of default. Given the lack of good accounting data in the pre-1963 period, such accounting variables are not calculable. I therefore use a simpler version of these regressions that only controls for size and book-to-market.

Finally, I examine the cross-sectional mean and standard deviation of dA_{t-1}/A_{t-2}, as well as the adjusted R^2 from the regressions in Panel B for the full time series. The purpose is to determine if there is some sort of change in the time series of either investment or expected investment that is centered in the years around 1963. I see no particular pattern and therefore do not report these in a table.

4. Asset pricing tests

Since the asset pricing tests are fairly standard, I follow the procedures in Fama and French (2015), generally replicating (but not re-reporting) their results for the 1963-2013 period before proceeding to the analysis of the 1940-1963 period. I dispense with alternative measures of profitability and investment, and only present results for OP/BE and dA_{t-1}/A_{t-2}. This allows for straightforward comparisons with the 1963-2013 period. The results using gross profitability and dB_{t-1}/B_{t-2} are not meaningfully different and not reported in tables.

4.1 Test portfolios and factor construction

I start by constructing 5x5 portfolios based on independent sorts on various combinations of size, value, operating profitability, and investment. The sample is identical to that used in Tables 1 and 2. As before, portfolios are formed every June and rebalanced annually. Panel A of Table 4 shows average monthly value-weighted excess returns for portfolios that use size and one other variable. The value premium is present in both small and big stocks. Aside from the very smallest group of stocks, there is a profitability premium in all other size quintiles. In big stocks, the low profitability portfolio has an average excess return of 0.86 percent per month, rising to 1.10 percent per month for the high profitability portfolio. Despite this, the increase in excess returns across profitability portfolios is not monotonic, especially in the middle three size quintiles (i.e. outside small and big stocks). This is in contrast to the monotonic increase in excess returns across size and profitability portfolios reported in Table 4 of Novy-Marx (2013) and Table 1 of Fama and French (2015). The differences in average excess returns across size and investment portfolios are miniscule and appear to be randomly distributed. For instance, in small stocks, the low investment quintile has an excess return of 1.33 percent while the high investment quintile has an excess return of 1.21 percent, a spread of only 0.12 percent per month. In big stocks, the spread is -0.13 percent. Given the lack of a premium associated with investment in the Fama-MacBeth regressions (Table 1) and univariate portfolios sorts (Table 2), this is not particularly surprising.

Panel B contains excess returns for 5x5 portfolios formed from pairs of value, operating profitability, and investment. Controlling for book to market ratios, profitability generates much larger variation in returns. In the low B/M group, the low OP/BE portfolio has an excess return of 0.12 percent per month while the high OP/BE portfolio earns 1.16 percent per month, a spread of 1.04 percent per month. In extreme value firms (high B/M), the spread is 1.02 percent per month. Holding investment roughly constant, the profitability premium is also large: in low dA_{t-1}/A_{t-2} firms, the spread between high and low OP/BE firms is 0.31 percent per month, and in high dA_{t-1}/A_{t-2} firms, the spread rises to 0.51 percent per month. As before, there are no spreads in portfolios formed on dA_{t-1}/A_{t-2}, holding profitability or value constant.

Size effects could play a role here so I also examine returns to portfolios formed on pairs of B/M, OP/BE and dA_{t-1}/A_{t-2}, within two size groups (small and big stocks). Since the number of securities in the cross-section is smaller in this sample period, I use 3x3 rather than 5x5 portfolio sorts within each size group to ensure adequate diversification. The returns of these 2x3x3 portfolios confirm the above patterns in both large and small stocks – the combination of value and profitability generates large spreads in both small and large cap stocks, with particularly large spreads in small stocks. I do not report these portfolio returns to avoid redundancy but the results are available on request.

For the right hand side of the asset pricing regressions, I form profitability and investment factors following the procedures in Fama and French (2015). I elect to use only the simpler 2x3 sorts, instead of the 2x2x2x2 sorts for two reasons. First, Fama and French (2015) advocate the 2x3 sorts because they isolate exposures to value, profitability and investment just as well as the more complicated 2x2x2x2 sorts that jointly control for all other variables. Second, the number of stocks in the 1940-1963 period is smaller than in the 1963-2013 period. As a result, 2x3 sorts produced better diversified portfolios; in this sample period, 2x2x2x2 sorts sometimes produce portfolios with empty cells.

Panel A of Table 5 shows average monthly excess returns and standard deviations for the 2x3 portfolios. The naming convention for the portfolios is the same as Fama and French (2015). The first letter “S” or “B” refers to big or small stocks. For B/M, the second letter (L, N, H) refers to low, neutral and high B/M ratios. For OP/BE, the second letter (W, N, R) refers to weak, neutral or robust profitability. And finally for investment, the second letter (C, N, A) refers to conservative, neutral or aggressive investment.

In the size and B/M portfolios, the value premium is easily visible in both small and big stocks. The BL portfolio has an average monthly excess return of 1.03 percent while the BH portfolio has a return of 1.60 percent. In the size and profitability portfolios, the profitability premium again shows up but with a nuance. In small stocks, the SW portfolio has an average return of 1.35 percent. The SR portfolio has a higher return (1.48 percent) but in fact the SN portfolio has an even greater return (1.55 percent). This is the non-monotonicity observed in the finer 5x5 portfolios (Panel A) that also shows up in these coarser sorts. In big stocks, however, the monotonic pattern in returns across profitability portfolios reappears: the BW, BN, and BR portfolios have average returns of 1.07, 1.09 and 1.22 percent respectively. Finally, as expected, the variation in the 2x3 portfolio returns for investment are tiny. The spread between conservative versus aggressive portfolio returns is just 0.03 percent in small stocks and 0.04 percent in big stocks.[12]

Panel B shows average monthly returns and standard deviations of the factor building blocks created from the 2x3 portfolios. Specifically, the size spreads across B/M, OP/BE and dA_{t-1}/A_{t-2} (labelled SMB_{bm}, SMB_{op}, SMB_{inv}), as well as equivalent HML, RMW and CMA spreads in small and big stocks. The HML spread is large in both small and big stocks (about 0.55 percent in each). The RMW spread is also similar in small and big stocks (0.13 and 0.15 percent respectively). The CMA spread is, as expected, puny in both small stocks (0.03 percent) and big stocks (-0.04 percent).

[12] The standard deviations of each of the 2x3 portfolio returns follow familiar patterns, generally larger for high B/M portfolios, lower for robust profitability portfolios, and larger for aggressive investment portfolios. Readers interested in comparing portfolio returns in Panel B to the 1963-2013 sample should compare to Table A1 in Fama and French (2015).

Panel C contains the final 2x3 factors built from these portfolios. SMB has an average monthly return of 0.29 percent, coincidently the same as the 0.29 percent reported by Fama and French (2015) for the 1963-2013 period. HML has an average return of 0.43 percent per month, higher than the 0.37 percent for the 1963-2013 period. The big differences come in RMW and CMA. RMW has an average monthly return of 0.14 percent over the 1940-1963 period. Fama and French (2015) find that RMW has a much higher average return of 0.25 percent per month in 1963-2013. (In their 2x2 factors that use the full cross-section of stocks, the average monthly return for RMW drops to 0.17 percent.) Linnainmaa and Roberts (2016) report an average monthly return of -0.36 percent for RMW in the 1926-1944 period. But in their 1945-1963 period, which closely overlaps with my 1940-1963 sample, the average return for RMW in their data is almost identical (0.15 percent per month) to mine (0.14 percent per month). Consistent with the lack of an investment premium, CMA has an average monthly return of -0.01 percent in the 1940-1963 period. This is also very similar to the average return of -0.08 percent per month reported by Linnainmaa and Roberts (2016) for 1945-1963.

4.2 Asset pricing tests

Table 6 contains tests of the performance of three-, four- and five-factor models for the 1940-1963 period. The test assets are the 5x5 portfolios in Table 4. The table is organized in a manner similar to that of Table 5 in Fama and French (2015) so that readers can easily follow the parallel set of results. The first column shows factors that augment MKT and SMB in time series regressions. For example, the first row, labelled "HML", is simply the three-factor model. The second row, labelled "HML, RMW" is a four-factor model that augments the three-factor model with RMW. For each factor model, I show the value of the GRS test, the associated p-value, and the average absolute value of the intercepts. The final column shows the average absolute value of the intercept scaled by the absolute value of the average return on each portfolio minus the average of the portfolio returns ($A|\alpha_i|/A|\overline{r_i}||$). The denominator is a scalar that describes the variation in test portfolio returns, so the quotient represents the percentage of the dispersion in test portfolio returns left unexplained by each factor model.

The GRS test frequently rejects the null hypothesis that some linear combination of the various factor portfolios is on the minimum variance boundary. From an economic and practical standpoint, it is more interesting to ask whether the profitability and investment factors are useful in (a) shrinking intercepts, and/or (b) shrinking the dispersion in intercepts relative to average test portfolio returns.

The answers are fairly clear. Take, for instance, the size and B/M portfolios in panel A. The average absolute intercept in the three-factor model is 0.157. Adding RMW shrinks that by about 2 basis points (to 0.135). And the percentage of the dispersion in test portfolio returns unexplained by the model drops by 10 percentage points with the addition of RMW (from 0.79 in the three-factor model to 0.69 in a four-factor model that includes RMW). In contrast, adding CMA (but not RMW), leaves the average absolute intercept virtually unchanged at 0.154. And the value of $A|\alpha_i|/A|\overline{r_i}||$ is virtually unchanged from 0.79 in the three-factor model to 0.78 in a four-factor model that includes CMA. A similar story emerges for the size and OP/BE portfolios (Panel B). The three-factor model has an average absolute intercept of 0.13. Adding RMW reduces the intercept by 2.5 basis points to 0.119, and $A|\alpha_i|/A|\overline{r_i}||$ drops from 0.70 to 0.64. In contrast, adding CMA only decreases the intercept by 0.4 basis points to 0.126, and $A|\alpha_i|/A|\overline{r_i}||$ drops very marginally to 0.68. Panel C simply reinforces evidence presented earlier – sorts on size and dA_{t-1}/A_{t-2} do not contain a pattern in returns so the inclusion or exclusion of various factors does not influence average absolute intercepts of test portfolios that are basically noise. Panels D through E contains results from test portfolios that ignore size but control for pairs of value, profitability and investment. Again, except when investment is part of the test portfolio, adding RMW reduces the average absolute intercepts.

4.3 HML redundancy

A closer look at the results in Table 6 shows the importance of HML in this sample period. Consider the size and B/M portfolios in Panel A. A four-factor model that includes MKT, SMB, RMW and CMA has an average absolute intercept of 0.176. Adding HML, the five-factor model reduces that by almost 4 basis points to 0.138. The last column shows that the percent of variation in the left-hand-side portfolio returns left unexplained by the above successive models drops enormously from 0.89 to 0.70. The implication is that not only is HML not redundant in pricing these portfolios, but it is quite powerful in explaining their excess returns.

Fama and French (2015) estimate spanning regressions of each factor on the others to jointly characterize the dependence in factors and understand their marginal contributions. In their 1963-2013 period, a regression of HML on the remaining factors has an intercept of -0.04 percent per month with a t-statistic of 0.47. In contrast, MKT, SMB, RMW and CMA have intercepts of 0.82, 0.39, 0.43 and 0.28 percent per month respectively, all with t-statistics well above 2.00. The implication is not that there is no value premium (far from it), but that HML does not improve the mean-variance efficient tangency portfolio that could be produced via a combination of MKT, SMB, RMW and CMA. It is in this sense that Fama and French (2015) refer to the redundancy of HML and suggest that "It will be interesting to examine whether this result shows up in U.S. data for the pre-1963 period…" (p. 12).

Panel A of Table 7 shows such factor regressions for 1940-1963. The intercepts for the MKT and RMW regressions are large, with t-statistics above 2.00. The intercept for SMB is surprisingly small (0.14 percent per month) with a t-statistic of 1.34. Perhaps not surprisingly, the intercept for CMA is only 0.05 percent with a t-statistic of 0.92. Most importantly, HML has an intercept of 0.35 percent per month with a t-statistic of 3.70. In this sample period, HML does not appear to be redundant.

At least part of the explanation for the difference in intercepts in the HML regressions between the two periods resides in the slopes on CMA. In the 1963-2013 period, Fama and French (2015) report that the HML regression has a slope on CMA of 1.04 (t-statistic=23.03), which means that CMA absorbs much of the variation in HML. In the 1940-1963 period, CMA itself is weak at best, and the slope on CMA is only 0.30 (t-statistic=3.20), failing to absorb as much of the return variation in HML. In other words, the weakness in CMA leaves the door open for HML to do the work of explaining return variation. And consistent with the three factor model results in Table 3, RMW loads negatively on HML (RMW has a slope of -0.85 with a t-statistic of 11.91), which pushes the intercept in the HML regression above its average return.

The 1940-1963 sample contains 276 months, substantially shorter than the 1963-2013 period. But lower power should make it harder to reject the null hypothesis that HML is redundant.

To increase power as much as possible, I also estimate regressions for the full 1940-2015 period. Panel B shows the results of those regressions. The intercept for the HML regression is now positive (0.11 percent per month) but with a t-statistic of only 1.58. The difference in intercepts in the HML regressions in the two time periods suggest that looking at subperiods within the 1963-2015 period might also be useful. I split the post-1963 sample into three equal subperiods that correspond to roughly 210 months each. The results of these regressions are reported in Panels C through E. The intercept in the HML regression is only reliably negative the last subperiod, July 1998 to December 2015, suggesting that the redundancy of HML is far from pervasive.

There are at least two other reasons to be believe that the redundancy in HML is specific to the sample in Fama and French (2015). First, Fama and French (2016b) find that after constructing RMW from a cleaner cash profitability measure, the intercept in the HML spanning regression rises to 0.30 (t-statistic=3.49). Second, in an out-of-sample test using international data, the intercepts in HML spanning regressions in regions outside the US are reliably positive (Fama and French (2015b)).

5. Conclusion

I examine the profitability and investment premium in the 1940-1963 period. Reassuringly, the profitability premium is similar in magnitude and significant to the post-1963 samples in most studies. Not so reassuringly, the investment premium is absent regardless of whether I employ growth in assets or book equity to measure investment. Even extending the sample back to 1926 does not seem to resurrect the results in post-1963 samples. Why might this be the case? It could be that the problem is low power. That seems unlikely, however, since the 1926-1963 series is reasonably long and the cross-section quite complete. Moreover, the fact that the value and profitability premiums are present suggests that power is not the problem.

Unlike the post-1963 period, HML is not redundant in the five-factor model prior to 1963. Could one make the argument that it is CMA, not HML, that is redundant? That depends on how one views the pre- versus post-1963 samples. Prior to 1963, CMA does not appear to be useful in asset pricing tests, but after 1963 it appears to improve the mean-variance efficient tangency portfolio. The issue is especially relevant if one is interested in comparing the relative efficacy of the Fama and French (2015) and the Hou, Xue, and Zhang (2015) factor models; both employ a version of CMA but the latter does not employ HML. A head-to-head comparison of the two models in the pre-1963 period would be interesting but that requires accounting data for the Hou, Xue and Zhang (2015) profitability factor (income before extraordinary items and quarterly book equity) that is not available. It would also be interesting to assess the difference between operating and cash profitability in this sample period (Ball et al. (2016), Fama and French (2016b)). But again, the accounting data required to properly measure accruals is unlikely to be of a high enough quality for that to be a successful endeavor.

References

Aharoni, G., Grundy, B., Zeng, Q., 2013, Stock returns and the Miller-Modigliani valuation formula: revisiting the Fama-French analysis, *Journal of Financial Economics* 110, 347-357.

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

Ball, R., J. Gerakos, J.T. Linnainmaa, V. Nikolaev, 2016, Accruals, cash flows, and operating profitability in the cross section of stock returns, working paper, University of Chicago.

Berle, A. A., and G.C. Means, 1932, The modern corporation and private property, New York, MacMilan.

Carey, John L, 1969, The rise of the accounting profession: From technician to professional, 1896-1936. New York, American Institute of Certified Public Accountants.

Cooper, M.J., Gulen, H., and Schill, M., 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

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

Fama, E., French, K., 2006, Profitability, investment and average returns, *Journal of Financial Economics* 82, 491-518.

Fama, E., French, K., 2015, A five-factor asset pricing model, *Journal of Financial Economics* 16, 1-22.

Fama, E., French, K., 2015b, International tests of a five-factor asset pricing model, working paper, University of Chicago.

Fama, E., French, K., 2016a, Dissecting anomalies with a five-factor model, *Review of Financial Studies* 29, 69-103.

Fama, E., French, K., 2016b, Choosing factors, Working paper, Dartmouth College.

Greer, H.C., 1928, How to understand accounting, The Ronald Press Company, New York.

Hooper, W.E., 1916, The accounting system prescribe for railroads by the Interstate Commerce Commission, The Annals of the American Academy of Political and Social Science, Vol. 63, 222-231.

Madsen, P.E., 2011, How standardized is accounting? *The Accounting Review* 86, 1679-1708.

Miller, M.H., Modigliani, F., 1961, Dividend policy, growth and the valuation of shares, *Journal of Business* 34, 411-433.

Linnainmaa, J. T., Roberts, M., 2016, The history of the cross section of stock returns. Working paper, University of Chicago.

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

Patton, W.A., 1932, Accountants' handbook, The Ronald Press Company, New York.

Patton, W.A., and A.C. Littleton, 1940, An introduction to corporate accounting standards, American Accounting Association.

Titman, S., Wei, K., Xie, X., 2004, Capital investments and stock returns, *Journal of Financial and Quantitative Analysis* 39.677-700.

Table 1

Fama-MacBeth regressions of returns

The table shows average slopes from Fama and MacBeth (1973) regressions of returns on two measures of profitability, and investment. The profitability measures are (a) GP/AT, gross profitability measured as revenues minus cost of goods sold, scaled by total assets, and (b) OP/BE, operating profitability, measured as revenues minus cost of goods sold, minus selling, general and administrative expenses, minus interest expense, scaled by book equity. The investment measures are (a) dA_{t-1}/A_{t-2}, the percentage change in total assets from fiscal year t-2 to fiscal year t-1, and (b) $dB_{t-1}/B_{t-2,}$ the percentage change in book equity from fiscal year t-2 to fiscal year t-1. The regressions control for size (log(ME)), book-to-market (log(B/M), prior month returns ($R_{1,0}$), and momentum ($R_{2,12}$). The sample is separated into large and small firms based on the median NYSE breakpoint in June of each year. Independent variables are trimmed at the 99% and 1% level. Newey-West *T*-statistics appear in parentheses.

	All Stocks				Large Stocks				Small Stocks			
Panel A: July 1940-June 1963												
GP/AT	0.70	-	0.65	-	0.68	-	0.67	-	0.78	-	0.70	-
	(3.66)		(3.55)		(2.81)		(2.93)		(3.49)		(3.08)	
OP/BE	-	0.91	-	0.77	-	1.07	-	0.99	-	0.75	-	0.60
		(3.24)		(3.01)		(3.71)		(3.76)		(2.01)		(1.80)
dA_{t-1}/A_{t-2}	0.01	-0.25	-		-0.06	-0.36	-	-	0.07	-0.13	-	-
	(0.04)	(1.17)			(0.23)	(1.48)			(0.26)	(0.47)		
dB_{t-1}/B_{t-2}	-	-	0.26	0.22	-	-	0.07	-0.01	-	-	0.48	0.46
			(1.34)	(1.30)			(0.27)	(0.01)			(1.99)	(2.26)
$R_{2,12}$	1.29	1.28	1.32	1.30	1.46	1.45	1.53	1.49	1.22	1.22	1.22	1.23
	(5.60	(5.44)	(5.79)	(5.60)	(5.20)	(5.09)	(5.43)	(5.24)	(5.05)	(5.04)	(5.13)	(5.14)
$R_{1,0}$	-8.48	-8.48	-8.52	-8.49	-7.91	-7.79	-8.04	-7.92	-8.94	-9.03	-8.95	-8.93
	(12.07)	(11.74)	(12.05)	(11.69)	(9.74)	(9.12)	(9.79)	(9.18)	(11.73)	(11.83)	(11.83)	(11.78)
log(B/M)	0.36	0.38	0.37	0.39	0.31	0.35	0.33	0.36	0.40	0.40	0.40	0.39
	(3.67)	(3.22)	(3.72)	(3.25)	(2.89)	(2.74)	(3.00)	(2.86)	(3.88)	(3.22)	(3.70)	(3.12)
log(ME)	-0.07	-0.09	-0.07	-0.09	-0.04	-0.05	-0.03	-0.05	-0.16	-0.18	-0.16	-0.17
	(1.75)	(2.08)	(1.73)	(2.06)	(1.12)	(1.47)	(1.01)	(1.35)	(1.84)	(2.03)	(1.78)	(1.91)
Panel B: July 1926-June 1963												
dA_{t-1}/A_{t-2}		-0.93		-		-0.18		-		-2.23		-
		(1.97)				(0.60)				(2.37)		
dB_{t-1}/B_{t-2}		-		-0.28		-		-0.24		-		-0.29
				(0.80)				(0.62)				(0.53)
$R_{2,12}$		1.08		1.06		1.57		1.59		0.69		0.65
		(3.40)		(3.23)		(5.28)		(5.36)		(2.46)		(2.41)
$R_{1,0}$		-8.84		-8.88		-6.71		-6.85		-10.42		-10.51
		(11.04)		(11.80)		(7.52)		(7.59)		(11.03)		(11.04)
log(B/M)		0.07		0.09		0.11		0.13		0.01		0.02
		(0.60)		(0.82)		(0.82)		(1.01)		(1.12)		(1.13)
log(ME)		-0.15		-0.15		-0.04		-0.03		-0.41		-0.43
		(2.26)		(2.30)		(1.03)		(0.81)		(2.30)		(2.47)

Table 2

Excess returns and 3-factor model estimates for portfolios sorted on measures of profitability and investment

Each June, five portfolios (labelled Low through High) are formed on two measures of profitability and investment. The profitability measures are (a) GP/AT, gross profitability measured as revenues minus cost of goods sold, scaled by total assets, and (b) OP/BE, operating profitability, measured as revenues minus cost of goods sold, minus selling, general and administrative expenses, minus interest expense, scaled by book equity. The investment measures are (a) dA_{t-1}/A_{t-2}, the percentage change in total assets from fiscal year t-2 to fiscal year t-1, and (b) dB_{t-1}/B_{t-2}, the percentage change in book equity from fiscal year t-2 to fiscal year t-1. The table shows average excess returns, and estimates from time series regressions of portfolio returns on the three-factor model. *T*-statistics appear in parentheses.

	R^e	α	MKT	SML	HML	R^e	α	MKT	SML	HML
Panel A: Portfolios formed on measures of profitability, July 1940-June 1963										
			GP/AT					OP/BE		
Low	0.96	-0.17	0.87	0.27	0.32	1.00	-0.32	0.96	0.26	0.50
	(4.03)	(2.14)	(39.36)	(6.37)	(9.28)	(3.75)	(3.85)	(43.11)	(5.96)	(14.14)
2	0.98	-0.16	0.95	0.10	0.24	0.95	-0.13	0.91	0.13	0.21
	(4.13)	(2.36)	(50.26)	(2.84)	(7.81)	(4.23)	(2.15)	(54.28)	(3.93)	(7.81)
3	1.09	-0.12	1.08	-0.08	0.15	0.98	-0.10	0.97	0.03	0.10
	(4.29)	(1.77)	(57.38)	(2.35)	(5.02)	(4.30)	(1.67)	(62.04)	(1.06)	(4.06)
4	1.08	-0.03	1.09	-0.03	-0.07	1.12	0.04	1.04	-0.03	-0.03
	(4.29)	(0.51)	(58.30)	(1.05)	(2.37)	(4.68)	(0.60)	(65.63)	(0.96)	(1.28)
High	1.11	0.24	0.96	0.17	-0.29	1.10	0.05	1.09	0.09	-0.20
	(4.95)	(3.31)	(50.76)	(4.65)	(9.53)	(4.40)	(0.72)	(63.29)	(2.89)	(7.36)
High-Low	0.15	0.41	0.08	-0.10	-0.61	0.10	0.37	0.13	-0.16	-0.70
	(1.07)	(3.69)	(2.97)	(1.78)	(12.95)	(0.70)	(3.58)	(5.06)	(3.00)	(16.10)
Panel B: Portfolios formed on measures of investment, July 1940-June 1963										
			dA_{t-1}/A_{t-2}					dB_{t-1}/B_{t-2}		
Low	1.09	-0.07	0.99	0.26	0.18	0.93	-0.21	0.93	0.21	0.24
	(4.35)	(0.98)	(46.43)	(6.34)	(5.34)	(3.90)	(2.81)	(46.50)	(5.25)	(7.55)
2	1.01	-0.10	0.92	0.15	0.22	0.91	-0.15	0.89	0.08	0.19
	(4.40)	(1.66)	(55.93)	(4.63)	(8.38)	(4.10)	(2.35)	(50.55)	(2.47)	(7.00)
3	1.01	-0.06	0.97	-0.07	0.12	1.06	-0.01	0.97	-0.04	0.09
	(4.50)	(1.23)	(67.26)	(2.49)	(5.35)	(4.70)	(0.11)	(62.26)	(1.50)	(3.78)
4	1.09	0.02	1.03	-0.01	-0.05	1.21	0.09	1.09	0.03	-0.05
	(4.59)	(0.42)	(62.77)	(0.36)	(1.93)	(4.87)	(1.52)	(66.98)	(1.03)	(2.07)
High	1.02	-0.09	1.12	0.15	-0.17	1.07	-0.07	1.13	0.20	-0.14
	(3.92)	(1.38)	(62.92)	(4.32)	(6.01)	(4.04)	(0.99)	(59.77)	(5.50)	(4.79)
High-Low	-0.07	-0.02	0.14	-0.11	-0.35	0.14	0.14	0.20	-0.01	-0.38
	(0.73)	(0.16)	(5.53)	(2.27)	(8.75)	(1.16)	(1.35)	(7.19)	(0.04)	(8.67)
Panel C: Portfolios formed on measures of investment, July 1926-June 1963										
			dA_{t-1}/A_{t-2}					dB_{t-1}/B_{t-2}		
Low	0.90	-0.09	1.09	0.22	0.15	0.79	-0.24	1.11	0.52	0.05
	(2.23)	(0.80)	(49.36	(5.98)	(4.43)	(1.79)	(1.43)	(35.35)	(9.55)	(1.05)
2	0.76	-0.12	1.00	0.07	0.10	0.63	-0.27	0.92	0.11	0.28
	(2.13)	(1.22)	(55.28)	(2.27)	(3.84)	(1.78)	(3.23)	(58.95)	(4.35)	(11.55)
3	0.77	-0.08	0.95	-0.01	0.18	0.93	0.07	0.97	0.16	0.08
	(2.23)	(0.87)	(52.65)	(0.27)	(6.42)	(2.63)	(0.62)	(45.31)	(4.50)	(2.55)
4	0.74	-0.05	1.00	-0.01	-0.03	0.81	0.05	0.98	0.01	-0.05
	(2.26)	(0.69)	(72.61)	(0.33)	(1.82)	(2.55)	(0.62)	(64.97)	(0.61)	(2.30)
High	0.60	-0.14	1.05	0.01	-0.23	0.66	-0.11	1.09	0.02	-0.24
	(1.83)	(1.78)	(69.47)	(0.74)	(10.00)	(1.95)	(1.30)	(68.84)	(0.90)	(9.86)
High-Low	-0.30	-0.05	-0.03	-0.20	-0.38	-0.13	0.13	-0.01	-0.49	-0.29
	(1.67)	(0.34)	(1.19)	(4.42)	(9.13)	(0.53)	(0.64)	(0.49)	(7.57)	(5.05)

Table 3

Univariate and multivariate regressions to predict profitability and investment

The table shows average slopes from Fama-MacBeth regressions to predict operating profitability ($OP_{t+\tau}$ / BE_t) and investment ($dA_{t+\tau}$ / A_t), one two and three years ahead (τ=1, 2, 3). Panel A contains univariate regressions. Panel B contains multivariate regressions that control for size (log(ME)) and book-to-market ratios (log(B/M)). Book equity (BE_t) and assets (A_t)) are measured at the end of the fiscal year t. T-statistics appear in parentheses.

	1964-2015			1940-1963		
τ	1	2	3	1	2	3
Panel A: Univariate regressions						
Predicting $dA_{t+\tau}$ / A_t						
Intercept	11.33	25.51	46.67	6.34	14.23	23.77
	(16.32)	(18.38)	(19.97)	(8.32)	(10.10)	(14.54)
dA_t / A_t	0.25	0.43	0.62	0.21	0.37	0.45
	(13.88)	(11.30)	(10.24)	(6.64)	(6.00)	(5.92)
Adj-R^2	0.03	0.02	0.02	0.03	0.03	0.03
Predicting $OP_{t+\tau}$ / BE_t						
Intercept	0.05	0.08	0.11	0.05	0.09	0.12
	(10.66)	(1.77)	(11.01)	(5.74)	(6.41)	(6.84)
OP_t / BE_t	0.87	0.86	0.88	0.94	0.92	0.93
	(45.13)	(32.35)	(23.26)	(30.29)	(16.41)	(14.22)
Adj-R^2	0.38	0.21	0.14	0.58	0.40	0.29
Panel B: Multivariate regressions						
Predicting $dA_{t+\tau}$ / A_t						
Intercept	11.39	34.75	66.89	8.86	21.14	36.89
	(9.73)	(12.46)	(13.89)	(8.48)	(10.69)	(14.84)
dA_t / A_t	0.16	0.23	0.33	0.15	0.25	0.27
	(9.66)	(8.41)	(8.28)	(4.67)	(4.19)	(3.89)
Log(ME)	-1.00	-3.86	-7.58	-0.60	-1.71	-3.28
	(4.47)	(7.28)	(8.95)	(3.01)	(4.21)	(5.96)
Log(B/M)	-13.38	-27.99	-42.97	-4.62	-9.63	-15.19
	(15.79)	(16.68)	(17.81)	(13.08)	(15.42)	(20.67)
Adj-R^2	0.07	0.07	0.07	0.07	0.07	0.07
Predicting $OP_{t+\tau}$ / BE_t						
Intercept	0.03	0.08	0.13	0.09	0.16	0.23
	(3.00)	(5.89)	(7.15)	(7.15)	(7.53)	(8.28)
OP_t / BE_t	0.82	0.79	0.81	0.83	0.77	0.74
	(42.12)	(28.15)	(22.40)	(22.92)	(12.91)	(10.84)
Log(ME)	0.01	0.00	-0.01	-0.01	-0.01	-0.02
	(2.90)	(0.12)	(1.49)	(2.54)	(2.51)	(3.25)
Log(B/M)	-0.03	-0.05	-0.07	-0.05	-0.07	-0.10
	(5.58)	(6.66)	(7.24)	(9.67)	(10.87)	(11.27)
Adj-R^2	0.45	0.27	0.21	0.59	0.41	0.32

Table 4

Average monthly excess returns for 5x5 portfolios, July 1940-June 1963

Five size portfolios (small through big) are formed based on market capitalization at the end of each June. Independently, five portfolios (low through high), are formed based on book-to-market, operating profitability (OP/BE) and investment (dA_{t-1}/A_{t-2}) respectively at the end of each June. Panel A shows excess returns for these 5x5 portfolios. In panel B, 5x5 portfolios are formed at the intersection of pairs of book-to-market, operating profitability and investment.

Panel A: 5x5 Portfolios formed on size and B/M, size and OP/BE, and size and dA_{t-1}/A_{t-2}

	Low	2	3	4	High
Size and B/M portfolios					
Small	1.07	1.08	1.29	1.45	1.83
2	0.97	1.25	1.21	1.41	1.86
3	1.03	1.13	1.20	1.41	1.47
4	0.98	1.13	1.37	1.26	1.63
Big	0.95	0.87	1.29	1.21	1.43
Size and OP/BE portfolios					
Small	1.59	1.40	1.57	1.43	1.54
2	1.09	1.45	1.40	1.37	1.32
3	1.02	1.23	1.32	1.39	1.24
4	1.06	0.98	1.05	1.27	1.22
Big	0.87	0.86	1.01	0.93	1.10
Size and dA_{t-1}/A_{t-2} portfolios					
Small	1.33	1.69	1.55	1.54	1.21
2	1.25	1.30	1.50	1.23	1.35
3	1.33	1.32	1.25	1.25	1.18
4	1.18	1.23	1.13	1.20	1.07
Big	0.86	0.85	1.05	0.99	0.99

Panel B: 5x5 Portfolios formed on pairs of B/M, OP/BE, and dA_{t-1}/A_{t-2}

B/M and OP/BE portfolios					
	Low OP/BE	2	3	4	High OP/BE
Low B/M	0.12	0.34	0.95	0.97	1.16
2	0.65	0.67	0.89	0.97	1.47
3	0.93	0.84	1.31	1.31	1.73
4	0.86	1.15	1.19	1.34	1.99
High B/M	1.03	1.08	1.46	2.03	2.15
OP/BE and dA_{t-1}/A_{t-2} portfolios					
	Low dA_{t-1}/A_{t-2}	2	3	4	High dA_{t-1}/A_{t-2}
Low OP/BE	0.79	1.10	1.12	1.08	0.68
2	1.14	0.91	1.01	0.98	0.99
3	1.04	0.96	1.13	1.06	0.97
4	1.05	1.01	1.08	1.02	1.07
High OP/BE	1.10	0.91	1.05	1.16	1.19
dA_{t-1}/A_{t-2} and B/M portfolios					
	Low B/M	2	3	4	High B/M
Low dA_{t-1}/A_{t-2}	0.70	1.18	1.01	1.27	1.26
2	0.63	0.91	1.02	1.14	1.27
3	0.95	0.90	1.19	1.29	1.66
4	0.96	1.00	1.36	1.08	1.48
High dA_{t-1}/A_{t-2}	1.02	0.99	1.15	1.28	1.61

Table 5

Average monthly excess returns for portfolios and factors, July 1940-June 1963

In panel A, portfolios are constructed at the intersection of two size groups (small and big) using the median NYSE breakpoint and three groups based on the 30th and 70th percentile of B/M, OP/BE, and dA_{t-1}/A_{t-2}. For B/M, these are labelled small-low (SL), small-neutral (SN), small-high (SH), big-low (BL), big-neutral (BN) and big-high (BH). For OP/BE, these are labelled small-weak (SW), small-neutral (SN), small-robust (SR), big-weak (BW), big-neutral (BN), and big-robust (BR). For dA_{t-1}/A_{t-2}, these are labelled small-conservative (SC), small-neutral (SN), small-aggressive (SA), big-conservative (BC), big-neutral (BN), and big-aggressive (BA). Panel B contains returns for factor building blocks defined in Fama and French (2015. Panel C contains SMB, HML, RMW and CMA factors built from the 2x3 portfolios in Panel A.

Panel A: 2x3 Portfolios formed on size and BM, size and OP/BE, and size and dA_{t-1}/A_{t-2}

Size & BM	SL	SN	SH	BL	BN	BH
Mean	1.17	1.35	1.73	1.03	1.29	1.60
Std. Dev.	4.59	4.48	5.49	3.78	3.54	4.80
Size & OP/BE	SW	SN	SR	BW	BN	BR
Mean	1.35	1.55	1.48	1.07	1.09	1.22
Std. Dev.	4.82	4.48	4.70	4.12	3.72	4.02
Size & dA_{t-1}/A_{t-2}	SC	SN	SA	BC	BN	BA
Mean	1.45	1.50	1.42	1.12	1.11	1.16
Std. Dev.	4.82	4.49	4.79	3.93	3.69	4.19

Panel B: 2x3 Factor building blocks

	SMB_{bm}	SMB_{op}	SMB_{inv}	HML_s	HML_b	HML_{s-b}
Mean	0.10	0.33	0.32	0.55	0.55	0.00
Std. Dev.	1.94	1.83	1.92	2.47	2.84	2.30
	RMW_s	RMW_b	RMW_{s-b}	CMA_s	CMA_b	CMA_{s-b}
Mean	0.13	0.15	-0.02	0.03	-0.04	0.01
Std. Dev.	1.79	2.09	2.11	1.52	1.61	0.01

Panel C: 2x3 Factors

Factors	SMB	HML	RMW	CMA
Mean	0.29	0.43	0.14	-0.01
Std. Dev.	1.83	2.32	1.64	1.26

Table 6

Tests of three-, four-, and five factor models, July 1940-June 1963

Test portfolios in Panels A, B and C are 25 size and B/M, size and OP/BE, and size and dA_{t-1}/A_{t-2} portfolios from Panel A of Table 4. The first column shows the factors that augment the MKT and SMB in time series regressions. The second and third columns show the GRS statistic and its p-value. The fourth column shows the average absolute value of the intercepts. The fifth column shows the average absolute value of the intercept over the average absolute value of $\bar{r}_i$, which is the average return on portfolio *i* minus the average of the portfolio returns.

	GRS	p (GRS)	$A\|\alpha_i\|$	$\frac{A\|\alpha_i\|}{A\|\bar{r}_i\|}$
Panel A: Size - B/M portfolios				
HML	2.21	0.00	0.157	0.79
HML, RMW	1.86	0.01	0.135	0.69
HML, CMA	2.18	0.00	0.154	0.78
RMW, CMA	2.42	0.00	0.176	0.89
HML, RMW, CMA	1.89	0.01	0.138	0.70
Panel B: Size - OP/BE portfolios				
HML	1.69	0.02	0.130	0.70
RMW	1.23	0.21	0.105	0.57
HML, RMW	1.25	0.19	0.119	0.64
HML, CMA	1.68	0.03	0.126	0.68
RMW, CMA	1.25	0.19	0.108	0.58
HML, RMW, CMA	1.26	0.18	0.121	0.65
Panel C: Size - dA_{t-1}/A_{t-2} portfolios				
HML	1.72	0.02	0.122	0.82
CMA	1.75	0.02	0.121	0.81
HML, RMW	1.74	0.02	0.121	0.80
HML, CMA	1.76	0.02	0.121	0.81
RMW, CMA	1.78	0.01	0.124	0.83
HML, RMW, CMA	1.70	0.02	0.121	0.81
Panel D: OP/BE - B/M portfolios				
HML	1.31	0.15	0.203	0.73
HML, RMW	1.17	0.27	0.176	0.75
HML, CMA	1.31	0.15	0.202	0.73
RMW, CMA	1.20	0.23	0.227	0.76
HML, RMW, CMA	1.07	0.37	0.177	0.76
Panel D: OP/BE - dA_{t-1}/A_{t-2} portfolios				
HML	1.68	0.03	0.146	1.80
HML, RMW	1.41	0.09	0.149	1.83
HML, CMA	1.58	0.04	0.145	1.79
RMW, CMA	1.42	0.09	0.139	1.70
HML, RMW, CMA	1.40	0.10	0.150	1.85
Panel E: dA_{t-1}/A_{t-2} - B/M portfolios				
HML	1.40	0.10	0.141	0.55
HML, RMW	1.59	0.04	0.144	0.48
HML, CMA	1.41	0.09	0.140	0.55
RMW, CMA	1.63	0.03	0.146	0.62
HML, RMW, CMA	1.56	0.04	0.145	0.48

Table 7

Regression of one factor on the remaining four factors

For the July 1940 to June 1963 period, the SMB, RMW and CMA factors are built from the 2x3 portfolios in Table 6. For the July 1963-December 2015 period, these factors (along with R_M-R_F and HML) are taken from Ken French's website. T-statistics appear in parentheses.

	α	R_M-R_F	SMB	HML	RMW	CMA	Adj-R^2
Panel A: July 1940 – June 1963							
R_M-R_F	0.51	-	0.31	0.86	0.56	-0.99	0.21
	(2.39)		(2.67)	(6.77)	(2.87)	(4.78)	
SMB	0.14	0.08	-	0.14	-0.06	0.21	0.13
	(1.34)	(2.67)		(2.07)	(0.68)	(1.90)	
HML	0.35	0.16	0.11	-	-0.85	0.30	0.60
	(3.70)	(6.77)	(2.07)		(11.91)	(3.20)	
RMW	0.26	0.05	-0.02	-0.40	-	-0.41	0.62
	(4.14)	(2.87)	(0.68)	(11.91)		(6.73)	
CMA	0.05	-0.07	0.06	0.12	-0.34	-	0.46
	(0.92)	(4.78)	(1.90)	(3.20)	(6.73)		
Panel B: July 1940 – December 2015							
R_M-R_F	0.81	-	0.26	0.21	-0.34	-1.04	0.21
	(6.27)		(5.39)	(3.43)	(5.30)	(11.36)	
SMB	0.28	0.11	-	0.02	-0.41	-0.07	0.15
	(3.22)	(5.39)		(0.47)	(9.84)	(1.12)	
HML	0.11	0.06	0.01	-	0.03	1.01	0.43
	(1.58)	(3.43)	(0.47)		(1.03)	(25.41)	
RMW	0.38	-0.08	-0.23	0.03	-	-0.28	0.16
	(5.79)	(5.30)	(9.84)	(1.03)		(5.75)	
CMA	0.16	-0.11	-0.01	0.40	-0.12	-	0.51
	(3.78)	(11.36)	(1.12)	(25.41)	(5.75)		
Panel C: Subperiod: July 1963- December 1980							
R_M-R_F	0.32	-	0.47	-0.26	-0.59	-0.74	0.29
	(1.20)		(5.54)	(1.63)	(2.53)	(3.43)	
SMB	0.49	0.27	-	0.39	-0.44	-0.65	0.26
	(2.46)	(5.54)		(3.27)	(2.55)	(4.00)	
HML	0.16	-0.05	0.12	-	-0.44	0.70	0.61
	(1.42)	(1.63)	(3.27)		(4.70)	(8.67)	
RMW	0.26	-0.05	-0.06	-0.21	-	-0.38	0.53
	(3.34)	(2.53)	(7.98)	(4.70)		(6.23)	
CMA	0.19	-0.07	-0.11	0.38	-0.42	-	0.66
	(2.34)	(3.43)	(4.00)	(8.67)	(6.23)		
Panel D: Subperiod: January 1981- June 1998							
R_M-R_F	1.50	-	0.09	-0.53	-0.44	-0.77	0.29
	(5.61)		(0.84)	(3.40)	(2.24)	(3.53)	
SMB	0.39	-0.03	-	0.29	-0.70	-0.18	0.18
	(2.17)	(0.84)		(2.94)	(5.95)	(1.29)	
HML	0.39	-0.09	0.13	-	-0.24	0.87	0.61
	(3.17)	(3.40)	(2.94)		(2.86)	(11.27)	
RMW	0.54	-0.05	-0.20	-0.15	-	-0.19	0.22
	(5.97)	(2.24)	(5.95)	(2.86)		(2.46)	
CMA	0.22	-0.07	-0.04	0.43	-0.15	-	0.60
	(2.63)	(3.53)	(1.29)	(11.27)	(2.46)		
Panel E: Subperiod: July 1998 - December 2015							
R_M-R_F	0.87	-	0.03	0.56	-0.86	-0.80	0.36
	(3.32)		(0.36)	(5.20)	(8.13)	(5.41)	
SMB	0.47	-0.01	-	0.09	-0.61	0.18	0.27
	(2.37)	(0.36)		(1.05)	(7.55)	(1.57)	
HML	-0.39	0.20	0.05	-	0.50	0.83	0.54
	(2.43)	(5.20)	(1.05)		(7.74)	(11.09)	
RMW	0.53	-0.28	-0.35	0.45	-	-0.15	0.57
	(3.59)	(8.13)	(7.55)	(7.74)		(1.72)	
CMA	0.32	-0.15	0.06	0.45	-0.09	-	0.46
	(2.74)	(5.41)	(1.57)	(11.09)	(1.72)		