Value and profitability premium: from the perspective of duration

Wah Yip Chu[1]

This version: Jan 2017

[1]Norwegian Business School (BI), Department of Financial Economics. E-mail: wah.y.chu@bi.no

Abstract

Traditional duration based model, though successfully explains the value premium, counterfactually predicts that profitable stocks, which have long duration, underperform the unprofitable stocks. This is because the traditional model identifies the stocks' exposures to both cash flow and discount rate risks with only one measure - duration. In contrast, this article allows the duration to be orthogonal to the book-to-market (or profitability) ratios, so that we have two measures that truly span the two dimensional risk metric. I find that the value premium arises mainly from the difference, in terms of a compensation for cash flow risk, between long duration value and long duration growth stocks; while the profitability premium arises mainly from the difference, in terms of a compensation for discount rate risk, between short duration profitable and short duration unprofitable stocks. It is this additional degree of freedom that accommodates both the value and profitability premiums under the same risk-based framework.

Keywords: Asset pricing, ICAPM, Value premium, Profitability premium, Term structure of equity, Cross-section of stock returns, Long-run risk, Linear multifactor models

JEL classification: G10; G12

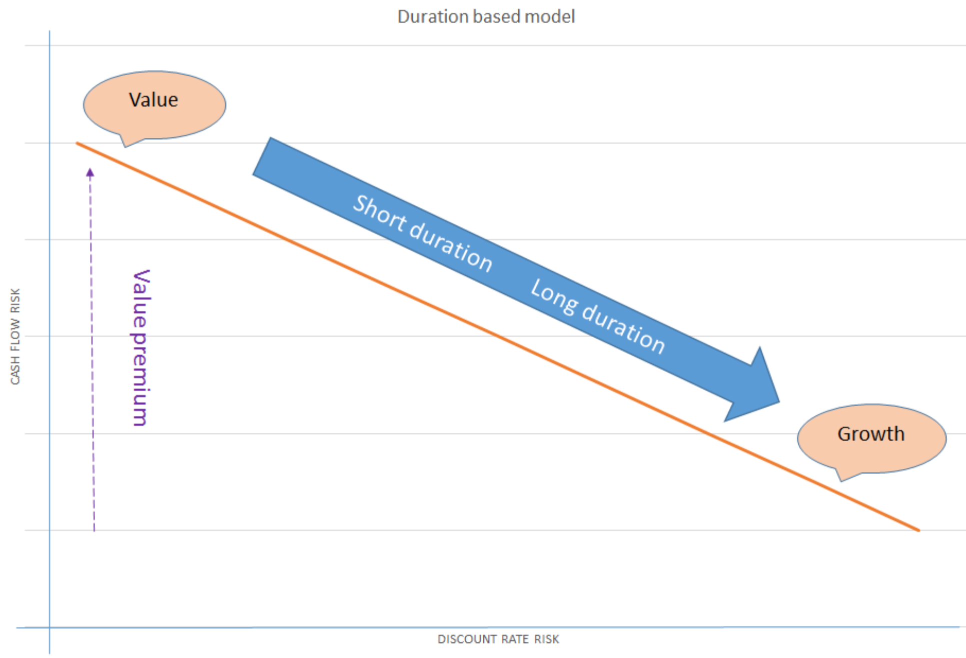
1 Introduction

The value premium is considered as the most important anomaly in the cross sectional asset pricing literature, motivating extensive researches that seek explanations. One of the leading theories is Lettau and Wachter (2007)'s duration-based model. In their model, investors view short-duration assets, which are more sensitive to cash flow shock, riskier than long-duration assets, which are more sensitive to discount rate shock. Value firms have short durations, and consequently are compensated with higher risk premium than long duration growth firms. This model, though successfully fits various empirical facts, [1] fails to explain another anomaly —profitability (Fama and French, 2006; Novy-Marx, 2013). What is interesting about profitability is that profitable firms, similar to growth firms, are associated with long run growth in profits, earnings, free cash flows, and dividends. According to Lettau and Wachter (2007), the profitable firms, which have longer duration than unprofitable firms do, should under-perform the unprofitable firms. Novy-Marx (2013) however finds the opposite: profitable firms, measured by gross profit to asset ratio (GPA), actually generate significantly higher returns than unprofitable firms. How can we understand such empirical finding that contradicts the prediction of duration based model?

Some latest research might help reconcile the facts and the theory. For example, it is possible that profitability anomaly is simply a behavior phenomenon not captured by risk based models. The relevant evidence can be found in Wang and Yu (2013) and Lam, Wang, and Wei (2014), who argue that profitability anomaly arises from mispricing. It is also possible that our traditional interpretation of book-to-market as a measure of value versus growth is inappropriate to begin with. Chen (2014) shows that growth stocks do not really exhibit significantly higher cash flow growth than value stocks do. Alternatively, one could seek explanations for both the value and profitability anomaly using other models such as Hou, Xue, and Zhang (2015), in which the authors propose a characteristics-based factors model that relates stocks' expected return to firm characteristics. Hou, Xue, and Zhang

[1] See Chu, Cooper, Maio, Oded (2016)

Figure 1: Graphical illustration for traditional duration based model



(2015) successfully prices both the value and profitability premiums, however, is silent about the nature of risks captured by these premiums. [2]

This article attempts to fill the gap in the literature by jointly investigating the value and profitability anomalies from a risk-based perspective. I argue that the profitability premium mainly compensates for discount rate risk, while the value premium mainly compensates for cash flow risk, in the spirit of Campbell and Vuolteenaho (2004). [3] This is the first article that suggests a risk-based explanation for profitability premium. More importantly, this article suggests an alternative duration based mechanism that can potentially accommodate both the value and profitability premiums. The main intuition of the mechanism is best illustrated by comparing figure 1 and 2.

Figure 1 shows a graphical illustration for the traditional duration-based model. In this

[2] Lin and Zhang (2013) explains: "different from Fama and French (1993, 1996) characteristics-based factors are not necessarily risk factors... the evidence that characteristics dominate covariances in horse races does not necessarily mean mispricing. Instead, measurement errors in covariances are more likely to blame."

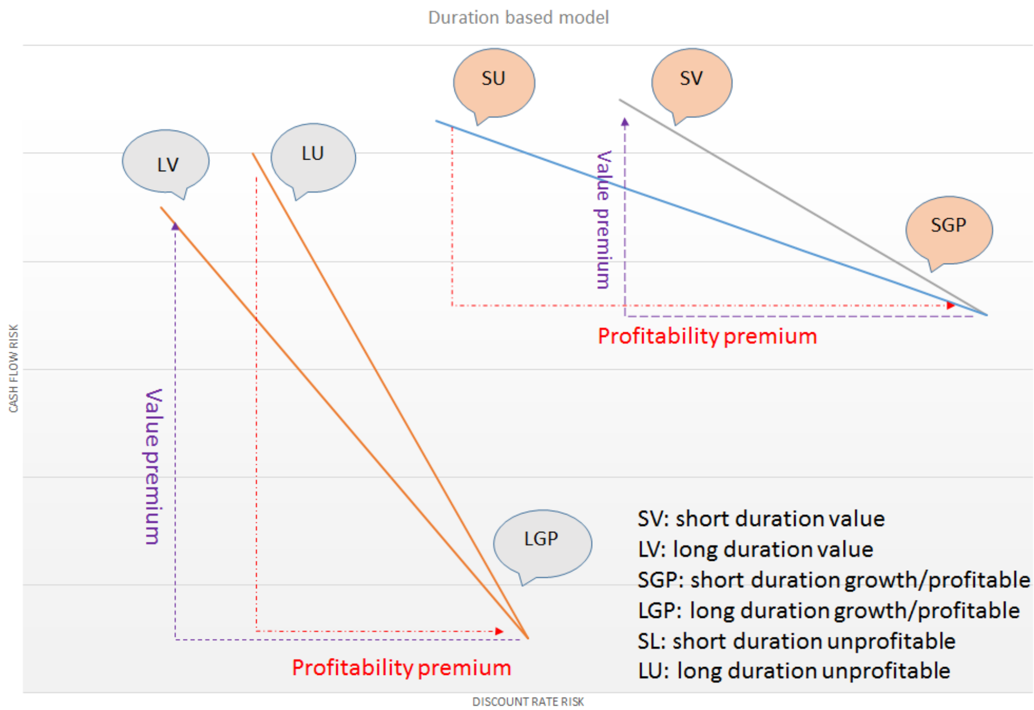
[3] Campbell and Vuolteenaho (2004) suggest a version of ICAPM (Merton, 1973) by decomposing risk premiums into two components, namely the part attributed to cash flow news and the other part attributed to discount rate news.

model, stocks are subject to two dimensions of risks, namely the cash flow risk and discount rate risk, represented by the two axis. The duration (DUR) of a stock measures how much of cash flow versus discount rate risk the stock is exposed to. Obviously a stock can only have either a long or short duration. This implies a stock with high cash flow risk must have relatively low discount rate risk, or vice versa. As a result, the stock market, even though is modelled with such a two dimensional system, essentially is described by just a negative sloping line on figure 1, one end with high cash flow risk while the other end with high discount rate risk. The book-to-market ratio (BM) is just a proxy for duration that measures this line. High BM stocks have short duration, subject to high cash flow risk, while low BM stocks have long duration, subject to high discount rate risk. In order to obtain a positive value premium, the price of cash flow risk needs to be unconditionally higher than the price of discount rate risk. To an extreme, only the cash flow risk matters, this is the reason why the model predicts the unprofitable stocks, having higher cash flow risk than the profitable stocks, should outperform profitable stocks, which however is counterfactual.

The main problem with the traditional duration based model is that only one risk measure, the duration, is used to capture two dimensions of risks. My model differs from the traditional model in that I orthogonalize duration and BM (or GPA), so that I can use the two risk measures, DUR and BM (or GPA), to span the two dimensional risk metric. It is this additional degree of freedom that accommodates both of the value premium, and the profitability premium. We can visualize this idea more clearly in figure 2. In this figure, there exist three type of stocks, including the value stocks, unprofitable stocks, and growth/profitable stocks (the growth stocks and profitable stocks are equivalent in this setting) [4], and each type of stocks is further divided into two groups by their duration orthogonalized to BM or GPA. Hence, there are six portfolios correspondingly represented by the six balls.

[4] Firms with high book-to-market ratio are likely to have high profitability: I find that almost half of constituent stocks in the growth and profitable quintiles are overlapped. This treatment is crucial to understand the interaction between value and profitability premium as shown below.

Figure 2: Graphical illustration of
the relation among value, unprofitable and growth/profitable stock returns



In this model, the value premium arises from the difference, in terms of the compensation for cash flow risk, between the value stocks and growth/profitable stocks, but note that there are two sources of value premium, one from the long duration stocks, the other one from the short duration stocks. The size of value premium of the long duration stocks however is larger than that of the short duration stocks because long duration stocks capture more quantity of cash flow risk. Similarly, the profitability premium arises from the difference, in terms of the compensation for discount rate risk, between the growth/profitable stocks and unprofitable stocks, but this time the short duration stocks contribute more to the profitability premium.

For this model to work, it requires different price of risk for long duration and short duration stocks. Suppose there is only one pair of risk prices, and that the price of cash flow risk dominates the price of discount rate risk as in the traditional duration model, then the profitability premium that captures discount rate risk can never arise. In another case, if the price of cash flow risk and discount rate risk are similar in size, the premium that captures cash flow risk would cancel out the premium that captures discount rate risk. Only if the cash flow risk price dominates for long durations stocks, and discount rate risk price dominates for short duration stocks can both the value and profitability premium survives.

Another justification for my model is that it explains the empirical fact that the value and profitability premiums are negatively correlated. [5] When the market go down, the growth/profitable stocks, no matter long- or short-duration, gain more cash flow and discount rate risks relative to value or unprofitable stocks (growth/profitable stocks are mostly the large market capitalization stocks, hence comove more with the value weighted market portfolio), i.e. both of SGP and LGP in figure 2 move towards the top right corner, consequently the value premium becomes smaller, while profitability premium becomes larger.

Empirically, I first show that the value spread in returns is stronger for the long duration stocks than for the short duration stocks, while the profitability spread in returns is stronger for the short duration stocks than for the long duration stocks. This is done by double sorting the firms by BM and duration, as well as by GPA and duration. Then, using the approaches in Campbell, Polk, and Vuolteenaho (2010), I estimate the cash flow and discount rate betas separately for long- and short-duration stocks. I find that the cash flow betas align with the portfolio returns sorted by BM, and the cash flow beta spread is stronger for the long duration stocks; while the discount rate betas align with the portfolio returns sorted by GPA, and the discount rate beta spread is stronger for the short duration stocks. In sum, for the BM portfolios, the pattern of the portfolios' cash flow betas is consistent with the pattern of the portfolio's returns. For the GPA portfolios, the pattern of the portfolios' discount rate betas is consistent with the pattern of the portfolio's returns. Such consistency holds across different duration.

The prices of the cash flow and discount rate risks are estimated with the cross sectional regressions of average portfolio returns on the estimated cash flow and discount rate betas. My test assets include a set of 10-by-3 portfolios sorted by BM and duration, as well as

[5] Novy-Marx (2013) shows that a profitability strategy, which takes a long-short position in the extreme portfolios sorted by profitability, appears to be a good hedge against the value strategy.

a set of 10-by-3 portfolio sorted by GPA and duration. The estimations are conducted separately for different duration groups. I find that for the longest duration stocks, the price of cash flow risk is as large as 4.82%, while the price of discount rate risk is insignificantly different from zero. However, for the short duration stocks, the price of cash flow risk is 0.9%, while the price of discount rate risk is 0.51%. The high discount rate risk price for short duration stocks, which is comparable to the cash flow risk price, gives rise to the profitability premium. This two factor model can explain over 90% of variations of the BM and GPA portfolio returns.

Finally, to show that my model can explain the fact that the value and profitability premiums are negatively correlated. I first show that the growth and profitable stocks are equivalent to each other by identifying individual stocks for the highest GPA quintile that also appear in the lowest BM quintile. It is the variation of these overlapping stocks that drives both the value and profitability premiums to the opposite directions. I show that the return of the growth/profitable stocks is negatively correlated with value premium, but positively correlated with profitability premium.

The remainder of the paper is organized as follows. Section 2 discusses the data and the methodology. Section 3 to 6 present the main empirical results. Section 7 discusses the interpretations and asset pricing implications. Section 8 concludes.

2 Data and methodology

As discussed in the introduction, one important feature in this article is to orthogonalize DUR with respect to BM (or GPA), so that we have two orthogonal measures to truly span the two dimensional risk metric. This section provides some theoretical background and empirical evidence that justify the empirical strategy.

2.1 Data

In this article, I use all NYSE, Amex, and NASDAQ nonfinancial firms (excluding firms with four-digit SIC codes between 6000 and 6999) listed on the CRSP monthly stock return files and the Compustat annual industrial files. My sample period covers from June 1968 to Dec 2015, in order to ensure a reasonable number of firms in each month, especially in the earlier part of the sample. To mitigate backfilling biases, a firm must be listed on Compustat for 2 years before it is included in the data set (Fama and French, 1993). To minimize the impact of outliers, I winsorize all variables at the 1% and 99% level.

To account for the delisting bias in the CRSP database, I follow Beaver, McNichols, and Price (2007) to incorporate those delisted returns into monthly return data. Delisting data is recorded in daily CRSP database, so that the exact delisting date can be identified, I include only the observations with delisting code larger than 199. If the delisting date happens to be on the last day of the month, then the delisted return is set as the replacement value, which is equal to average delisted returns across all delisted firms with the same delisting code. If delisting date happens to be in the middle of the month, then delisted return is equal to the last monthly CRSP return compounded with replacement value.

At the end of June of each year t, I use NYSE breakpoints to split stocks into portfolios based on firm characteristics, and calculate monthly portfolio returns and the corresponding betas from July of year t to June of t +1. The definition of the firm characteristics are detailed in Appendix. The five Fama & French factors, the one-month Treasury- bill rate come from the Fama & French data library on Ken French's webpage.

Table 1 reports a summary statistics of in terms of the average firm characteristics for portfolios sorted by the book-to-market ratio (BM) and gross profitability (GPA) respectively in panel A and B. In panel A, we see the growth firms firms are associated with high profitability, high sales growth (SG), high investment (IA), large market capitalisation (ME), and long duration (DUR). The procedure of constructing these measures is detailed in the appendix. In panel B, we see that the profitable firms, are associated with high book-to-market ratio , low sales growth, and low duration. Investment and market capitalisation across the profitability portfolios have no significant difference. One important interpretation of panel B is that the profitable firms should have lower return returns than unprofitable firms according to the duration based model. Next subsection give more details about the idea of duration model and how it motivates my empirical strategy.

2.2 Background for duration and cross sectional anomalies

This article is greatly inspired by Campbell and Vuolteenaho (2004). They derive a version of Merton (1973)'s Intertemporal Capital Asset Pricing Model (ICAPM) in which the expected stock returns are determined not by their overall beta with the market, but by two separate betas, one with permanent cash-flow shocks to the market, and the other with temporary shocks to market discount rates:

$$E(r_{i,t} - r_{f,t}) + \frac{\sigma_i^2}{2} = \lambda_{i,t}^{CF} \beta_{i,t}^{CF} + \lambda_{i,t}^{DR} \beta_{i,t}^{DR}, \tag{1}$$

where $\beta_{i,t}^{CF}$ and $\beta_{i,t}^{DR}$ respectively represent the quantity of cash flow and discount rate risk; while $\lambda_{i,t}^{CF}$ and $\lambda_{i,t}^{DR}$ respectively represent the price of cash flow and discount rate risk. One key point of Campbell and Vuolteenaho (2004)'s model is that the cash flow shock is considered riskier than the discount rate shock, and therefore the cash flow risk has higher price. They show that this model can explain the value premium, because value stock, subject to higher cash flow risk than the growth stocks, are compensated with higher returns.

A natural follow up question is: why would the value stocks be subject to high cash flow risk? The duration based models can give some insights (see Lettau and Wachter, 2007, 2011). The main message of the duraiton based models is that the short duration stocks covary more with cash flows, while the long duration stocks covary more with discount rate. The value stocks, considered having short duration, therefore are subject to higher cash flow risk, while growth stocks, considered having long duration, are subject to higher discount rate risk. In short, we can express the relation between DUR and stock returns by:

$$R_{i,t} = a + b_1 DUR_{i,t} + \varepsilon_{i,t}, \tag{2}$$

where $R_{i,t}$ stands for the stock returns.

One important assumption of these papers is that the BM is a perfect proxy for duration, however such assumption does not necessarily hold. Weber (2016) and Chen (2014) finds that growth stocks (low book-to-market stocks) do not have substantially higher future cash-flow growth rates than value stocks. To see this empirically, I look at the firm level correlations between BM(or GPA) and duration. Table 2 shows a correlations matrix between cash flow duration and other accounting variables. Since both the DUR and BM are scaled by the market value, their correlation could be due to the common scaling factor. Correlations among the variables in a panel data could exist in two ways: a time series correlation conditional on a specific firm i, $CORR_i(X_i, Y_i)$; or alternatively a cross sectional correlation conditional on a specific time t, $CORR_t(X'_t, Y'_t)$. Panel A of Table 2 shows an average of time series correlation, i.e. $\frac{1}{N}\sum_i^N CORR_i(X_i, Y_i)$. Panel B of Table 2 shows an average of cross sectional correlation, i.e. $\frac{1}{T}\sum_t^T CORR_t(X_t, Y_t)$. We see that DUR have negative correlation with BM: average time series correlation equals -0.34, and average cross sectional correlation equals -0.44. The correlation between DUR and GPA is weak with only series correlation correlation -0.12, and average cross sectional correlation -0.08.

The fact that DUR and BM(or GPA) are not highly correlated suggests that DUR contains other information not captured by BM (or GPA). Formally, we can represent the relation between DUR and BM by:

$$DUR_{i,t} = a + b_1 BM_{i,t} + b_2 F_{i,t} + \xi_{i,t}, \tag{3}$$

where $F_{i,t}$ is another component orthogonal to BM ratio. [6] This pose a question to the

[6] From the perspective of Lettau and Wachter (2007, 2011), $F_{i,t}$ does not exist, i.e. DUR and BM contain

traditional duration based model: if we substitute equation 3 into 4, which of the BM or F that is orthogonal to BM drives the cross sectional returns? This motivates my study the relation of stock returns with respect to both BM and the part of duration that is orthogonal to BM:

$$R_{i,t} = a + b_1 BM_{i,t} + b_2 F_{i,t} + e_{i,t}. \tag{4}$$

However, before the tests, I need to estimate F first by orthogonalising DUR with respect to BM and GPA as shown below.

2.3 Orthogonalising Duration

As discussed, correlation between DUR and the explanatory variables could exist in both time series and cross section, therefore a two-stage orthogonalisation is conducted to cater for the both time series and cross sectional relation. First, I run a cross sectional regression of DUR on BM and GPA for each time period t.

$$DUR_{i,s} = a + b_1 BM_{i,s} + b_2 GPA_{i,s} + \varepsilon_{i,s}, \quad \forall s = t. \tag{5}$$

I choose to use BM and GPA as my explanatory variables because value and profitability are the two main anomalies in my analysis.

Second, I run a time series regression of the residuals from the first stage regression, $\varepsilon_{i,t}$, on the explanatory variables again for each firm i with its whole history of data:

$$\varepsilon_{j,t} = a + b_1 BM_{j,t} + b_2 GPA_{j,t} + \xi_{j,t}, \quad \forall j = i. \tag{6}$$

The residual ξ obtained in the second stage regression is used as my real measure for cash flow duration, for which I call it "orthogonalized duration" (DUR_r).

For the rest of the paper, duration refers to as the "orthogonalized duration" unless

the same set of information.

specified. My focus is to examine the stock returns in response to BM and DUR_r, as well to GPA and DUR_r.

3 Univariate Analysis

In the coming two sections. I aim at showing the relation between profitability premium and discount rate risk , as well as the relation between value premium and cash flow risk. This section first conduct a univariate analysis to illustrate the monotonic relation between the portfolios returns and the corresponding risk measure. Specifically, I sort the firms into ten deciles by BM and GPA respectively. Then for each decile I compute their stock returns, as well as the cash flow and discount rate betas using two different approaches: a VAR estimation and a direct beta measurement, following Campbell, Polk, and Vuolteenaho (2010).

3.1 A VAR Approach

In this subsection, I use a VAR approach to estimate cash-flow and discount-rate news following Campbell and Vuolteenaho (2004) and Campbell, Polk, and Vuolteenaho (2010). The main idea of this method is to decompose stock returns into two components, one attributed to the cash flow news and the other one attributed to the discount rate news, using a VAR model as in Campbell (1991). These two components are the only shocks that move prices. The detail VAR model is detailed in appendix.

Then, correspondingly, the cash-flow and discount-rate betas can be estimated with respect to the the fitted values of the market's cashflow and discount-rate news. Specifically, I use sample covariances and variances in the formulas (7) and (8), allowing for one additional lag of the news terms. The additional lag is motivated by the possibility that, especially during the early years of our sample period, not all stocks in our test-asset portfolios were traded frequently and synchronously.

$$\beta_{i,CF} = \frac{COV(r_{i,t}, N_{CF,t})}{VAR(N_{CF,t-N_{DR,t}})} + \frac{COV(r_{i,t}, N_{CF,t-1})}{VAR(N_{CF,t-N_{DR,t}})} \tag{7}$$

$$\beta_{i,DR} = \frac{COV(r_{i,t}, N_{DR,t})}{VAR(N_{CF,t-N_{DR,t}})} + \frac{COV(r_{i,t}, N_{DR,t-1})}{VAR(N_{CF,t-N_{DR,t}})} \tag{8}$$

Table 3 reports the stock returns, the cash flow and discount rate beta for the portfolios sorted by BM and GPA respectively in panel A and B. There are 5 rows in each panel. The first row shows the value weighted average of excess returns for the ten BM portfolios, with the growth stocks at the left and value stocks at the right of the table, while the spread between the extreme portfolios are presented in the the column. The second row shows the cash flow beta, computed from the VAR method, for each decile, the third row shows the discount rate beta, computed form the VAR method for each decile. The fourth and fifth row show respectively the cash flow and discount rate beta computed with the direct measurement method, for which I will discuss in the next subsection. Standard errors from bootstrapping are reported in parentheses.

In panel A, first we see that the value spread in terms of excess returns is 0.51% per month, while in panel B the profitability spread in terms of excess returns is 0.32% per month. Although my sample here removes the observations without proper value for duration, different from Novy-Marx (2013), the result are consistent his finding. Moreover, in panel A, the value stocks have higher cash flow betas than the growth stocks, with a significant spread. In contrast, in panel B, cash flow beta is decreasing in profitability. This is why profitability premiums appears so puzzling from the perspective of duration based model, because the profitable stocks, being exposed to less cash flow risk, earn higher returns, which is inconsistent with Campbell and Vuolteenaho (2004) and Lettau and Wachter (2007). However, the discount rate betas perhaps can help reconcile the puzzle, as we see that the profitable stocks have higher discount rate beta than the unprofitable stocks, with a significant spread. This result suggests that profitability premium perhaps captures the discount rate risk. If this is the case, the price of discount rate risk should be not be unconditionally lower than the price of cash flow risk as in Campbell and Vuolteenaho (2004), otherwise the profitable stocks, even though have high discount rate beta than the unprofitable stocks, can never be compensated with a return high enough to give rise to a profitability premium. I discuss such conditional risk prices in latter sections.

3.2 A direct measurement strategy

As a robustness check, I also use the direct measurement method, constructing direct proxies for portfolio-level and market-level cash-flow news and for market-level discount-rate news, which is detailed in the appendix. Then the cash flow and discount rate betas are computed as the regression coefficient of the portfolio-level proxies for cash-flow news onto the proxies for the market's cash-flow news:

$$CF_{k,t} = \alpha_k^{CF} + \beta_k^{CF} CF_{M,t} + \varepsilon_{k,t}, \tag{9}$$

where $CF_{M,t}$ and $CF_{k,t}$ are measures of cash flow news respectively at market and portfolio level. Similarly, I compute the discount rate beta as the regression coefficient of the portfolio-level proxies for discount rate news onto the proxies for the market's discount rate news:

$$DR_{k,t} = \alpha_k^{DR} + \beta_k^{DR} DR_{M,t} + \varepsilon_{k,t}, \tag{10}$$

where $DR_{M,t}$ and $DR_{k,t}$ are measures of discount rate news respectively at market and portfolio level. The subscript k indicates portfolios group k sorted by firm characteristics, t indicates the year at which the portfolio is formed, M indicates the aggregates of firm along the characteristic dimension. Table 3 also reports the cash flow and discount rate beta computed with this direct measurement approach in the fourth and fifth rows. The monotonic patterns of these rows are consistent with those computed with the VAR approach.

In sum, there are two main messages in this section. First, my result in panel A about value premium, using an updated sample with period from 1968 to 2015, are comparable to the results reported in Campbell, Polk, and Vuolteenaho (2010), who also use these two approaches to compute the cash flow beta. Second, discount rate beta aligns with the corresponding portfolio returns sorted by GPA, which suggests that the profitability premium might captures the discount rate risk. The next section shows further evidence to support this argument.

4 Bivariate Analysis

Section 2.2 shows that DUR actually contains other information not captured by BM, and I extract such information by orthogonalising DUR with respect to BM and GPA. The purpose of such orthogonalization is to test equation 4, examining the stock returns along both the orthogonalized DUR and BM (or GPA) dimensions. The motivation of this test exactly corresponds to the model shown in figure 2, , in which I illustrate the mechanism that that gives rise the both value and profitability premiums. I conducts the tests using double sorts.

4.1 Portfolio returns

4.1.1 Value premium

Hypothesis 1 *Value premium, measured by the return difference between value and growth stocks, should be larger for long orthogonalized duration assets than for short orthogonalized duration assets.*

Hypothesis 1 describes the value premium according to figure 1: the value premium is stronger for the long duration stocks than the short duration stocks. To test this hypothesis, I double sort on the firms independently by BM in to 5 groups and orthogonalized DUR into 3 groups, and look at the stock returns of each portfolios.

The panel A of table 4, reports the returns for the 5-by-3 value weighted average portfolios. The portfolio returns are generally increasing in BM across the rows. The spreads between the two extreme BM portfolios, as indicated in the last column, is monotonically increasing in duration from 0.11% to 0.66% per month. Standard errors from bootstrapping are reported in parentheses. This result suggests that a value strategy, that takes a long-short positions in the extreme portfolios sorted by BM, should be better compensated with long-term assets than on short-term assets. The difference between the long duration value premium and short duration value premium is significant with t-value equals 2.6.

Another interesting observation is that the low BM quintiles have their returns monotonically decreasing in duration, while the high BM quintiles have their returns monotonically increasing in duration. This observation is related to the literature about term structure of equity (van Binsbergen, Brandt, and Koijen, 2012; van Binsbergen, Hueskes, Koijen, and Vrugt, 2013; Weber, 2016), in which they find that the term structure of equity is downward sloping.

In this article, duration can be considered as a measure for cash flow timing , suggested by Weber (2016), and the returns for the stocks across different duration can be considered as the term structure of equity. In my case, value premium exhibits an upward sloping term structure. This article however differs from Weber (2016) in two ways. First, Weber (2016) look at the market excess returns for each duration group, while I look at the value and profitability premium for each duration groups. Second, Weber (2016) uses raw measures of duration, while my measure of duration, as shown in section 2.1, is orthogonalized with respect to firm characteristics.

4.1.2 Profitability premium

Hypothesis 2 *Profitability premium, measured by the return difference between profitable and unprofitable stocks, should be larger for short orthogonalized duration assets than for long orthogonalized duration assets.*

Hypothesis 2 describes the profitability premium according to figure 1: the profitability premium is stronger for the short duration stocks than for the long duration stocks. To show this, I repeat the same exercises as in panel A but instead of sorting the firms by duration and GPA, I sort the firms by duration and BM, and the result is shown in the panel B.

We see that portfolio returns are generally increasing in GPA in general. The profitability spread, as indicated in the last column range from 0.03% to 0.42% per month, and more importantly is monotonically decreasing in duration. This means a profitability strategy, that takes a long-short positions in the extreme portfolios sorted by profitability, should be better compensated when used on short duration stocks than on long duration stocks. Such difference is significantly large with t-value equals -1.9.

As above, we also examine profitability quintiles across duration. I find that all of the five profitability quintiles have the returns that are decreasing in duration. As mentioned, one could interpret stock duration as a measure for cash flow timing, and accordingly all the 5 profitability quintiles of stock exhibit downward sloping term structure of equity, consistent with Weber (2016).

4.1.3 Unorthogonalized results

To see the effect of orthgonalization, I repeat the same double sorting analysis of table 4 but with the raw duration without orthogonalization for sorting. The result is reported in table 5. We see from the last column that the value spread still exists, but such spread is no longer increasing in duration, in which the difference in value premium between long and short duration group is -0.06% with insignificant. Similarly, in panel B the difference in profitability premium between long and short duration group shown in the last columns are no longer significant. Such a big contrast of results in the two tables also suggest that the orthogalised duration could contain much more information about asset prices different from what BM and GPA can captures.

4.2 Cash flow and discount rate beta

Previous subsection shows the monotonic relation between the value and profitability premium across duration. This section argues that such monotonic return patterns can be explained by their exposures to the cash flow and discount rate risks.

4.2.1 Value premium and cash flow beta

Hypothesis 3 *The value premium mainly captures the cash flow risk. Accordingly, the cash flow beta spread between value and growth stocks should aligns with their corresponding returns, and such beta spread should becomes stronger across the orthogonalized duration.*

Hypothesis 3 describes the mechanism that drives the value premium according to figure 1: the value premium is mainly attributed to the long duration stocks, which captures more cash flow risk. To test the hypothesis 3, I again double sort the stocks independently into 5 groups by BM and 3 groups by orthogonalized duration. I use only 3 groups of duration, because the available data for computing cash flow and discount rate news is much less than that for computing returns. I tradeoff number of portfolios for the number of stocks in each portfolio. Then for each portfolio, I compute the cash flow and discount rate betas using the VAR approach as mentioned. Table 6 shows the corresponding results. In general, panel A shows that the cash flow betas are increasing in BM across the rows, which is consistent with the table 3. Such beta spreads is the strongest for the long duration stocks, but the beta spreads of the short duration stocks are also strong in this test, which equals to 0.07 with t-value 2.1. In contrast panel B of the table shows the discount rate betas computed with the VAR approach. We see that the discount rate betas are decreasing in BM across the rows, which are again consistent with the table 3, and such discount rate beta spread is the strongest for the short duration stocks, which equals to -0.27 with t-value 3.6.

As a robustness check, table 7 repeats table 6 except that the betas are now computed with direct measurement approach. We see now that the spread in cash flow betas is also monotonically increasing in duration, consistent with the patterns of value premium across duration. This result can be considered as an additional restriction for testing equation 4 in the sense that the equation is tested conditional on each duration.

4.2.2 Profitability premium and discount rate beta

Hypothesis 4 *The profitability premium mainly captures the discount rate risk. Accordingly, the discount rate beta spread between profitable and unprofitable stocks should aligns with the profitability premium, and such beta spread should becomes weaker across the orthogonalized duration.*

Hypothesis 4 describes the mechanism that drives the profitability premium according to figure 1: the profitability premium is mainly attributed to the short stocks, which captures more discount rate risk. To examine this hypothesis 4, I again repeat the same exercises as in table 6, but instead of sorting stocks by BM and duration, now I sort the the stocks by GPA and duration. Then for each group of duration, I compute the cash flow and discount rate beta using the both the VAR and direct measurement approach mentioned.

Table 8 shows the results from the VAR approach. In general, shown in panel B, the discount rate betas increases across the GPA portfolios, which is consistent with the results in table 3. More importantly, such beta spreads also exhibit monotonic patterns along duration, consistent with the patterns of profitability premium as in table 4 panel B. The spread of discount rate beta for short duration stocks is higher than the spread of discount rate beta for long duration stocks by 0.13. In contrast, panel A of the table shows that the cash flow betas computed with VAR approach is decreasing across the GPA portfolios, which are again consistent with the table 3.

As a robustness check, table 9 repeats table 8 except that the discount rate betas are now computed with direct measurement approach. The discount rate beta spread over GPA is again decreasing in duration. All these results can again be considered as an additional restriction for testing equation 4 conditional on each duration, further supporting the model in figure 2.

5 Pricing Cash-Flow and Discount-Rate Betas

For my model to work, it requires different price of risks for long duration and short duration stocks. Suppose there is only one pair of risk prices for cash flow and discount rate risk, and that the price of cash flow risk dominates the price of discount rate risk as in Campbell and Vuolteenaho (2004), then the profitability premium that captures discount rate risk can never arise. In another case, if the price of cash flow risk and discount rate risk are similar in size, the premium that captures cash flow risk would cancel out the premium that captures discount rate risk. Only if the cash flow risk price dominates for long durations stocks, and discount rate risk price dominates for short duration stocks can both the value and profitability premium survives.

To estimate the prices of risks, I use a set of 10-by-3 portfolios sorted by BM and duration, and a set of 10-by-3 portfolios sorted by GPA and duration. Therefore there are 60 portfolios in total. Then I separately estimate the price of risks for each duration group, which means there are 20 portfolios used on the left-hand side of the first-order condition 1 for each duration group. I evaluate the performance of an unrestricted two-beta model that allows free risk prices for cash-flow. The risk free rate is obtained from Ken French's website. I estimates the parameters with a cross sectional regression:

$$\overline{R}_i - R_f = g_1 \widehat{\beta}_{i,CF} + g_2 \widehat{\beta}_{i,DR} \tag{11}$$

The cash flow and discount rate betas used in this estimation are computed from the VAR model instead of direct measurement method, because the VAR model gives more frequent observations.

Table 10 reports results. Because I estimate the risk prices separately for each of the 3 duration groups, I obtained 3 different pairs of cash flow and discount rate risk prices. The first column of the table indicates the duration group, the second and third column shows the estimates of cash flow and discount rate risk prices. Standard errors are reported in parentheses. The fourth column shows the R^2 of the model in explaining cross sectional returns.

Table 10 shows that the price of cash flow risk is monotonically increasing in duration from 0.9% to 4.82%, while conversely the price of discount rate risk is monotonically decreasing in duration from 0.51% to -0.26%. More importantly, the magnitude of the two risk prices at the same order of magnitude for the short duration group, while the magnitude of the two risk prices differ by one degree in order of magnitude. Such large difference between the price of cash flow risk and price of discount rate risk for long duration stocks is actually consistent with Campbell and Vuolteenaho (2004). In contrast, but discount rate risk cannot be ignored for the short duration stocks, because the price of discount risk is now almost the same as the price of discount rate risk. The performance of such two factor models performs well in explaining the 20 BM and GPA portfolios for each duration with the R^2 over 90%.

Figure 3 provides a visual summary of these results. The figure plots the predicted average excess return on the horizontal axis and the actual sample average excess return on the vertical axis. For a model with a 100-percent estimated R^2, all the points would fall on the 45- degree line displayed in each graph. The points in the figure denote the 60 BM-by-DUR and GPA-by-DUR portfolios

6 Hedging

6.1 Opposite loadings

It seems in the first place that I treat the long duration stocks and short duration stocks as two segregated market. This is not true. The two sides of the stocks are connected in a way that both of long duration and short duration growth/profitable stocks are highly correlated.

The motivation of jointly investigating value and profitability premium is based on Novy-Marx (2013), who points out an interesting fact that a profitability strategy, which takes a long-short position in the extreme portfolios sorted by profitability, appears to be a good hedge against the value strategy. This means the profitability and value strategies must possess opposite sign of loading towards the underlying state variables in order to hedge against each other.

To see this, we inspect again table 3. We see that the relation between portfolio returns and the corresponding cash flow (discount rate) betas sorted by BM is positive(negative), and conversely, the relation between portfolio returns and the corresponding cash flow (discount rate) betas sorted by GPA is negative (positive). Moreover, such opposite monotonic patterns exit even conditional on each orthogonalized duration as shown in table 4.

The dynamics between the value and profitability premiums over time series can also be be explained by this model. The intuition can be seen in figure 2 as mentioned in the introduction. The growth/profitable stocks are mostly the large market capitalization stocks, hence comove more with the value weighted market portfolio. When the market go down, these growth/profitable stocks, no matter long- or short-duration, gain more of both cash flow and discount rate risks relative to value or unprofitable stocks, i.e. both of SGP and LGP in figure 2 move towards the top right corner, consequently the value premium becomes smaller, while profitability premium becomes larger. Similarly, when the market go up, both of the SGP and LGP in figure 2 move towards the bottom left corner, consequently the value premium becomes larger, while profitability premium becomes smaller. It is the variation of the growth or profitability stock returns, which largely represent the market return, that drives both of the value and profitability premiums. We can formulate a testing hypothesis accordingly:

Hypothesis 5 *The return of the growth/profitable stocks are positively correlated with the profitability premium, and negatively correlated with value premium.*

To test such mechanism, I first show in table 11 the number of stocks that is assigned to both of the lowest BM portfolio and highest GPA portfolio, as well as the number of stocks that is assigned to either of the two portfolios. Panel A of the table shows for each quintile portfolio sorted by GPA the amount of stocks that happens to be assigned to a particular

BM quintile. For example, the cell at the bottom left corner shows that there are 37% of the stocks at the highest profitability quintile happens to be assigned to the lowest BM quintile as well. Panel B reports the size of these overlapped stock relative to the average size of the whole market in fractions. For example the cell at the bottom left corner shows that the size of the overlapped stocks are 2.16 times of the average market capitalisation. The main is that almost half of high GPA stocks happens to be low BM stocks. This result justify my treatment for growth and profitable stocks. Table 12 repeats 11 except that now I look at how many stocks for each portfolio sorted by BM are assigned to a particular GPA quintile.

Moreover, to see that it is these growth/profitable stocks that drives both of the value and profitability premiums, I show that the correlations among time series of the market return, the value premium and profitability premium and the returns of the growth/profitable stocks. We see that the value and profitability premiums are negatively correlated with a strong correlation that equals -0.63, consistent with Novy-Marx (2013), who otherwise show that a portfolio that combines both profitability and value strategy greatly improves the portfolio's performance in terms of Sharpe ratio compared with either the profitability or value strategy alone. More importantly the correlation between growth/profitable stock returns is negatively correlated with value premium (with correlation -0.42), but positively correlated with profitability (with correlation 0.40). This is consistent with my hypothesis.

6.2 Alternative strategies

Following this logic, any other risk premiums that captures the cash flow risk as BM does should be able to hedge against GPA strategy. Chu, Cooper, Maio, Oded (2016) surveyed various cross sectional anomalies and find that the cash flow beta is decreasing in investment portfolios, which suggest that investment premium could be another candidate that captures cash flow risk. This section demonstrates that applying investment strategy together with profitability strategy can also achieve similar hedging effect as that in Novy-Marx (2013), because the premiums that capture cash flow risk always move in opposite direction to the profitability premium that captures discount rate risk.

Hypothesis 6 *If investment, proxied by asset growth, capture cash flow risk as BM does, the investment strategy, similar to value strategy, can hedge against the profitability strategy as well.*

Figure 2 shows realized annual Sharpe ratio of the HML, CMA factor (Fama and French, 2015), and a 50/50 mix of the two over the preceding five years at the end of each month between June 1968 and December 2015(dashed line). I use the data obtained from Ken French website. The figure shows that RMW generally performed well in the periods when HML performed poorly, while HML generally performed well in the periods when RMW performed poorly. This figure is a direct comparison to Novy-Marx (2013)'s figure 1. indicating that profitability even measured with profitability used in Fama and French (2015) can also hedge against value.

The point I am trying to make comes from a comparison between figure 4 and 5. If my argument that the hedging function comes from the negative correlation between cash flow and discount rate risk, then replacing HML with another factor that captures cash flow risk would still serve the same hedging purpose again RMW. Figure 5 shows that it is the case, where I replace HML with the investment factor CMA.

7 Discussions

Finally, in this section I discuss some implications of this model that relates to long-run and short-run risks (Bansal and Yaron, 2004; Restoy and Weil, 2011). This section attempts to give an interpretation and discuss the benefits of viewing cross sectional anomalies from the timing perspective.

Why do we want to study the term structure of cross sectional anomalies? The term structure of anomalies tells us which of the long-run versus short-run risk the anomalies are exposed to, so that we can generalize the anomalies as the compensations for only two type of risks, in the spirit of Bansal and Yaron (2004) and Restoy and Weil (2011). This corresponds to Cochrane (2011), who points out that organizing many risk factors out of existing literature is important, because we want to sort some orders out of the chaotic "factor zoo". The benefits of the long-run versus short-run risk categorization can be seen more clearly in comparison with other asset pricing models. For example, ad hoc factor models win great successes in fitting empirical data, but lack theoretical supports for how the factors span the risk universe. Without a solid theoretical ground, these ad hoc factor models can never be claimed as the true models (or not). In other words, we can always argue that there exist other "anomalies" when we use the ad hoc factor models as the benchmark. Another classification frequently used in the literature is cash flow versus discount rate risks (Campbell, 1993, 1996). The major challenge faced by this classification is measurement. The discount rate is unobservable to begin with. Cash flow, which should be discounted at a proper discount rate, can hardly be disentangled from discount rate effect as well. In sum, it is difficult to identify the part of risk premium that is attributed to cash flow or discount rate risk. In contrast, classifying risks into long-run versus short-run is more implementable only if we have a proper measure of timing of expected cash flow. This article propose one of such methods as shown above.

There is a flurry of work now looking at the term structure of risk premiums, [7] which produces a new set of restrictions for model fitting. Empirical papers along this line of literature so far mainly focus on studying time series market excess return, except Weber (2016) who look into cross sectional evidence. This article differs from Weber (2016) in that he look at excess returns for each duration, while I look at the value and profitability premium for each duration. Certainly the methodology used in this article can be applied on other cross sectional anomalies as well. As a demonstration, I replicate the Fama and French (2015)'s factors, but separately for long and short duration group. I find that the HML, CMA and SMB constructed within long duration group have higher returns than

[7] See van Binsbergen and Koijen (2015) for a review of the literature.

that of short duration group, while RMW constructed within short duration group has higher return than that of long duration group. The differences between the long and short duration factors are significant and cannot be explained by the five factor model itself.

Similar to Fama and French (2015), I assign firms into a set of 2x3x2 portfolios independently sorted by duration, firm characteristics of interest, and size. The DUR breakpoint is the NYSE median duration, Size breakpoint is the NYSE median market cap, and the break points for BM, OP, Inv and MOM are the 30th and 70th percentiles within NYSE stocks sample. Then I construct the five factors using only the firms within each duration group. The value factor HML is the average of the two high B/M portfolio returns minus the average of the two low B/M portfolio returns. Equivalently, it is the average of small and big value factors constructed with portfolios of only small stocks and portfolios of only big stocks. The profitability and investment factors of the 2x3 sorts, RMW and CMA, are constructed in the same way as HML except the second sort is either on operating profitability (robust minus weak) or investment(conservative minus aggressive).Like HML, RMW and CMA can be interpreted as averages of profitability and investment factors for small and big stocks. The 2x3 sorts used to construct HML, RMW and CMA produce three Size factors, SMBB/M, SMBOP and SMBInv, which are computed as an average of three large portfolio returns minus the average of three small portfolio returns. The Size factor SMB from the three 2x3 sorts is defined as the average of $SMB_{B/M}$, SMB_{OP}, and SMB_{Inv}.

The time series average of these long and short factors are reported in table 14. Column (2) and (3) shows the average factor return for short and long duration firms correspondingly. The difference between the long and short factor is reported in column (4). For example, short duration HML strategy deliver 0.069% monthly return, while long duration HML strategy deliver 0.435% monthly return. The difference amounts to 0.365% per month. I also reports the volatility of these portfolios in column (5)-(7). Although the volatility of long and short duration portfolios appears statistically significant in some cases, (number in the parenthesis are bootstrapped standard errors) the economic difference is small, for example in case of HML, the volatility difference between long HML and short HML is only 0.094% per month. This suggest that risk in terms return volatility does not account for the return difference. For easier comparison, in the last two columns I also report the returns and volatilities of (Fama and French, 2015)'s factor, for which the data comes from Ken French's website.

Finally in table 15, I run a time series regression of the long duration minus short duration factor returns on Fama and French five factors model. The Jensen's alpha shows up to be significant in case of HML, RMW, meaning that these regular factor models are not able to explain the difference between long and short duration portfolios. Note that since the duration measure needs to be orthogonalised both in time and cross section before analysis, our results are limited to in-sample only, and therefore the portfolio strategy mentioned in this section is not implementable in practise, but for the purpose of analysis, this methodology is sufficient.

8 Conclusion

In this article, I suggest a risk-based explanation for both the value and profitability premiums. I argue that value premium is mainly a compensation for cash flow risk among the long duration stocks, while profitability premium is a compensation for discount rate risk among the short duration stocks. My duration measure is orthogonalized with respect to BM and GPA, which is the main difference from the traditional duration based model that assumes duration and BM are perfect proxies for each other.

The orthogonalization step allows me to analyze the stock returns across both the duration and BM (or GPA). My empirical strategy relies on a set of double sort by duration and BM, as well as by duration and GPA. For each portfolio, I compute the value weighted returns and their corresponding cash flow and discount rate betas. I also estimates the price of cash flow and discount rate risks separately for different duration. The performance of the model is comparable to other ad hoc factor models. Finally, I discuss about the implication of this model in terms of term structure of equity, which can be interesting for future research.

Appendices

A Variable Definitions

ME . The size of a firm used for portfolio formation in year t is simply its market capitalisation (ME) at the end of June of year t.

BM . The book-to-market (BM) ratio of year t is defined the book equity for the fiscal year ending in calendar year t-1 over the market equity as of December t-1, where book equity (BE) as total stockholders' equity plus deferred taxes and investment tax credit (if available) minus the book value of preferred stock. Based on availability, I use the redemption value, liquidation value, or par value (in that order) for the book value of preferred stock. I prefer the shareholders' equity number as reported by Compustat. If these data are not available, I calculate shareholders' equity as the sum of common and preferred equity. If neither of the two are available, I define shareholders' equity as the difference between total assets and total liabilities.

OP . Operating profit(OP) is revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense all divided by book equity.

Inv . Asset growth (IA) is used as a proxy for investment (Inv). IA is computed as the change in total assets (Compustat annual item AT) from the fiscal year ending in year t-2 to the fiscal year ending in t-1, divided by t-2 total assets.

E/P . To construct the Basu (1983) earnings-to-price (E/P) deciles, we use NYSE breakpoints to split stocks into deciles based on E/P at the end of June of each year t . E/P is calculated as income before extraordinary items (Compustat annual item IB) for the fiscal year ending in calendar year t-1 divided by theME(from Compustat or CRSP) at the end of December of t-1. Stocks with negative earnings are excluded. Monthly value-weighted decile returns are calculated from July of year t to June of t +1, and the deciles are rebalanced in June of t +1

SG . Following Lakonishok, Shleifer, and Vishny (1994), we measure sales growth (SG) in June of year t as the weighted average of the annual SG ranks for the prior 5 years, 5 j=1 (6-j)Rank(t-j). The SG for year t-j is the growth rate in sales (COMPUSTAT annual item SALE) from fiscal year ending in t-j-1 to fiscal year ending in t-j . Only firms with data for all 5 prior years are used to determine the annual SG ranks. For each year from t-5 to t-1, we rank stocks into deciles based on their annual SG, and then assign rank i (i =1,...,10) to a firm if its annual SG falls into the ith decile. At the end of June of each year t , we use NYSE breakpoints to assign stocks into deciles based on SG, and calculate monthly value-weighted decile returns from July of year t to June of t +1.

PIA . Following Lyandres, Sun, and Zhang (2008), we measure PIA as changes in gross property, plant, and equipment (Compustat annual item PPEGT) plus changes in inventory (item INVT) scaled by lagged total assets (itemAT).At the end of June of each year t , we use NYSE breakpoints to assign stocks into deciles based on PIA for the fiscal year ending in calendar year t-1, and calculate monthly value-weighted decile returns from July of year t to June of t +1.

IG . Following Xing (2008), we measure investment growth (IG) for the portfolio formation year t as the growth rate in capital expenditure (Compustat annual item CAPX) from the fiscal year ending in calendar year t-2 to the fiscal year ending in t-1. At the end of June of each year t , we use NYSE breakpoints to split stocks into deciles based on IG, and calculate monthly value-weighted decile returns from July of year t to June of t +1.

GPA . Following Novy-Marx (2013), we measure gross profits-to-assets (GP/A) as total revenue (Compustat annual item REVT) minus cost of goods sold (item COGS) divided by total assets (itemAT, the denominator is current, not lagged, total assets). At the end of June of each year t , we use NYSE breakpoints to sort stocks into deciles based on GP/A for the fiscal year ending in calendar year t-1. Monthly value-weighted decile returns are calculated from July of year t to June of t +1, and the deciles are rebalanced in June of t +1.

ROE . ROE is income before extraordinary items (Compustat quarterly item IBQ) divided by 1-quarter-lagged book equity. Book equity is shareholders equity, plus balance-sheet deferred taxes and investment tax credit (item TXDITCQ) if available, minus the book value of preferred stock. Depending on availability, we use stockholders equity (item SEQQ), or common equity (item CEQQ) plus the carrying value of 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 preferred stock.

B Duration definition

The basic cash-flow-duration is defined as in Dechow, Sloan, and Soliman (2004) resembling the traditional Macaulay duration for bonds and hence reflects the weighted average time to maturity of cash flows:

$$DUR_{i,t} = \frac{\sum_{s=1}^{T} s \times CF_{i,t+s}/(1+r)^s}{P_{i,t}} \tag{12}$$

where $DUR_{i,t}$ is the duration of firm i at the end of fiscal year t, $CF_{i,t+s}$ denotes the cash flow at time $t+s$, $P_{i,t}$ is the current price, and r is the expected return on equity, which is set to be a constant. Allowing for firm-specific discount rates ceteris paribus amplifies cross-sectional differences in the duration measure because high-duration firms tend to have lower returns on equity. Variation over time in return on equity, however, does not affect the cross-sectional ordering, and hence had no effect on my later results.

Formula (12) cannot be directly used, however, as stocks do not have a well-defined finite maturity, $t+T$, as bonds do. To address this issue, one can split the equation into two parts - a finite detailed forecasting period and an infinite terminal value:

$$DUR_{i,t} = \frac{\sum_{s=1}^{T} s \times CF_{i,t+s}/(1+r)^s}{P_{i,t}} + (T + \frac{1+r}{r})\frac{P_{i,t} - \sum_{s=1}^{T} CF_{i,t+s}/(1+r)^s}{P_{i,t}}. \tag{13}$$

Formula (13) is computable if we could make accurate forecast of cash flow in the first term, and then assume the second term is paid out as level perpetuity.

To make forecast on cash flow, we start with the following accounting identity:

$$CF_{i,t+s} = E_{i,t+s} - (BV_{i,t+s} - BV_{i,t+s-1}), \tag{14}$$

which says that cash flow to equity for a firm i at each period $t+s$ is equal to the firm's accounting earning at the end of fiscal year $t+s$, $E_{i,t+s}$, minus any change of book value of equity over the year $t+s$, $BV_{i,t+s} - BV_{i,t+s-1}$. Re-arranging the right-hand side of equation (14) gives:

$$CF_{i,t+s} = BV_{i,t+s-1} \times [\frac{E_{i,t+s}}{BV_{i,t+s-1}} - \frac{BV_{i,t+s} - BV_{i,t+s-1}}{BV_{i,t+s-1}})]. \tag{15}$$

Now the two fraction terms in the bracket of equation (15), namely the return on equity (ROE), $E_{i,t+s}/BV_{i,t+s-1}$, and the growth in book equity, $(BV_{i,t+s} - BV_{i,t+s-1})/BV_{i,t+s-1}$, are forecastable if we assume them stationary. I model ROE as a first-order autoregressive process with an autocorrelation coefficient of 0.57 and a long-run mean of 0.12, and the growth in book equity as a first-order autoregressive process with an autocorrelation coefficient of 0.24 and a long-run mean of 0.06. For the starting year (t =0), I measure ROE as income before extraordinary items (item IB) divided by 1-year-lagged book equity (item CEQ), and the book equity growth rate as the annual change in sales (item SALE). Finally, we use a forecasting period of $T = 10$ years and a cost of equity of $r = 0.12$.

C Proxy for cash flow news

I use accounting return on equity (ROE) to construct direct proxies for firm-level and market cash-flow news, and the price-earnings ratio to construct a proxy for market discount-rate news, following Campbell, polk, (2009), Cohen, Polk, and Vuolteenaho (2009).

After portfolio formation, we track the subsequent cash flow proxy (defined below) of our portfolios from year $t+1$ to $t+5$ by keeping the same firms in each portfolio while allowing their weights to drift with returns (as would be implied by a buy-and-hold investment strategy). Because we perform a new sort every year, our final annual data set is three-dimensional: the number of portfolios formed in each sort or characteristic, times the number of years we follow the portfolios, times the time dimension of our panel. Such portfolio formation methodology have been used by Fama and French (1995), Cohen, Polk, and Vuolteenaho (2003, 2009), and Campbell, Polk, and Vuolteenaho (2010), among others..

Specifically, after we sort the sample firms into ten portfolios for each anomaly, we compute the cash-flow beta (β_k^{CF}) for each portfolio k as the slope coefficient of a regression of portfolio's cash-flow measure on the corresponding market portfolio's cash-flow,

$$CF_{k,t} = \alpha_k^{CF} + \beta_k^{CF} CF_{M,t} + \varepsilon_{k,t}, \tag{16}$$

where $CF_{k,t}$ and $CF_{M,t}$ represent the cash-flow proxies for portfolio k and the market portfolio (sorted in year t), respectively. To asses the statistical significance of the beta estimates, we use Newey-West t-ratios (Newey and West, 1987), computed with N lags, where N denotes the horizon used in the construction of the cash-flow proxy (five years).

Following Cohen, Polk, and Vuolteenaho (2009) and Campbell, Polk, and Vuolteenaho (2010), we estimate direct measures of cash-flow news, rather than using indirect cash-flow measures implied by a first-order vector autoregression (VAR). [8] We use two main measures

[8] Campbell (1991), Campbell and Ammer (1993), Campbell and Vuolteenaho (2004), Bernanke and Kuttner (2005), Maio (2013a), among many others, follow the VAR approach to estimate cash-flow news at the aggregate level. Under this approach, cash-flow news is the residual component of the stock return decomposition. Vuolteenaho (2002), Hecht and Vuolteenaho (2006), Campbell, Polk, and Vuolteenaho (2010), and

of cash-flow news. The first measure is similar to those employed by Ball and Brown (1969) and Beaver, Kettler, and Scholes (1970):

$$CF^1_{k,t} \equiv \sum_{j=1}^{N} \rho^{j-1} R^{CF}_{k,t,t+j} \quad \text{where} \quad R^{CF}_{k,t,t+j} = \sum_{i \in k} w_{i,t,t} \frac{X_{i,t,t+j}}{ME_{i,t,t+j-1}}. \tag{17}$$

In the formula above, the first subscript (k or i) represents respectively a portfolio or a firm; the second subscript t indicates the year when the portfolio is sorted; and the third subscript $t+j$ indicates the year when the variable is measured. In this definition, the cash-flow for portfolio k formed at time t ($CF_{k,t}$) is defined as the sum of discounted future return on equity ($R^{CF}_{k,t,t+j}$). ρ is a discount factor, linked to the average log dividend-to-price, which we set to 0.975, as in Cohen, Polk, and Vuolteenaho (2009). The portfolio is tracked for N years after the formation. We set $N = 5$ in our main estimation, but also present the results for $N = 1$ in our sensitivity analysis section. Using $N = 5$ avoids the short-term noise and volatility in earnings, which affects negatively the efficiency of the regression estimates. On the other hand, unlike Cohen, Polk, and Vuolteenaho (2009), we do not consider horizons longer than five years. The reason hinges on our relatively short sample and the fact that using longer horizons would reduce significantly the number of truly independent observations leading to statistical-inference problems in the regression above.[9]

The second equality further defines the cash-flow return on equity at portfolio level. For simplicity, we drop the first and second subscript in the text discussion that follows. We first compute for each firm the return on equity as the ratio of clean surplus earnings X_t for the fiscal year ending in calendar year t to the market value of equity at the end of December of

Maio (2014) employ the same approach to estimate cash-flow news at the stock or portfolio level. On the other, Chen and Zhao (2009), Maio (2014), and Maio and Philip (2015) employ an alternative VAR-based identification in which cash-flow news is estimated directly within the VAR setup, rather than backed-up as the residual from the return decomposition.

[9]This stems from the fact that both variables in the regression contain overlapped terms, which is incorporated in the regression residuals, and as a consequence, the usual t-ratios tend to over-reject the null hypothesis of zero slopes. See Valkanov (2003), Torous, Valkanov, and Yan (2004), Boudoukh, Richardson, and Whitelaw (2008), and Hjalmarsson (2011) for a discussion in the context of predictive long-horizon regressions.

year $t-1$, where $X_t = BE_t - BE_{t-1} + D_t^{gross}$, and D_t^{gross} denotes gross dividends computed from the difference between CRSP total stock returns and returns excluding dividends. This firm-level return on equity $(X_t/ME_t - 1)$ is winsorized at the 1% level. Then we take a weighted average of the return on equity within the portfolio k. We employ both the equal-weighting and value-weighting schemes for our empirical tests. As indicated by the third subscript of $w_{i,t,t}$, the weight assigned for each firm is determined at time t even though the clean surplus earnings and market equity take the values at time $t+j$ and $t+j-1$, respectively. This is because the same portfolio is tracked for N years after the formation, as discussed previously.

D Proxy for discount rate news

Again, I follow Cohen, Polk, and Vuolteenaho (2009) to use annual increments in the market's log P/E ratio, $ln(P/E)_M$. This reflects the findings of Campbell and Shiller (1988a, 1988b), Campbell (1991), and others, that discount-rate news dominates cash-flow news in aggregate returns and price volatility. The resulting news variable is

$$-DR_{M,DR,t+1} = \sum_{k=1}^{K}[\rho^{k-1}\Delta_{t+k}R^{DR}_{k,t,t+j}] \quad \text{where} \quad R^{DR}_{k,t,t+j} = \sum_{i\in k} w_{i,t,t}\ln(P/E). \tag{18}$$

Then I compute the discount rate beta (β_k^{DR}) for each portfolio k as the slope coefficient of a regression of portfolio's cash-flow measure on the corresponding market portfolio's cash-flow,

$$DR_{k,t} = \alpha_k^{DR} + \beta_k^{DR} DR_{M,t} + \varepsilon_{k,t}, \tag{19}$$

where $DR_{k,t}$ and $DR_{M,t}$ represent the discount rate proxies for portfolio k and the market portfolio (sorted in year t), respectively. To asses the statistical significance of the beta estimates, we use Newey-West t-ratios (Newey and West, 1987), computed with N lags, where N denotes the horizon used in the construction of the cash-flow proxy (five years).

To proxy for short run news, I use leading one year of accounting ROE and PE ratios; to proxy for long run news, I use $k = 2$ up to 5 years to emphasize longer-term trends that correspond more closely to the revisions in infinite-horizon expectations that are relevant for stock prices.

E Aggregate VAR

In specifying the aggregate VAR, we follow Campbell and Vuolteenaho (2004) by choosing the same four state variables. Consequently, our VAR specification is one that has proven successful in cross-sectional asset pricing tests. However, we implement the VAR using annual data, rather than monthly data, in order to correspond to our estimation of the firm-level VAR, which is more naturally implemented using annual observations.

E.1 State variables

$$r_{t+1} - E_t r_{t+1} = (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta d_{t+1+j} - (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j r_{t+1+j} \tag{20}$$

$$= N_{CF,t+1} - N_{DR,t+1} \tag{21}$$

where r_{t+1} is a log stock return, d_{t+1} is the log dividend paid by the stock, Δ denotes a one period change, E_t denotes a rational expectation at time t, and ρ is a discount coefficient. N_{CF} denotes news about future cash flows (i.e., dividends or consumption), and N_{DR} denotes news about future discount rates (i.e., expected returns).

To implement this decomposition, we follow Campbell (1991) and estimate the cash-flow news and discount-rate-news series using a VAR model. This VAR methodology first estimates the terms $E_t r_{t+1}$ and $(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}$ and then uses r_{t+1} and equation (1) to back out the cash-flow news.

first estimating the terms $E_t r_{t+1}$ and $(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{t+1+j}$ and then using realizations of r_{t+1} and Equation (2) to back out the cash-flow news.

We assume that the data are generated by a first-order VAR model

$$z_{t+1} = a + \Gamma z_t + u_{t+1} \tag{22}$$

where z_{t+1} is an m-by-1 state vector with r_{t+1} as its first element, a and Γ are an m-by-1 vector and m-by-m matrix of constant parameters, and u_{t+1} an is i.i.d. m-by-1 vector of shocks. Of course, this formulation also allows for higher-order VAR models via a simple redefinition of the state vector to include lagged values.

Provided that the process in Equation (20) generates the data, $t+1$ cash-flow and discount-rate news are linear functions of the $t+1$ shock vector:

$$N_{DR,t+1} = e1'\lambda u_{t+1} \tag{23}$$

$$N_{CF,t+1} = (e1'\lambda + e1'\lambda)u_{t+1} \tag{24}$$

Above, $e1$ is a vector with the first element equal to unity and the remaining elements equal to zero. The VAR shocks are mapped to news by λ, defined as $\lambda \equiv \rho\Gamma(I-\rho\Gamma)^{-1}$, and $e1'\lambda$ captures the long-run significance of each individual VAR shock to discount-rate expectations. The greater the absolute value of a variable's coefficient in the return prediction equation (the top row of Γ), the greater the weight the variable receives in the discount-rate-news formula. More persistent variables should also receive more weight, which is captured by the term $(I-\rho\Gamma)^{-1}$.

The aggregate-VAR state variables are defined as follows. First, the excess log return on the market (reM) is the difference between the annual log return on the CRSP value-weighted stock index (rM) and the annual log risk-free rate, constructed by CRSP as the return from rolling over Treasury bills with approximately three months to maturity. We take the excess return series from Kenneth French's Web site (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/d library.html). The term yield spread (TY) is provided by Global Financial Data and is computed as the yield difference between ten-year constant-maturity taxable bonds and short-term taxable notes, in percentage points. Keim and Stambaugh (1986) and Campbell (1987) point out that TY predicts excess returns on longterm bonds. These papers argue that since stocks are also long-term assets, TY should also forecast excess stock returns, if the expected returns of longterm assets move together. Fama and French (1989) show that TY tracks the business cycle, so this variable may also capture cyclical variation in the equity premium.

We construct our third variable, the log smoothed price-earnings ratio (PE), as the log of the price of the S&P 500 index divided by a ten-year trailing moving average of aggregate earnings of companies in the index. Graham and Dodd (1934), Campbell and Shiller (1988b, 2003), and Shiller (2000) advocate averaging earnings over several years to avoid temporary spikes in the price-earnings ratio caused by cyclical declines in earnings. This variable must predict low stock returns over the long run if smoothed earnings growth is close to unpredictable. We are careful to construct the earnings series to avoid any forward-looking interpolation of earnings, ensuring that all components of the time t earnings-price ratio are contemporaneously observable. This is important because look-ahead bias in earnings can generate spurious predictability in stock returns while weakening the explanatory power of other variables in the VAR system, altering the properties of estimated news terms. Fourth, we compute the small-stock value spread ($V\,S$) using the data made available by Kenneth French on his Web site. The portfolios, which are constructed at the end of each June, are the intersections of two portfolios formed on size (market equity, ME) and three portfolios formed on the ratio of book equity to market equity (BE/ME). The size breakpoint for year t is the median NYSE market equity at the end of June of year t. BE/ME for June of year t is the book equity for the last fiscal year ending in t - 1 divided by ME for December of t - 1. The BE/ME breakpoints are the 30th and 70th NYSE percentiles. At the end of June of year t, we construct the small-stock value spread as the difference between the log(BE/ME) of the small high-book-to-market portfolio and the log(BE/ME) of the small low-book-to-market portfolio, where BE and ME are measured at the end of December of year t - 1.

We include V S because of the evidence in Brennan, Wang, and Xia (2001), Campbell and Vuolteenaho (2004), and Eleswarapu and Reinganum (2004) that relatively high returns for small growth stocks predict low returns on the market as a whole. This variable can be motivated by the ICAPM itself. If small growth stocks have low and small value stocks have high expected returns, and this return differential is not explained by the static CAPM, the ICAPM requires that the excess return of small growth stocks over small value stocks be correlated with innovations in expected future market returns. There are other more direct stories that also suggest that the small-stock value spread should be related to market-wide discount rates. One possibility is that small growth stocks generate cash flows in the more distant future and, therefore, their prices are more sensitive to changes in discount rates, just as coupon bonds with a high duration are more sensitive to interest-ratemovements than are bonds with a low duration (Cornell 1999; Lettau and Wachter 2007). Another possibility is that small growth companies are particularly dependent on external financing and thus are sensitive to equity market and broader financial conditions (Ng, Engle, and Rothschild 1992; Perez-Quiros and Timmermann 2000). Finally, it is possible that episodes of irrational investor optimism (Shiller 2000) have a particularly powerful effect on small growth stocks.

A Tables

Table 1: Summary statistics of deciles portfolios sorted by duration

Panel A										
Variable	Low	2	3	4	5	6	7	8	9	High
BM	0.18	0.33	0.45	0.56	0.68	0.79	0.92	1.08	1.31	1.95
	(0.003)	(0.005)	(0.008)	(0.009)	(0.011)	(0.012)	(0.014)	(0.016)	(0.021)	(0.028)
GPA	0.56	0.45	0.38	0.34	0.31	0.27	0.24	0.21	0.20	0.19
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
SG	84.98	77.34	72.25	71.23	68.24	66.51	65.31	64.62	63.53	59.38
	(0.254)	(0.245)	(0.223)	(0.239)	(0.277)	(0.283)	(0.349)	(0.242)	(0.308)	(0.254)
IA	0.22	0.17	0.14	0.15	0.11	0.12	0.12	0.09	0.12	0.09
	(0.003)	(0.002)	(0.003)	(0.004)	(0.003)	(0.005)	(0.003)	(0.002)	(0.005)	(0.004)
ME	45091	36818	44435	35472	24439	24692	17237	14273	12522	7802.0
	(1870.4)	(1758.9)	(2193.4)	(1933.8)	(1383.4)	(1393.6)	(935.21)	(815.19)	(565.24)	(683.09)
DUR_10	17.70	16.93	16.38	15.90	15.52	15.18	14.67	14.29	13.67	12.58
	(0.025)	(0.035)	(0.047)	(0.059)	(0.063)	(0.072)	(0.081)	(0.087)	(0.110)	(0.122)
Dur_resid	0.17	0.03	0.00	-0.10	-0.09	-0.06	-0.19	-0.16	-0.25	-0.18
	(0.022)	(0.018)	(0.018)	(0.015)	(0.017)	(0.017)	(0.021)	(0.021)	(0.034)	(0.033)
Panel B										
Variable	Low	2	3	4	5	6	7	8	9	High
BM	0.96	0.91	0.80	0.71	0.65	0.58	0.48	0.40	0.34	0.27
	(0.016)	(0.013)	(0.011)	(0.011)	(0.010)	(0.009)	(0.009)	(0.008)	(0.006)	(0.005)
GPA	0.07	0.13	0.18	0.23	0.28	0.34	0.40	0.47	0.58	0.80
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
SG	66.32	65.53	68.88	69.08	70.21	71.79	73.60	74.99	77.50	81.91
	(0.372)	(0.321)	(0.250)	(0.305)	(0.363)	(0.316)	(0.298)	(0.280)	(0.221)	(0.295)
IA	0.20	0.14	0.15	0.16	0.14	0.13	0.15	0.15	0.15	0.16
	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
ME	30560	25381	18595	24034	34924	29948	45366	43124	41734	36526
	(2183.3)	(1574.4)	(512.63)	(1230.1)	(1734.1)	(1548.6)	(2511.8)	(2134.9)	(1811.1)	(1341.9)
DUR_10	15.35	15.12	15.32	15.41	15.60	15.89	16.25	16.64	16.86	17.17
	(0.096)	(0.065)	(0.060)	(0.066)	(0.058)	(0.057)	(0.053)	(0.045)	(0.041)	(0.035)
Dur_resid	0.05	-0.08	-0.05	-0.09	-0.01	-0.02	0.01	0.03	0.00	0.05
	(0.022)	(0.023)	(0.023)	(0.021)	(0.021)	(0.016)	(0.016)	(0.020)	(0.019)	(0.021)

Table 2: Correlations

Panel A							
variable	DUR_10	Dur_resid	BM	GPA	logME	ffOP	ROE
BM	-0.34	-0.00	1.00	-0.17	-0.56	-0.23	-0.23
	(-64.71)	(-0.06)	(.)	(-40.66)	(-147.9)	(-54.00)	(-51.02)
DUR_10	1.00	0.54	-0.34	-0.12	0.17	-0.17	-0.37
	(.)	(165.11)	(-65.20)	(-25.66)	(32.23)	(-34.83)	(-76.31)
Dur_resid	0.54	1.00	-0.00	-0.00	-0.09	-0.14	-0.33
	(165.68)	(.)	(-0.21)	(-1.34)	(-29.46)	(-40.51)	(-87.44)
GPA	-0.12	-0.00	-0.17	1.00	0.05	0.53	0.37
	(-25.64)	(-1.28)	(-40.64)	(.)	(10.56)	(128.49)	(82.57)
ROE	-0.37	-0.33	-0.23	0.37	0.22	0.66	1.00
	(-76.36)	(-87.37)	(-51.34)	(82.67)	(47.38)	(173.91)	(.)
ffOP	-0.17	-0.14	-0.24	0.53	0.20	1.00	0.66
	(-34.68)	(-40.48)	(-54.17)	(129.31)	(43.84)	(.)	(173.55)
logME	0.16	-0.09	-0.56	0.05	1.00	0.20	0.22
	(32.05)	(-29.35)	(-147.9)	(10.64)	(.)	(43.88)	(47.15)

Panel B							
variable	DUR_10	Dur_resid	BM	logME	GPA	ffOP	ROE
BM	-0.44	0.02	1.00	-0.31	-0.13	-0.12	-0.12
	(-14.61)	(1.44)	(.)	(-20.59)	(-15.89)	(-4.50)	(-5.00)
DUR_10	1.00	0.53	-0.44	0.03	-0.08	-0.26	-0.39
	(.)	(25.42)	(-14.61)	(1.23)	(-4.35)	(-7.25)	(-8.61)
Dur_resid	0.53	1.00	0.02	-0.02	-0.03	-0.12	-0.28
	(25.42)	(.)	(1.44)	(-2.06)	(-2.49)	(-12.16)	(-21.12)
GPA	-0.08	-0.03	-0.13	-0.02	1.00	0.35	0.28
	(-4.35)	(-2.49)	(-15.89)	(-2.39)	(.)	(34.44)	(30.13)
ROE	-0.39	-0.28	-0.12	0.25	0.28	0.71	1.00
	(-8.61)	(-21.12)	(-5.00)	(18.18)	(30.13)	(42.56)	(.)
ffOP	-0.26	-0.12	-0.12	0.27	0.35	1.00	0.71
	(-7.25)	(-12.16)	(-4.50)	(16.14)	(34.44)	(.)	(42.56)
logME	0.03	-0.02	-0.31	1.00	-0.02	0.27	0.25
	(1.23)	(-2.06)	(-20.59)	(.)	(-2.39)	(16.14)	(18.18)

Table 3: Relation between discount rate risk and profitability.

Panel A: BM											
variable	Low	2	3	4	5	6	7	8	9	High	High-Low
Eret	0.38	0.54	0.57	0.53	0.55	0.60	0.69	0.64	0.66	0.89	0.51
	(0.218)	(0.205)	(0.202)	(0.207)	(0.196)	(0.193)	(0.193)	(0.189)	(0.202)	(0.231)	(0.196)
CFbeta	0.10	0.11	0.16	0.15	0.14	0.12	0.16	0.15	0.15	0.19	0.09
	(0.042)	(0.037)	(0.037)	(0.038)	(0.038)	(0.035)	(0.038)	(0.036)	(0.039)	(0.044)	(0.033)
DRbeta	0.99	0.96	0.88	0.91	0.83	0.80	0.76	0.77	0.79	0.87	-0.12
	(0.075)	(0.064)	(0.053)	(0.071)	(0.067)	(0.061)	(0.072)	(0.075)	(0.076)	(0.089)	(0.078)
CFbeta1	0.46	0.78	1.04	1.30	1.39	1.52	1.49	1.56	1.63	1.20	0.74
	(0.030)	(0.048)	(0.079)	(0.112)	(0.069)	(0.062)	(0.059)	(0.112)	(0.241)	(0.182)	(0.201)
DRbeta1	0.95	0.98	1.01	0.80	0.86	0.77	0.68	0.81	0.80	0.63	-0.32
	(0.060)	(0.066)	(0.061)	(0.068)	(0.092)	(0.084)	(0.101)	(0.146)	(0.104)	(0.162)	(0.176)

Panel B: GPA											
variable	Low	2	3	4	5	6	7	8	9	High	High-Low
Eret	0.34	0.42	0.46	0.42	0.58	0.54	0.50	0.47	0.56	0.66	0.32
	(0.199)	(0.188)	(0.205)	(0.208)	(0.206)	(0.209)	(0.221)	(0.214)	(0.202)	(0.201)	(0.142)
CFbeta	0.14	0.11	0.13	0.15	0.15	0.16	0.13	0.14	0.09	0.09	-0.05
	(0.038)	(0.034)	(0.037)	(0.041)	(0.038)	(0.042)	(0.043)	(0.040)	(0.037)	(0.038)	(0.025)
DRbeta	0.81	0.77	0.89	0.89	0.86	0.90	0.99	0.93	0.89	0.89	0.09
	(0.065)	(0.069)	(0.082)	(0.073)	(0.074)	(0.058)	(0.067)	(0.066)	(0.060)	(0.062)	(0.052)
CFbeta1	1.96	1.42	1.36	0.98	0.98	1.02	0.94	0.68	0.66	0.50	-1.46
	(0.099)	(0.090)	(0.076)	(0.093)	(0.047)	(0.058)	(0.043)	(0.036)	(0.031)	(0.038)	(0.090)
DRbeta1	0.80	0.97	0.95	0.94	1.03	0.89	0.99	1.01	0.95	1.00	0.20
	(0.076)	(0.139)	(0.089)	(0.058)	(0.066)	(0.072)	(0.078)	(0.050)	(0.067)	(0.066)	(0.089)

Table 4: Double sorting by duration and anomalies (VW, orthogonalized)

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: BM

Duration	Low	2	3	4	High	5-1
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.51	0.61	0.63	0.66	0.62	0.11
	(0.25)	(0.21)	(0.22)	(0.20)	(0.22)	(0.21)
2	0.49	0.51	0.56	0.72	0.93	0.43
	(0.20)	(0.20)	(0.19)	(0.18)	(0.20)	(0.18)
3	0.20	0.47	0.65	0.57	0.86	0.66
	(0.24)	(0.23)	(0.21)	(0.22)	(0.25)	(0.20)
long-short	-0.31	-0.13	0.02	-0.09	0.24	0.55
	(0.15)	(0.13)	(0.13)	(0.12)	(0.15)	(0.21)

Panel B: GPA

Duration	Low	2	3	4	High	1-5
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.43	0.51	0.60	0.66	0.85	0.42
	(0.20)	(0.23)	(0.23)	(0.24)	(0.24)	(0.17)
2	0.42	0.49	0.52	0.55	0.65	0.23
	(0.19)	(0.20)	(0.20)	(0.21)	(0.19)	(0.14)
3	0.34	0.30	0.51	0.34	0.37	0.03
	(0.23)	(0.22)	(0.23)	(0.24)	(0.24)	(0.18)
long-short	-0.09	-0.21	-0.08	-0.32	-0.47	-0.38
	(0.13)	(0.13)	(0.14)	(0.14)	(0.16)	(0.20)

Table 5: Double sorting by duration and anomalies (VW, unorthogonalized)

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: BM						
Duration	Low	2	3	4	High	5-1
rank3DUR	_1	_2	_3	_4	_5	_6
1	0.59	0.77	0.68	0.77	0.76	0.16
	(0.38)	(0.27)	(0.23)	(0.20)	(0.20)	(0.31)
2	0.67	0.58	0.55	0.61	0.58	-0.10
	(0.24)	(0.20)	(0.19)	(0.18)	(0.24)	(0.20)
3	0.44	0.49	0.50	0.47	0.54	0.10
	(0.21)	(0.22)	(0.24)	(0.31)	(0.33)	(0.25)
long-short	-0.15	-0.28	-0.17	-0.30	-0.21	-0.06
	(0.27)	(0.18)	(0.15)	(0.20)	(0.23)	(0.40)

Panel B: GPA						
Duration	Low	2	3	4	High	1-5
rank3DUR	_1	_2	_3	_4	_5	_6
1	0.62	0.69	0.88	0.86	1.00	0.38
	(0.19)	(0.22)	(0.22)	(0.25)	(0.26)	(0.20)
2	0.32	0.55	0.61	0.62	0.93	0.61
	(0.18)	(0.20)	(0.20)	(0.21)	(0.21)	(0.16)
3	0.30	0.21	0.42	0.46	0.54	0.24
	(0.26)	(0.24)	(0.23)	(0.22)	(0.20)	(0.16)
long-short	-0.32	-0.48	-0.47	-0.40	-0.47	-0.15
	(0.19)	(0.15)	(0.16)	(0.16)	(0.18)	(0.26)

Table 6: Double sorting by duration and anomalies (BM, VAR)

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: CF beta

Duration	Low	2	3	4	High	5-1
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.12	0.16	0.17	0.19	0.19	0.07
	(0.047)	(0.039)	(0.042)	(0.040)	(0.040)	(0.038)
2	0.11	0.15	0.15	0.16	0.15	0.04
	(0.036)	(0.038)	(0.035)	(0.036)	(0.038)	(0.032)
3	0.13	0.15	0.12	0.17	0.20	0.07
	(0.043)	(0.042)	(0.039)	(0.040)	(0.045)	(0.034)
long-short	0.01	-0.01	-0.05	-0.01	0.01	0.00
	(0.029)	(0.022)	(0.025)	(0.020)	(0.028)	(0.039)

Panel B: DR beta

Duration	Low	2	3	4	High	1-5
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	1.11	0.92	0.92	0.79	0.84	-0.27
	(0.086)	(0.071)	(0.076)	(0.074)	(0.077)	(0.075)
2	0.92	0.87	0.75	0.69	0.72	-0.20
	(0.061)	(0.057)	(0.057)	(0.062)	(0.082)	(0.071)
3	1.07	1.03	0.84	0.87	0.92	-0.15
	(0.077)	(0.081)	(0.069)	(0.089)	(0.089)	(0.085)
long-short	-0.04	0.12	-0.09	0.08	0.07	0.12
	(0.053)	(0.044)	(0.056)	(0.059)	(0.057)	(0.079)

Table 7: Double sorting by duration and anomalies (BM, CF)

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: CF beta						
Duration	Low	2	3	4	High	5-1
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.66	0.93	1.08	1.10	1.19	0.53
	(0.048)	(0.049)	(0.077)	(0.052)	(0.095)	(0.118)
2	0.63	1.02	1.29	1.33	1.38	0.67
	(0.034)	(0.040)	(0.060)	(0.109)	(0.132)	(0.108)
3	0.48	1.01	1.39	1.62	1.63	0.75
	(0.045)	(0.110)	(0.126)	(0.074)	(0.196)	(0.148)
long-short	-0.13	-0.01	0.14	0.18	-0.04	0.22
	(0.107)	(0.062)	(0.112)	(0.125)	(0.150)	(0.155)

Panel B: DR beta						
Duration	Low	2	3	4	High	1-5
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.98	1.04	0.88	0.78	0.70	-0.28
	(0.081)	(0.130)	(0.099)	(0.089)	(0.128)	(0.182)
2	0.94	0.89	0.84	0.83	0.66	-0.27
	(0.072)	(0.067)	(0.077)	(0.170)	(0.124)	(0.121)
3	1.07	1.16	0.73	0.67	0.63	-0.38
	(0.072)	(0.132)	(0.109)	(0.097)	(0.182)	(0.139)
long-short	-0.13	-0.14	-0.31	-0.26	-0.23	-0.10
	(0.107)	(0.159)	(0.116)	(0.084)	(0.188)	(0.253)

Table 8: Double sorting by duration and anomalies (GPA, VAR)

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: CF beta

Duration	Low	2	3	4	High	5-1
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.16	0.19	0.19	0.13	0.11	-0.04
	(0.037)	(0.046)	(0.043)	(0.043)	(0.042)	(0.031)
2	0.13	0.14	0.15	0.15	0.11	-0.02
	(0.038)	(0.038)	(0.039)	(0.043)	(0.038)	(0.025)
3	0.12	0.14	0.17	0.14	0.11	-0.01
	(0.041)	(0.044)	(0.044)	(0.044)	(0.043)	(0.028)
long-short	-0.03	-0.05	-0.02	0.01	-0.00	0.03
	(0.024)	(0.024)	(0.028)	(0.028)	(0.028)	(0.035)

Panel B: DR beta

Duration	Low	2	3	4	High	1-5
rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.80	0.98	0.98	1.05	1.00	0.20
	(0.067)	(0.084)	(0.082)	(0.085)	(0.087)	(0.055)
2	0.73	0.85	0.84	0.90	0.84	0.11
	(0.070)	(0.071)	(0.058)	(0.060)	(0.061)	(0.058)
3	0.95	0.90	0.95	1.05	1.01	0.07
	(0.088)	(0.083)	(0.064)	(0.078)	(0.078)	(0.069)
long-short	0.13	-0.08	-0.01	0.01	0.01	-0.13
	(0.061)	(0.042)	(0.052)	(0.054)	(0.057)	(0.081)

Table 9: Double sorting by duration and anomalies (GPA, CF)

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: CF beta

rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	1.26	1.05	0.91	1.00	0.76	-0.50
	(0.111)	(0.070)	(0.045)	(0.056)	(0.051)	(0.130)
2	1.51	1.08	0.98	0.90	0.69	-0.74
	(0.156)	(0.054)	(0.043)	(0.030)	(0.035)	(0.140)
3	1.88	1.16	0.95	0.68	0.48	-1.06
	(0.094)	(0.105)	(0.066)	(0.044)	(0.025)	(0.107)
long-short	0.18	-0.08	-0.11	-0.47	-0.38	-0.56
	(0.120)	(0.116)	(0.095)	(0.069)	(0.051)	(0.140)

Panel B: DR beta

rank3Dur_resid	_1	_2	_3	_4	_5	_6
1	0.67	0.86	1.03	1.10	0.94	0.27
	(0.105)	(0.075)	(0.117)	(0.123)	(0.099)	(0.183)
2	1.05	0.96	0.88	0.94	0.94	-0.01
	(0.142)	(0.049)	(0.055)	(0.069)	(0.075)	(0.127)
3	0.77	1.09	1.04	1.00	0.99	0.13
	(0.096)	(0.113)	(0.103)	(0.089)	(0.084)	(0.085)
long-short	-0.00	-0.09	-0.25	-0.36	-0.14	-0.14
	(0.119)	(0.109)	(0.170)	(0.159)	(0.116)	(0.211)

Figure 3: Performance of the model

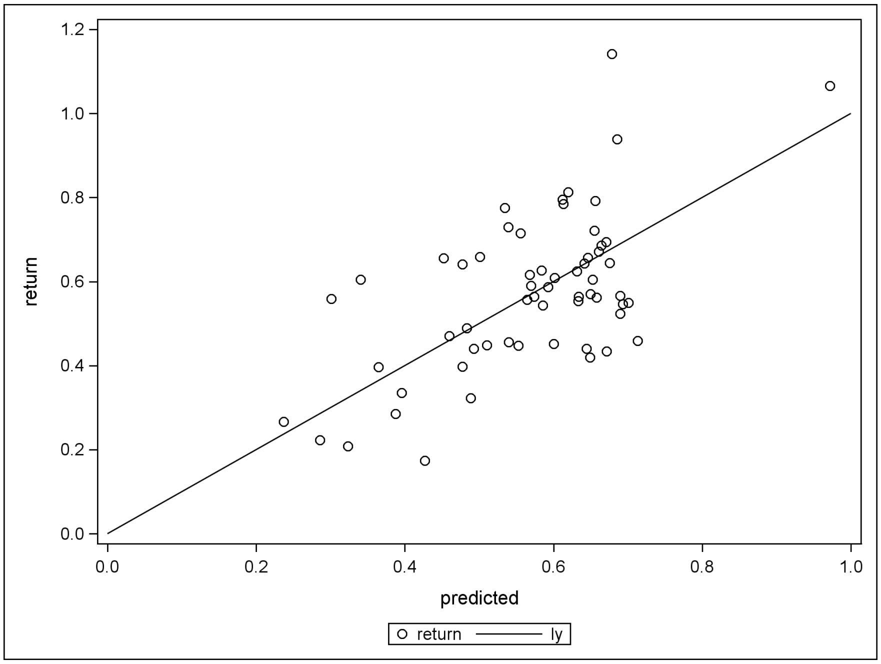


Table 10: Price of risks

Returns of long minus short factors are regressed on Fama and French five factors.

rank3Dur_resid	CFlamda	DRlamda	_RSQ_
(1)	(2)	(3)	(4)
1	0.90	0.51	0.96
	(0.671)	(0.117)	
2	3.24	0.17	0.94
	(1.340)	(0.235)	
3	4.82	-0.26	0.92
	(0.974)	(0.154)	

Table 11: Double sorting by BM and GPA

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: fraction					
rank5GPA	_1	_2	_3	_4	_5
1	0.20	0.12	0.16	0.21	0.31
2	0.15	0.15	0.17	0.22	0.30
3	0.18	0.19	0.22	0.19	0.24
4	0.24	0.22	0.19	0.16	0.20
5	0.37	0.21	0.15	0.12	0.15

Panel B: size					
rank5GPA	_1	_2	_3	_4	_5
1	0.65	1.22	1.25	1.24	0.81
2	1.19	1.50	1.32	1.10	0.46
3	1.60	1.45	1.27	0.80	0.22
4	2.28	1.31	0.60	0.25	0.11
5	2.16	0.65	0.28	0.14	0.06

Table 12: Double sorting by BM and GPA

This table reports excess return for the 5x5 portfolios sorted by duration and anomalies including book to market ratio(BM) in panel A, size (ME) in panel B, operating profit (OP) in panel C, and investment (Inv) in panel D. A return spread between the two extreme portfolios is reported at the last row and last column in each panel. Newey-West t-ratios, computed with N lags, are reported in parentheses. T-ratios marked with ***, **, * indicate that the spreads in betas are statistically significant at the 1%, 5%, and 10% levels, respectively.

Panel A: fraction					
rank5BM	_1	_2	_3	_4	_5
1	0.14	0.11	0.14	0.22	0.40
2	0.12	0.14	0.19	0.26	0.29
3	0.16	0.16	0.23	0.24	0.22
4	0.22	0.21	0.21	0.20	0.18
5	0.25	0.22	0.19	0.19	0.16

Panel B: size					
rank5BM	_1	_2	_3	_4	_5
1	0.27	0.71	0.86	1.25	1.27
2	0.87	1.44	1.28	1.09	0.54
3	1.18	1.67	1.43	0.63	0.30
4	1.55	1.69	1.07	0.34	0.19
5	2.08	1.32	0.58	0.28	0.20

Table 13: Correlations

Panel A				
NAME	MktRF	valuePrem	profPrem	Eret
MktRF	1	-0.162400631	0.057063405	0.903915486
valuePrem	-0.162400631	1	-0.627125158	-0.421429859
profPrem	0.057063405	-0.627125158	1	0.404170655
Eret	0.903915486	-0.421429859	0.404170655	1

Table 14: Construction of Fama-French factors

Firms are first sorted independently by duration into 2 groups. For each duration group, I follow Fama French(2015) to construct SMB, HML, RMW, and CMA factors. Column (2) and (3) shows the mean factor return of the short and long duration group; column (4) shows the mean difference between the (2) and (3); column (5) and (6) shows the standard deviation of factor return of short and long duration group column (7) shows the mean difference between (5) and (6). column (8) and (9) shows the mean and standard deviation of Fama and French factor for comparison. Bootstrapped standard errors of each of the moments are in parentheses.

Panel B: Value Weight								
factors	short	long	long - short	short STD	long STD	longSTD-shortSTD	FF5	FF5 STD
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
BM	0.069	0.435	0.365	3.275	3.369	0.094	0.345	2.849
	(0.14)	(0.14)	(0.12)	(0.18)	(0.15)	(0.13)	(0.11)	(0.12)
IA	0.196	0.395	0.2	2.416	2.501	0.085	0.310	2.007
	(0.10)	(0.10)	(0.11)	(0.09)	(0.10)	(0.12)	(0.08)	(0.08)
ME	0.118	0.287	0.169	3.347	2.928	-0.419	0.257	3.057
	(0.14)	(0.12)	(0.08)	(0.19)	(0.13)	(0.14)	(0.12)	(0.14)
ffOP	0.427	0.05	-0.377	2.664	2.587	-0.077	0.252	2.114
	(0.11)	(0.11)	(0.10)	(0.17)	(0.17)	(0.11)	(0.08)	(0.15)
gpa	0.34	0.052	-0.288	2.648	2.833	0.186		
	(0.11)	(0.12)	(0.12)	(0.09)	(0.10)	(0.11)		
roe	0.248	-0.235	-0.483	2.841	2.724	-0.117		
	(0.12)	(0.11)	(0.12)	(0.14)	(0.15)	(0.12)		

Table 15: Time series regression

Returns of long minus short factors are regressed on Fama and French five factors.

Dep Var	Intercept	Mktrf	SMB	HML	RMW	CMA	RSQ
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BM	0.375	0.003	0.064	-0.093	-0.107	0.123	0.024
	(2.98)	(0.10)	(1.48)	(-1.57)	(-1.73)	(1.35)	
GPA	-0.243	0.001	-0.037	0.113	-0.032	-0.209	0.012
	(-2.00)	(0.03)	(-0.90)	(1.98)	(-0.55)	(-2.39)	
IA	0.221	-0.053	0.038	-0.213	-0.109	0.300	0.055
	(1.90)	(-1.92)	(0.94)	(-3.89)	(-1.92)	(3.57)	
ME	0.141	-0.003	-0.061	0.254	0.096	-0.212	0.132
	(1.81)	(-0.18)	(-2.27)	(6.95)	(2.52)	(-3.78)	
ROE	-0.412	0.056	0.002	-0.072	-0.103	-0.144	0.054
	(-3.28)	(1.89)	(0.04)	(-1.22)	(-1.68)	(-1.59)	
ffOP	-0.251	-0.060	-0.003	0.053	-0.092	-0.270	0.029
	(-2.29)	(-2.32)	(-0.08)	(1.02)	(-1.72)	(-3.41)	

Figure 4: Sharpe Ratio of RMW vs CMA

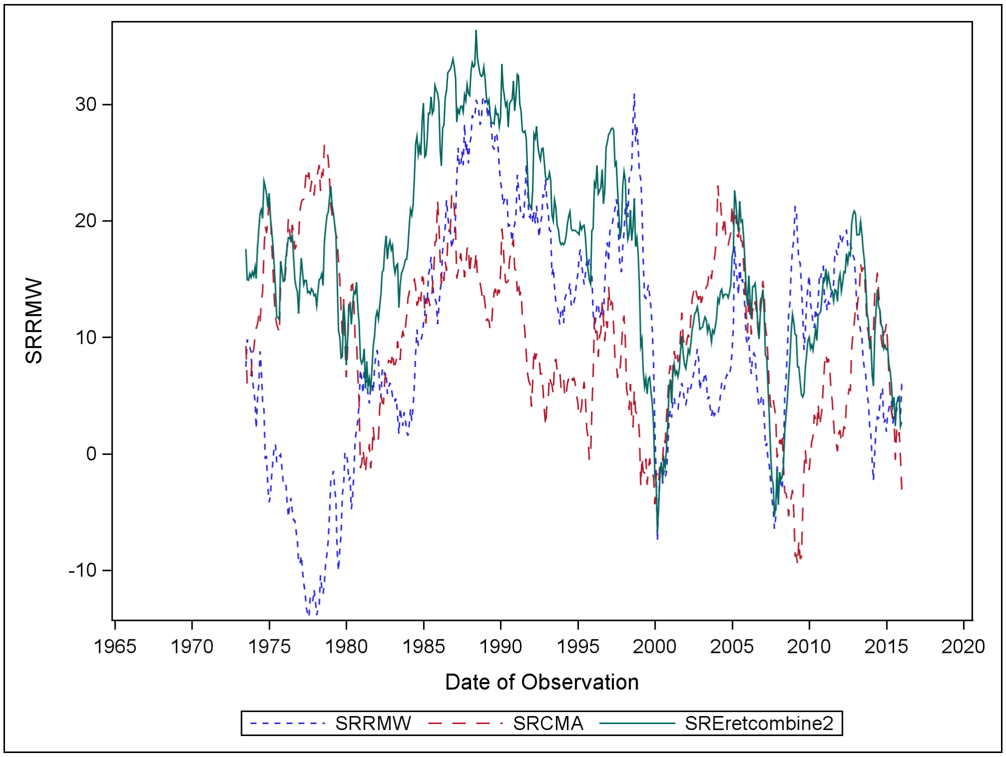
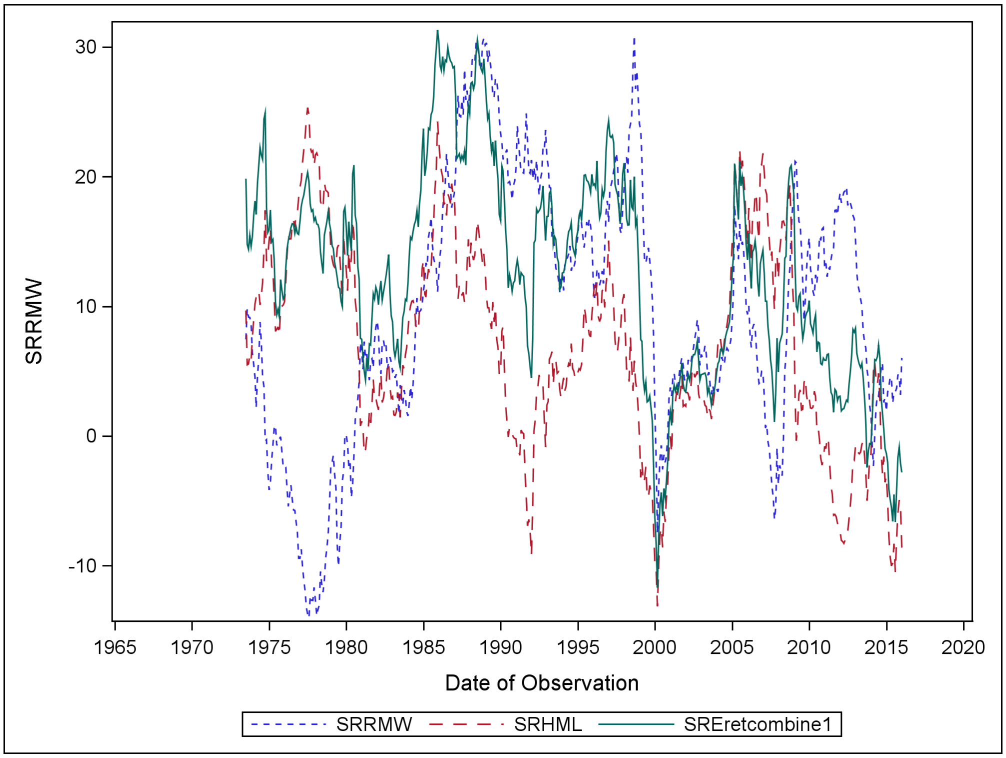


Figure 5: Sharpe Ratio of HML vs CMA



References

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

Beaver, William, Maureen McNichols, and Richard Price, 2007, Delisting returns and their effect on accounting-based market anomalies, *Journal of Accounting and Economics* 43, 341–368.

Campbell, J. Y., C. Polk, and T. Vuolteenaho, 2010, Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns, *Review of Financial Studies* 23, 305–344.

Campbell, John, 1996, Understanding Risk and Return, *The Journal of Political Economy* 104, 298–345.

Campbell, John, and Tuomo Vuolteenaho, 2004, Bad Beta, Good Beta, *THE AMERICAN ECONOMIC REVIEW* 96, 1249–1275.

Campbell, John Y., 1991, A Variance Decomposition for Stock Returns, *The Economic Journal* 101, 157.

Campbell, John Y., 1993, Intertemporal Asset Pricing without Consumption Data, *The American Economic Review* 83, 487–512.

Chen, Huafeng Jason, 2014, Do Cash Flows of Growth Stocks Really Grow Faster?, *Available at SSRN 1903904* .

Cochrane, John H., 2011, Presidential address: Discount rates, *The Journal of Finance* 66, 1047–1108.

Dechow, Patricia M., Richard G. Sloan, and Mark T. Soliman, 2004, Implied equity duration: A new measure of equity risk, *Review of Accounting Studies* 9, 197–228.

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

Fama, Eugene F., and Kenneth R. French, 1996, Multifactor Explanations of Asset Pricing Anomalies, *The Journal of Finance* 51, 55.

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

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

Hou, K., C. Xue, and L. Zhang, 2015, Digesting Anomalies: An Investment Approach, *Review of Financial Studies* 28, 650–705.

Lam, F. Y., Shujing Wang, and Kuo-Chiang John Wei, 2014, The Profitability Premium: Macroeconomic Risks or Expectation Errors?, *Available at SSRN 2479232* .

Lettau, Martin, and Jessica A. Wachter, 2007, Why is long-horizon equity less risky? a duration-based explanation of the value premium, *The Journal of Finance* 62, 55–92.

Lettau, Martin, and Jessica A. Wachter, 2011, The term structures of equity and interest rates, *Journal of Financial Economics* 101, 90–113.

Lin, Xiaoji, and Lu Zhang, 2013, The investment manifesto, *Journal of Monetary Economics* 60, 351–366.

Merton, Robert C., 1973, An Intertemporal Capital Asset Pricing Model, *Econometrica* 41, 867.

Newey, Whitney K., and Kenneth D. West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703.

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

Restoy, F., and P. Weil, 2011, Approximate Equilibrium Asset Prices, *Review of Finance* 15, 1–28.

van Binsbergen, Jules, Wouter Hueskes, Ralph Koijen, and Evert Vrugt, 2013, Equity yields, *Journal of Financial Economics* 110, 503–519.

van Binsbergen, Jules H., and Ralph SJ Koijen, 2015, The term structure of returns: Facts and theory, Technical report, National Bureau of Economic Research.

van Binsbergen, Jules van, Michael Brandt, and Ralph Koijen, 2012, On the Timing and Pricing of Dividends, *American Economic Review* 102, 1596–1618.

Wang, Huijun, and Jianfeng Yu, 2013, Dissecting the profitability premium, in *AFA 2013 San Diego Meetings Paper*.

Weber, Michael, 2016, Cash Flow Duration and the Term Structure of Equity Returns, Technical report, National Bureau of Economic Research.