Profitability Premium: Risk or Mispricing?

Current Draft: November 1^{st} 2015

Ryan Liu*

Abstract

This paper presents evidence that the stock return premium associated with profitability is hard to reconcile with risk-based explanations but is consistent with expectation errors. Firms with lower profitability are more volatile, suffer greater drawdowns and are more sensitive to macroeconomic conditions. This means that the profitable firms are actually less risky by most measures and perform better during economic downturns. In addition, there is a monotonic relationship between profitability and forecast error. Analysts tend to be overoptimistic for low profitability firms relative to high profitability firms. Surprisingly, this mis-expectation can persist even up to five years into the future.

JEL Classifications: E32, G12, G14

Keywords: Profitability Premium, Business Cycle, Analyst Forecast, Expectation Error

*Haas School of Business, University of California at Berkeley. Berkeley, CA, USA. Email: rliu@haas.berkeley.edu

1 Introduction

Previous research has shown that a firm's profitability exhibits strong link with the firm's expected stock return and predicts return as far as ten years ahead. The return variations generated by sorting based on profitability are not subsumed by the size and book-to-market factors. Moreover, a factor constructed based on profitability has been shown to explain a wide range of asset pricing "anomalies"[1]. This paper studies the behavior of the premium associated with profitability and tries to understand what is driving this premium.

Researchers have traditionally divided into two camps regarding why some groups of firms yield superior returns relative to other firms. Profitable firms' stock might perform better because it is fundamentally riskier, and therefore investors demand a higher risk premium for holding these stocks. Alternatively, the profitable firms might have higher return because these firms are mispriced due to incorrect expectation of the market. The mis-expectation might result from a myriad of cognitive biases that investors are subjected to. For example, they might extrapolate past performance or earnings too far into the future. They might overreact to good or bad news about and therefore tend to overbuy those that have done well in the past (glamour stocks) and oversell those that have done poorly. These two potential explanations have different implications for both academic research and industry applications. If the premium is truly driven by fundamental risk, then profitability should be incorporated into a benchmark model of return. This benchmark model can be used to better discount cash flows in capital budgeting applications and to test for alphas of new subsequent strategies and proposed factors. It can also be used to evaluate portfolio returns and fund performance. If the premium is just mispricing, market participants will eventually catch on to it. This implies that the premium will disappear in the future as more investors begin trading based on this idea.

The analysis in the paper is divided into two parts. In the first part, I investigate whether the profitability premium is due to systematic risk. I examine various measures that are traditionally associated with risk across ten portfolios sorted based on profitability. The profitable firms have quite consistently outperformed the unprofitable firms. They also tend to have lower return volatility and this leads to a higher Sharpe ratio. Moreover, these firms have lower drawdown in times of distress. Next, I decompose the beta of the portfolios with respect to the market into a component associated with cash flow news and a component

[1] See Novy-Marx (2013) and Ball et al. (2015)

associated with discount rate news. Campbell and Vuolteenaho (2004) argue that cash flow news only changes the wealth of the investors without affecting the investment opportunity set, while discount rate news affects both the wealth as well as the investment opportunity set. In an intertemporal asset pricing model, the cash flow news would have a higher price of risk since its effect is more permanent while the discount rate news provides a hedge component because of changing investment opportunity. A risk-averse, long-term investors would demand a higher premium for holding assets that covary with a market's cash flow news than assets that covary with a market's discount rate news. Therefore a firm's risk will not just depend on its overall market beta, but the composition of the beta with respect to cash flow (co-movement with market cash flow news) and discount rate beta (co-movement with the market discount rate news). Profitable firms might be riskier if they have higher covariations with the market's cash flow news, that is, higher cash flow beta. This is plausible since profitability is essentially a measure of how much money the firm can generate and how consistently they can generate that money. However, I find that unprofitable firms actually have higher cash flow beta relative to the profitable firms.

Economic theory suggests another way that the profitable firms can be riskier. Investors care the most about returns during bad times when their marginal utility of wealth is high. While the profitable firms have higher unconditional returns, investors might still avoid them if their returns are the lowest exactly during bad times. Stated in another way, the unprofitable firms might yield lower return if they do well in bad times and therefore provides a hedge to the investors. To study this possibility, I look at the cyclical variation of the premium across the business cycle and how the premium relate to common macroeconomic predictors of return. I find that the premium is actually higher during economic downturns. Thus the profitable firms do even better than the unprofitable firms during bad times when the marginal utility is the highest. To better capture the asymmetry in return behavior of profitable and unprofitable firms across the business cycle, I further adopt a two-state Markov-switching model that allows the conditional distribution of returns to vary with the state of the economy. This flexible econometric framework gives additional insights on the mechanisms generating the difference in expected returns. The expected profitability premium exhibit clear cyclical variation with a tendency to spike up in the beginning of recessions. Across the states, the unprofitable firms display the highest degree of asymmetry in their sensitivity to aggregate economic conditions.

In the second part of the paper, I investigate the mispricing hypothesis by examining the expectation error of the market participants. Specifically, I look at the difference between the earnings that are forecasted by sell-side analysts and the actual earnings realized across profitability-sorted portfolios. If the low return of the unprofitable firms relative to the profitable firms is due to investors being too optimistic about its future performance, I should find that the difference between forecast and actual earnings (expectation error) is larger for unprofitable firms than profitable firms. The results show that there is indeed a monotonically decreasing relationship across the ten deciles of profitability from low to high. The expectation error is larger for the unprofitable firms. Interestingly, this expectation error is quite persistent and last up to five years into the future. I also analyze potential reaons leading to the expectation error. Unlike the typical behavioral explanation given for the value premium in which the growth firms tend to be "glamorous" stocks that have done well in the past and the investors naively expect it to continue doing well, the unprofitable firms have actually done poorly over the few year prior to portfolio formation. Rather it is the profitable firms are the ones that have done well in the past, and they tend to continue perform well in the one year after portfolio formation. They also tend to be bigger firms with higher earnings yield, higher asset growth, lower stock issuance and lower distress. Overall, the investors seems to expect the performance of the profitable firms to mean-revert quicker than they actually do, and are willing to bet on the revival of the unprofitable firms despite low net income and poor current performance.

My paper relates to the literature on the role of profitability in determining stock return. Early papers such as Ball and Brown (1968) focus on accounting earnings and show that net income predicts the cross section of returns. However, later research by Fama and French (1992) and Fama and French (1996) show that information contained in earnings are completely subsumed by size and book-to-market. More recently, there has been a revival of interest in the link between profitability and stock return that relies on different measures. Novy-Marx (2013) argues that gross profitability (GP), defined as revenue minus cost of goods sold divided by book value of total asset, provides a cleaner measure of true economic profitability than earnings. This is because accounting earnings also captures investment expenses such as research and development (R&D), that actually helps increase future economic profit. A firm that is spending a high amount on R&D might have high future profitability, but low current earnings. He shows evidence that GP provides incremental information above size and book-to-market and works especially well when combined with the value signal. His results have attracted considerable attention both from academia and industry. Some asset managers have reported to change their investment strategies based on results. More importantly from an academic's perspective, Novy-Marx shows that a GP factor can explain a large set of asset pricing anomalies. Building on Novy-Marx's work, Ball et al. (2015) examine profitability closer and construct an alternative measure, operating profitability, that exhibits an even stronger link with firm performance than gross profit. Given the robust explanatory power of profitability beyond size and value and its relation with the other asset pricing anomalies, Fama and French (2014) try to incorporate profitability as an additional factor that summarizes the heterogeneity in stock returns. In a similar line of research, Hou et al. (2014), motivated by the q-theory of investment, propose an alternative factor model for cross section of expected returns consisting of a market factor, a size factor, an investment factor, and a profitability factor and shows that it outperforms the Fama-French 3-Factor model. My paper also fit into the large literature in using behavioral biases to explaining the cross section of stock returns. Early contributions such as Lakonishok et al. (1994) and La Porta (1996) focus on the value premium. A survey of this line of study can be found in Barberis and Thaler (2003). More recently, Engelberg et al. (2015) aggregated 97 stock return anomalies and investigate the earnings announcement period returns and analyst forecast errors. They found that the returns are much higher around announcement periods. In addition, the anomaly signals predict the forecast error. Their work suggests that some kind of expectation error might be driving these anomalies. However since all the predictors are aggregated, the results are in a black box and it is hard to distinguish exactly what anomalies are driving the results, or the relative importance of expectation error across the anomalies. My paper is also closely related to Wang and Yu (2013). They explore various form of conditional portfolio sorting and show that the premium exists primarily among firms with high uncertainty and information cost. They also argue that since the there is little evidence of long-run reversals, the premium is likely to be driven by underreaction. My paper differs from theirs along several dimensions. My main measure of profitability is different from theirs and have been shown to provide a stronger link with expected returns. My econometric framework is also more flexible and allows me to better capture non-linear effects. Most importantly, I directly show the expectation error across firms with different profitability and argues that the lack of long-run reversal is driven by the fact that the investors have not been paying attention to these newer measures and therefore they have not been correcting their mistakes (yet).

The rest of the paper proceeds as follows. Section 2 describes the data source as well as construction of the profitability measure. Section 3 investigates whether the profitability premium is consistent with systematic risk by examining a set of relevant variables. It also studies the cyclical behavior of the premium using the flexible Markov-switching model. Section 4 examines the mispricing story by relating the expected earnings to actual earnings. It also looks at firm characteristics to disentangle the reason behind the expectation error. Section 5 concludes.

2 Data and Profitability Measure

The stock return data is obtained from the Center for Research in Security Prices (CRSP). I only include ordinary common shares traded on NYSE, Amex, and NASDAQ. I combine that with firm level financial information from the Standard and Poor's Compustat database. The portfolios are constructed based on quintile and decile sort using the New York Stock Exchange (NYSE) breakpoints. As is standard in the literature, the portfolios are value-weighted and rebalanced each year at the end of June. The accounting data for all fiscal years ending in calendar year $t-1$ are matched to the returns for the period from July of year t to the June of year $t+1$. The sample consists of firms that have non-missing market value of equity and book-to-market. Following most of the papers in the literature, I exclude the financial firms that have a one-digit standard industrial classification code of six.

Profitability is meant to measure the productiveness of the firm's capital in generating cash flow. While the conceptual idea of profitability is clear, the literature has used various definitions and items in the financial statements in the past to capture it. Fama and French (2006) motivate their empirical analysis using the clean surplus accounting and define profitability to be net income divided by the book value of equity. Wang and Yu (2013) and Hou et al. (2014) use return-on-equity (ROE) as their main measure of profitability, defined as income before extraordinary items (Compustat item IB) divided by book equity. A recent paper by Novy-Marx (2013) that has received a lot of attention in both academia and industry argues that true economic profitability gets polluted the further down the income statement one goes. Furthermore, scaling the income items by book value of equity risks conflating firm's capital productivity with book-to-market which is related to the cross section fo stock return. He advocates gross profitability, defined as the ratio of revenue minus cost of goods sold over total asset, and shows that portfolios sorted on this measure exhibit significant variations in return and is not explained by the usual Fama-French 3-Factor model.

For the purpose of this study, I will mainly focus on an even more refined measure of profitability as given in Ball et al. (2015)

$$Profitability = \frac{REVT - COGS - (XSGA - XRD)}{AT} \tag{1}$$

where $REVT$ is the total revenue, $COGS$ is the cost of goods sold, $XSGA$ is selling, general & administrative expense, XRD is the R&D expense and AT is total asset. Ball et al. (2015) argue that this measure, which they call operating profitability, better matches current expenses with current revenues and leads to a significantly better predictor of future returns than the gross profitability measure. It differs from GP in that it includes selling, general & administrative expense which Ball et al. (2015) show has similar covariance with future return as the cost of goods sold. The R&D expenses are subtracted away because Compustat defines selling, general & administrative expense as the sum of firms' actual selling, general & administrative expense and their R&D expense. Since R&D expenses not just reduce current profit but also generates future revenues, undoing Compustat's adjustment leads to a cleaner measure of profitability. For robustness, I have also repeated the estimations using gross profitability. For all the subsequent results that follow, using gross profitability will not make a material difference to the conclusions.

I take the analysts' earnings forecast and expected growth rates from the Institutional-Brokers-Estimates-System (IBES) in the WRDS database. While not a perfect proxy for the expectation of market participants, these forecast and the accompanying reports are used by institutional investors in their decision making process and provides a valuable guide to how these investors think. Moreover, the technique used to come to these conclusions are standard among market participants and should be correlated with their unobserved belief. The IBES data aggregates the earnings and growth forecast issued by sell-side stock analysts. For each firm-year, I consider the mean analyst forecast of earnings. Since the earnings forecast are issued multiple times in a year, I only keep the closest forecast just prior to portfolio formation in July of year t and that the forecast end period will fall between July of year t to the June of year $t+1$.

3 Are Profitable Firms Riskier?

Broadly speaking, there are two interpretations to what exactly explains the cross-sectional determinants of stock return such as book-to-market, size, and so on. The first interpretation argues that the additional risk that value and small firms incur drives the additional return that they get on average. The alternative explanation argues that investors do not fully understand the relationship between these financial variables and subsequent future performances, and therefore they form erroneous expectations. The profitability premium is also subject to these two interpretations. In this section, I will analyze whether profitability premium is consistent with a risk-based explanation. There are two ways in which the profitable firms can be riskier than the unprofitable firms. Firstly, the profitable firms have to underperform in some states of the world. Secondly, the states that they underperform in are the "bad" states in which the marginal utility of wealth is high.

3.1 Traditional Risk Analysis

I begin by examining the performance of profitable firms and unprofitable firms over time and ask whether the underperformance is consistent. Figure 1 shows the return of high profitability firms relative to low profitability firms on a year by year basis from July 1963 to December 2013, with the first and last year only accounting for six months of return. The premium is calculated as the difference between the value-weighted return of firms in Decile 10 of profitability minus that of Decile 1.

[Figure 1 about here.]

The plot shows that profitable firms have outperformed unprofitable firms in 37 out of the 51 years in the sample. One possible explanation for such high percentage of outperformance that is consistent with systematic risk is that the premium takes the stairs up but the elevator down. That is, it is subject to a "rare disaster" shock that leads to extremely low returns. In the years of underperformance, the lowest return is -34%. This is lower than the highest return of 41% among the years of outperformance. The average return difference in the 37 years of outperformance is 10.8% while the average return difference in the 14 years that profitable firms underperform is 11.4%. Thus while profitable firms consistently yield higher returns than unprofitable firms, the few instances when they underperform do not lead to significantly higher negative returns.

Next, I examine a set of important factors that are usually connected to high risk premia. The results are shown in Panel A of Table 1. I start off by looking at the returns and volatility of each of the portfolios across a variety of holding periods ranging from one month to five years. Across all holding periods, profitable firms earn higher return than unprofitable firms The portfolio with the lowest profitability has the highest annualized standard deviation of 0.236 while the other portfolios have roughly the same standard deviation of 0.17. Volatility therefore cannot be a driver of the greater risk of the profitable firms under the risk hypothesis. Indeed the firms with lower volatility actually have higher average returns. Not surprisingly, the Sharpe ratio is the lowest at 0.084 for the lowest decile of profitability, and it increases to above 0.4 for the high profitable firms in the top three deciles. The next several rows look at the downside risk of the portfolios in more detail. One possibility for the high return of the profitable firms is that they might have to go through periods of enormous loss, and investors must be given a return premium in order to induce them to hold these stocks. Yet again the results do not seem to go in the direction of this hypothesis. I investigate the maximum drawdown, defined as the return from peak to trough, for holding periods of three months, six months, one year, and two years. In all cases, the low profitable firms tend to suffer significantly higher drawdowns and without exception, the lowest decile of profitability suffers the worst drawdown. For example, the worst three month drawdown for the lowest decile of profitability is -51% while the next highest drawdown is suffered by Decile 3 at -37%. The worst one year drawdown for the lowest profitability decile is -74%. This is more than 20% lower than any of the other portfolios and almost 30% lower than the most profitable decile.

[Table 1 about here.]

Previous research has shown that the profitability premium is not subsumed by the Fama-French 3-Factor model, and therefore it is not explained by CAPM. However, Campbell and Vuolteenaho (2004) have shown that the return on the market portfolio can be decomposed into news about future discount rate and news about future cash flow. In the first case, the wealth decreases but future investment opportunities improve while in the second case, wealth decreases and future investment opportunities remain uncahnged. Therefore these two components should have different effects on the risk premium. Specifically, the beta with respect to the market can be broken down into a discount rate beta and a cash flow beta. Firms with higher cash flow beta relative to discount rate beta would be riskier since their covariation is with cash flow news that does not affect investment opportunity set, even though they might have similar overall CAPM beta.

Panel B of Table 1 shows the relative make up of the betas across the profitability portfolios. I follow the procedure as outlined by Campbell and Vuolteenaho (2004). On average the low profitability firms tend to have higher overall beta than the high profitability firms. In particular, the Decile 1 portfolio has both substantially higher discount rate beta and cash flow beta than any of the other portfolios. Thus CAPM beta cannot explain the low returns of the low profit firms. More importantly, the relative weight of discount rate and cash flow beta cannot explain the variation in returns either. Decile 1 has a discount rate beta of 0.676 and a cash flow beta of 0.792, so the proportion of cash flow beta relative to the total beta is 54%. By comparison, the most profitable Decile 10 has a discount rate beta of 0.585 and a cash flow beta of 0.499 so that cash flow beta only makes up 46% of the total beta. Thus high profitability firms not only have lower total beta, but also lower proportion of cash flow beta. One caveat to the results is that beta decomposition has been shown to be very sensitive to model specification (see Chen and Zhao (2009)). This is because the cash flow news is usually back out as a residual after directly measuring the discount rate new, and therefore captures all kinds of modeling noise. My results do not preclude the possibility that profitable firms indeed have high cash flow beta but the VAR model that I used is misspecified.

3.2 Business Cycle Variations

In this section, I empirically analyze the time variation in the profitability premium and relate it to several well-known macroeconomic variables. I examine the data using the traditional regression approach to investigate whether the correlation between conditioning variables and future stock return exhibit systematic patterns across different profitability portfolios. In addition, I adopt a more flexible econometric model that allows me to capture asymmetric response of profitable and unprofitable firms' stock return to business conditions.

The discussion in the first part of this section suggests that profitable firms are consistently outperforming unprofitable firms with lower volatility and lower drawdown. Of course this does not automatically imply that profitable firms are not riskier. Economic theory suggests that the riskier firms are those that tend to underperform in bad times when the marginal utility of consumption is low, even if they have higher return unconditionally. I start off by analyzing the correlation of the relative return of profitable and unprofitable firms with respect to the aggregate market. Figure 2 plots the 12-month moving average of profitability premium (return of the most profitable decile minus return of the least profitable decile) along with the 12-months moving average of aggregate market return given by the CRSP value-weighted index. The NBER recessions are shaded.

[Figure 2 about here.]

The figure shows a clear negative relationship between the profitability premium and market return. During market downturns prior to the onset of recessions, the profitability premium tend to spike up, driven mainly by the low return of the unprofitable firms. The most notable exception is the 1973-1975 recession with its stagflation exacerbated by the oil crisis. Qualitatively, the relationship becomes more reliable after the 1980s. In the two most recent recessions, the aggregate market loss is accompanied by sharp upward spike in the profitability premium.

To understand the premium and how it changes through the business cycle, Table 2 shows some summary statistics and correlations of the premium related to value, size, and profitability.

[Table 2 about here.]

The size premium is significantly higher in recession than expansions, in agreement with the results in Perez-Quiros and Timmermann (2000). They argue that smaller firms are more susceptible to credit constraints and this constraint is especially likely to bind during bad economic times. Therefore, the expected return on small firms relative to large firms should be higher during recessionary states. The size premium also has a significant 32 percent correlation with the aggregate market return. The profitability premium displays some difference across the recessionary and expansionary states. The premium during recession of 0.69 percent is almost twice the premium during expansion of 0.36 percent. Interestingly, it also quite a significant raw negative correlation with the market of -33.4 percent, on par with the correlation associated with the size premium. Finally, the value premium display less cyclicality with the difference between recession and expansion being only 0.1 percent. The raw correlation with the market of -4.9 percent is also much less compared to size premium and profitability premium. This is consistent with the results in Chen et al. (2008) that the expected value premium is only weakly countercyclical.

Looking now at the correlation of the portfolios' return with aggregate macroeconomic variables in the tradition of Fama and French (1989). Fama and French (1989) explore the expected return on different stock and bond portfolios and their correlations with aggregate conditioning variables by running predictive regressions of the portfolio returns on the conditioning variables for various horizons. Table 3 displays the regressions of the profitability decile portfolios' excess return on a constant, the one-month Treasury bill rate, the default spread, and the dividend yield. The Newey-West standard errors adjusting for overlapping returns are reported in parentheses.

[Table 3 about here.]

The one-month Treasury bill rate I_t is widely used to proxy for the investors' unobserved expectation of future economic activity. Fama (1981) shows that a higher nominal T-Bill rate is indicative of an unobserved negative shock to real economic growth. Moreover, the Federal Reserve routinely lowers short term interest rate to stimulate growth in anticipation of economic downturns, leading many studies such as Fama and Schwert (1977), Campbell (1987), and Whitelaw (1994) to use it as a regressor in stock return predictions. These studies usually find a negative correlation between interest rate and future stock returns. More importantly for our study, interest rate serves as a proxy for firm's cost of debt capital and fluctuates as aggregate credit condition changes.

In agreement with results of long horizon regression studies such as Campbell and Shiller (1988) and Hodrick (1992), the point estimates of the Treasury bill coefficients becomes more negative and the p-value increases as the horizon of returns goes from 1 month to 12 months. This is also a general result that will hold for default spread and dividend yield. The previous studies interpret this as more predictability of stock return at longer horizons. More interestingly, the point estimate of the Treasury bill coefficient increases from the low profitability deciles to the high profitability decile. This suggests that stocks in the low profitability decile is much more sensitive to fluctuations general economic conditions and credit market conditions. Moreover as the horizon increases, the magnitude of the difference in the coefficients also increases. For 1-month return, the point estimate on the Treasury bill is -6.03 for the least profitable portfolio and -2.99 for the most profitable portfolio. Thus the least profitable portfolio's coefficient is about twice that of the most profitable portfolio. This difference increases monotonically with the regression horizon. For 12-month return regressions, point estimate on the least profitable portfolio is -54.47 while the same estimate for the most profitable portfolio is -15.46, for a factor of 3 to 4 times. Thus if we consider returns over longer periods of time, the association between economic growth and return of unprofitable firms becomes even stronger relative to the association between economic growth and return of profitable firms. This provides another perspective to the findings in Novy-Marx (2013) and Ball et al. (2015) that profitability premium is persistent. Part of the persistence might stem from the correlation between economic growth and long-term future return.

The default premium Def_t is given by the difference between the Moody's Aaa yield and Baa yield obtained from the St. Louis Federal Reserve database. Empirical macroeconomic research such as Stock and Watson (1989) has shown that default premium is one of the strongest business cycle forecasters. In stock return forecast, researchers as early as Keim and Stambaugh (1986) has found that default premium is positively correlated with future stock returns. Jagannathan and Wang (1996)'s study on conditional CAPM uses the default spread as the only conditioning variable. During bad times, investors will prefer bonds with more stable payout and lower default risk, thus widening the gap between the yield an investor can earn from a security with more credit risk relative to one with less credit risk. This variable captures the general credit constraintness of the economy. Unprofitable firms more affected by distress risk should be more sensitive to aggregate credit constraintness compared to profitable firms.

The results agree broadly with our intuition. In general, stock return of firms with low profitability is more sensitive to the default spread than the stock return of high profitability. For 1 month returns, the coefficient of lowest profitability portfolio's return on default premium is 0.67, which is about twice the coefficient of the highest profitability portfolio of 0.31. Again, this difference tend to increase with horizon. For the 12-month return regressions, the ratio of the coefficients on default premium between the first decile and the tenth decile is more than four times.

Finally, I include the aggregate dividend yield defined by sum of the previous 12 months' dividends of the value-weighted CRSP index divided by the current level of the index. High dividend yield indicates high discount rate and thus it is meant to capture the time variation in the aggregate risk premium. Most classic studies such as Campbell and Shiller (1988) and Fama and French (1988) include it in modeling expected stock returns. The pattern of the coefficient estimates are qualitatively similar to the Treasury bill and the default spread.

Increasing the horizon in general increases the magnitude and statistical significance of the coefficients. The lowest profitability decile' stock returns are more sensitive to the dividend yield than the stock return of the highest profitability decile by a factor of about two. This ratio also increases as the prediction horizon increase, albeit not as much compared to the coefficients on Treasury bill and default premium.

Overall, the results suggest a clear cyclical fluctuation in the profitability premium. The premium tend to increase during recessions and when the aggregate market is doing poorly. This is because the low profitability firms' stock returns are more sensitive to aggregate risk premium, economic growth and credit market conditions than the stock returns of the profitable firms. Low profitable firms' returns tend to decrease more during economic downturns than that of the high profitable firms. This increases the difference between the two groups of firms, and lead to a upward spike in the profitability premium. Thus the unprofitable firms perform worse in bad economic times when the marginal utility is high, so their low return cannot be explained away as a hedge for recessions. Building on the suggestive evidence in this part, I will next explicitly model the divergent response of the portfolios' stock returns across different states of nature.

3.2.1 Regime-Switching Model of Asymmetries

While the regressions as in the previous section provide insight about the differential response of profitability portfolios to aggregate market conditions, it cannot capture asymmetric response of the stock returns across recession and expansions. This is important for a full investigation of risk-based explanations. We have shown that the unprofitable firms have the same correlation structure as the profitable firms with respect to the macroeconomic variables but are unconditionally more sensitive. This does not automatically rule out the possibility that these unprofitable firms provide a hedge for bad times. For example if its correlation with macro conditions change during economic downturns, it can still be attractive for investors to hold these firms in their portfolios. In this section, I use the nonlinear Markov switching model to investigate the time variations in the expected stock return of profitable and unprofitable firms and their relation to macroeconomic variables across different states of the economy. This model is originally developed by Hamilton (1989) to investigate the fluctuation of macroeconomic variables through the business cycle, and has been used by Hamilton and Lin (1996) to study time variations in stock return and volatility. The model is later further extended by Gray (1996) to include time-varying transition probability. Perez-Quiros and Timmermann (2000) uses the model to investigate the relationship between firm size and cyclical stock return. In a similar paper, Gulen et al. (2011) applies the same framework to study cyclical variations and predictability of the value premium.

I will most closely follow the model of Perez-Quiros and Timmermann (2000) in my investigation. For simplicity there will be two states with the identity of the state determined by data. Let ρ_t denote the portfolio's excess return in period t. Let $\mathbf{X}_{t-1}$ be a vector of conditioning variables. I allow the intercept, slope coefficients and volatility of the excess returns to be a function of the latent state variable given by S_t

$$\rho_t = \beta_{0,S_t} + \boldsymbol{\beta}'_{S_t}\mathbf{X}_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma^2_{S_t}) \tag{2}$$

With two states, the latent variable S_t will take on values 1 and 2. So the coefficients and variances will be either $(\beta_{0,1}, \boldsymbol{\beta}'_1, \sigma^2_1)$ or $(\beta_{0,2}, \boldsymbol{\beta}'_2, \sigma^2_2)$.

The state transition probabilities follows a first-order Markov Chain

$$\begin{aligned}
p_t &= P(S_t = 1|S_{t-1} = 1, \mathbf{Y}_{t-1}) = p(\mathbf{Y}_{t-1}) \\
1 - p_t &= P(S_t = 2|S_{t-1} = 1, \mathbf{Y}_{t-1}) = 1 - p(\mathbf{Y}_{t-1}) \\
q_t &= P(S_t = 2|S_{t-1} = 2, \mathbf{Y}_{t-1}) = q(\mathbf{Y}_{t-1}) \\
1 - q_t &= P(S_t = 1|S_{t-1} = 2, \mathbf{Y}_{t-1}) = 1 - q(\mathbf{Y}_{t-1})
\end{aligned}$$

The conditioning vector Y_{t-1} are known at $t-1$ and affects the state transition probability between time $t-1$ and t.

I follow Filardo (1994) in allowing the state transition probability to be time-varying and dependent on the economic leading indicator. This allows the model to capture the change in investor's information about the transition probabilities. The transition probabilities are as follows

$$\begin{aligned}
p^i_t &= P(S^i_t = 1|S^i_{t-1} = 1, \mathbf{Y}_{t-1}) = \Phi(\gamma^i_1 + \gamma^i_2 \triangle CLI_{t-2}) \\
q^i_t &= P(S^i_t = 2|S^i_{t-1} = 2, \mathbf{Y}_{t-1}) = \Phi(\pi^i_1 + \pi^i_2 \triangle CLI_{t-2})
\end{aligned}$$

The variable $\triangle CLI_{t-2}$ is the year-on-year log-difference in the Composite Leading Indicator obtained from the OECD database, lagged two months to account for delayed data release. The function $\Phi(\cdot)$ is the cumulative density function of a standard normal random variable.

Given the assumptions about the normality of the innovation ε_t, the estimation of the parameters are done via the maximum likelihood method. Let $\boldsymbol{\theta}$ denote the vector of parameters, and Ω_{t-1} denote the information set at time $t-1$ that includes $\mathbf{X}_{t-1}$, ρ_{t-1}, $\mathbf{Y}_{t-1}$ and all of these variables' values prior to time $t-1$. Then the density conditional on being state j is given by

$$f(\rho_t|\Omega_{t-1}, S_t = j; \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma_j}} exp\left(\frac{-(\rho_t - \beta_{0,j} - \boldsymbol{\beta}_j'\mathbf{X}_{t-1})^2}{2\sigma_j}\right) \tag{3}$$

The log-likelihood function is given by

$$\mathcal{L}(\rho_t|\Omega_{t-1}; \boldsymbol{\theta}) = \sum_{t=1}^{T} log(\phi(\rho_t|\Omega_{t-1}; \boldsymbol{\theta})) \tag{4}$$

The density $\phi(\rho_t|\Omega_{t-1}; \theta)$ is calculated by summing the probability-weighted state densities $f(\cdot)$ across the two possible states

$$\phi(\rho_t|\Omega_{t-1}; \boldsymbol{\theta}) = \sum_{j=1}^{2} f(\rho_t|\Omega_{t-1}, S_t = j; \boldsymbol{\theta}) P(S_t = j|\Omega_{t-1}; \boldsymbol{\theta}) \tag{5}$$

where $P(S_t = j|\Omega_{t-1}; \boldsymbol{\theta})$ is the conditional probability of being in state j at time t given the conditioning information at time $t-1$.

The conditional state probabilities can be obtained recursively from the total probability theorem

$$P(S_t = i|\Omega_{t-1}; \boldsymbol{\theta}) = \sum_{j=1}^{2} P(S_t = i|S_{t-1} = j, \Omega_{t-1}; \boldsymbol{\theta}) P(S_{t-1} = j|\Omega_{t-1}; \boldsymbol{\theta}) \tag{6}$$

The conditional state probabilities can be calculated from Baye's Rule

$$P(S_{t-1} = j|\Omega_{t-1}; \boldsymbol{\theta}) = \frac{f(\rho_{t-1}|\mathbf{X}_{t-1}, \mathbf{Y}_{t-1}, \Omega_{t-2}, S_{t-1} = j; \boldsymbol{\theta}) P(S_{t-1} = j|\mathbf{X}_{t-1}, \mathbf{Y}_{t-1}, \Omega_{t-2}; \boldsymbol{\theta})}{\sum_{j=1}^{2} f(\rho_{t-1}|\mathbf{X}_{t-1}, \mathbf{Y}_{t-1}, \Omega_{t-2}, S_{t-1} = j; \boldsymbol{\theta}) P(S_{t-1} = j|\mathbf{X}_{t-1}, \mathbf{Y}_{t-1}, \Omega_{t-2}; \boldsymbol{\theta})} \tag{7}$$

The parameters can be estimated via a recursive iteration between the two conditional state probabilities. The specific implementation is done using the Markov Regime Switching Models package in Matlab written by Perlin (2014), with an augmentation by Ding (2012) that allow for time varying transition probabilities.

For each profitability decile indexed by i, we estimate a separate model with the excess return of each of the portfolios being dependent on an intercept term, lagged values of the one-month T-Bill rate, a default premium, and the aggregate dividend yield. This specification include the common regressors in the literature on stock predictability. The equation that we estimate is thus

$$\rho_t^i = \beta_{0,S_t}^i + \beta_{1,S_t}^i I_{t-1} + \beta_{2,S_t}^i Def_{t-1} + \beta_{3,S_t}^i Div_{t-1} + \varepsilon_t^i \tag{8}$$

where $\varepsilon_t \sim N(0, \sigma_{S_t}^2)$.

The one-month Treasury bill rate I_t, the default premium Def_t and the aggregate dividend yield Div_t are common predictors in stock return regressions and are all described in Section 3.2. One thing to emphasize is that unlike in Section 3.2, the goal here is to understand the asymmetric response of each of the portfolios return across economic expansions and downturns. That is, we want to study any differences in loading across recessionary and expansionary for profitable and unprofitable firms, both in sign and in magnitude.

The coefficients estimated by fitting the Markov switching model for each of the 10 profitability-sorted portfolios are reported in in Table 4. The table also reports the standard errors of the estimates, estimates of the variance in each states and the overall log likelihood of fit for the model. From the variance estimates, it is clear that State 1 is the low volatility state and State 2 is the high volatility state. It is tempting to interpret this as saying that State 1 correspond to economic expansion and State 2 correspond to recession. Before we can make this conclusion, it helps first to understand the states of the economy that the models identified through the latent approach. Figure 3 plots the model-estimated conditional probability of being in the high volatility state given the information in the previous period. The top plot shows the model-fitted probability estimated from the highest decile of profitability while the bottom plot shows the same probability estimated from the lowest decile of profitability. The recessions identified by the NBER are shaded.

[Figure 3 about here.]

The estimated probabilities are similar regardless of whether we use Decile 1 or Decile 10 of the profitability portfolio. The figure shows that the high volatility state picks up both the recessionary states as well as times when equity market volatility is really high. It also underlies the danger of hastily concluding that high volatility state is the same as recessionary state. There are several periods of high stock volatility that does not lead to recession. The two most obvious examples are the October 1987 stock crash, and the Asian Financial Crisis of 1997. The model identifies those periods as being in the high volatility state. Similarly, the recession of the early 1980s are not recognized as being a state of high volatility. Overall however, the state probabilities do display a strong relationship with the business cycle. Thus from now, I will interchangeably refer to the high volatility State 2 as recessionary states, and the low volatility State 1 as expansionary states.

[Table 4 about here.]

Keeping in mind the caveats regarding the states identified by the model, we can now examine the estimated coefficients from the model. Focus first on the volatility estimates. These parameters have very small standard errors, suggesting that they are quite precisely estimated. As we have seen earlier, the high volatitlity states are usually also the recessionary states, consistent with evidence given in Schwert (1989). More interestingly, the coefficients show that the return volatility decreases in profitability, and this effect is stronger during the high volatility states. During expansions, the volatility estimate for the low profitability portfolio is 0.0022 and the volatility estimate of the high profitability portfolio is 0.0015. The difference is only 0.0007. During recessions, however, the difference between low profitability portfolio of 0.021 and high profitability portfolio of 0.006 is 0.015. This is more than 20 times the difference during expansions. Thus the volatility of the unprofitable firms are more strongly influenced by aggregate economic condition than that of the profitable firms.

Consider now the mean equation. Since the recessionary states are much less common than expansionary states, these estimates are much less precise and none of the estimates for recessionary states are statistically significant. Focusing first on the coefficients on the Treasury bill. Consistent with the results in Section 3.2, I find that the low profitability portfolio is more sensitive to interest rate than the high profitability portfolio and that the point estimates of the coefficients are all negative. The flexible econometric specification allows me to extend the previous finding in two ways. First, note that during recessions, the sensitivity of the return with respect to the interest rate is flipped from negative to positive. This holds true for 9 of the 10 profitability portfolios. The absolute magnitude of the estimates are roughly the same as during expansions. Secondly, the difference in sensitivity between expansion and recession is much higher for the low profitability firms than the high profitability firms. For Decile 1, the difference in interest rate coefficient between expansion and recession is 14.46 while the same difference for Decile 10 is only 8.12.

Turning now to the coefficients on default premium. Most of the coefficients are negative for the recessionary states but they are all statistically insignificant. The positive coefficients estimated for expansionary states are mostly statistically significant and positive as in the simple linear model. However, the coefficients do not show a clear pattern across the portfolios. However during recessions, the difference in point estimates are greater than during expansions. The general conclusion is that the default premium's positive correlation with the return are only precisely estimated in periods of economic expansions, and if there is indeed a negative relationship between profitability and sensitivity with the default premium as suggested by the results in Section 3.2, it might mainly be driven by recessionary periods.

Finally, the coefficients on the dividend yield across the deciles are qualitatively similar to that of the interest rate. They have mostly positive point estimates and tend to decrease in magnitude as profitability increases, so the low profitability firms tend to have stronger correlation with the aggregate risk premium as proxied by the dividend yield during expansions. During recessions, this phenomenon is also observed through the point estimates, but again none of them are statistically significant. In addition, there is evidence of asymmetry across the two states for low profitability versus high profitability portfolios. The low profitability portfolio's sensitivity to dividend yield increase from -0.05 to 0.31 as the economy changes from recession to expansion, while the high profitability portfolio's sensitivity to dividend yield actually decreases from 0.24 to 0.17.

I perform a simple stylized trading exercise to determine the economic value of the switching model. The results are presented in Table 5; lowest and highest profitability portfolios represent decile 1 and 10 of the profitability portfolios. The conditional return for time t is recursively calculated using only information prior to time t. The switching portfolio consist of positions in which if the conditional return is positive, I go long in the portfolio. Otherwise, I will invest in one-month T-Bill. The table shows the annualized return and volatility in percentage, as well as the Sharpe ratio. In the full sample, the switching portfolio generates higher return and lower volatility for the low profit firms. This leads to a much higher Sharpe ratio of 0.48 compared to 0.048 for the buy-and-hold portfolio. For the high profit firms, switching portfolio results in lower average return. However since the volatility of the switching portfolio is lower than the buy-and-hold, the Sharpe ratio ends up being higher. During recessions, the buy-and-hold portfolio for low profit firms has negative return of -21.54 percent while the buy-and-hold portfolio for high profit firms has negative return of -0.87 percentage. The switching portfolio increases the negative return to -4.11 percentage for the low profit firms while simultaneously decreasing the volatility. For the high profit firms, the switching portfolio increases the average return to a positive 4.95 percentage while again decreasing the volatility. The pattern during expansions is similar in which the switching portfolio always decreases the volatility and it also increases the average return of the low profit portfolio. For the high profit portfolio, switching results in a lower average return but it compensates by reducing the volatility by a significant amount, so that the risk-and-return profile improves in the form of a higher Sharpe ratio.

[Table 5 about here.]

Given the model of expected return, I can discuss the conditional profitability premium and its relationship with the business cycle. Figure 4 displays the expected profitability premium implied by the model from 1963 to 2013, using the parameter estimates from Table 4. The premium is defined as the difference between conditional excess return of the high profitability portfolio minus the conditional excess return of the low profitability portfolio. As before, the NBER recessions are shaded.

[Figure 4 about here.]

The conditional premium is mostly positive. It tends to increase prior to and during the early stage of recessions, and decrease afterwards. The qualitative inference is that the expected profitability premium is cyclical: it goes up during recessions. This is consistent with the results based on linear regression. As the economy plunge deeper into recessions, the unprofitable firms are struck harder.

Variations in expected return by itself does not lead to a full understanding of the risk associated with the portfolios. Investors care about both return and the volatilities associated with the return. A popular measure that summarize the premium per unit of risk is the Sharpe Ratio. Previous papers on expected Sharpe Ratio using U.S. data such as Kandel and Stambaugh (1990) and Tang and Whitelaw (2011) find that it is strongly cyclical. Figure 5 shows the conditional Sharpe Ratio of high profitability and the low profitability portfolios with NBER recessions shaded.

[Figure 5 about here.]

The ratios for profitable and unprofitable firms are very similar, and they are both strongly cyclical, in line with the previous findings. The ratio increases rapidly in the final stage of the recession and then quickly drops off. Finally, notice that the Sharpe Ratio for the high profitability portfolio is almost always higher than that of the low profitability portfolio.

Thus both the cyclicality of the profitability premium and the unconditional premium are not consistent with a systematic risk explanation. The unprofitability portfolio seems to have higher volatilty, lower return, lower premium per unit of risk, and they tend to underperform during bad times when the investors care the most about their wealth.

So far, I have investigated a set of relevant payoff characteristics that are usually associated with risk premium. Based on these characteristics, there is no evidence suggesting that profitable firms are fundamentally riskier. An alternative explanation for the premium is that the profitable firms are mispriced relative to the unprofitable firms due to some systematic bias on the part of the market participants. The next section examines this possibility.

4 Is there Mispricing?

There are much evidence in psychology that individuals can form their expectation of the future based on simple heuristics that do not take into full account the underlying process. These naive expectations can lead to distortions in stock prices in a predictable way. Academics have long argued that the predictive power of many financial ratios such as book-to-market is due to the fact that they capture systematic errors in the way investors form expectations about future return. Based on this framework, one possible reason for the low return of unprofitable firms is that the investors are overoptimistic about their future performances relative to the profitable firms. These firms might be currently capturing the attention of market participants because they are new companies in "hot" sectors. While they have very low cash flows currently, investors are willing to hold them due to a naive expectation about their future growth that is too extreme, and therefore fail to materialize.

This story is similar to the growth and glamour story advocated by LSV in explaining the value/growth effect. While the mechanism is the same, the way that the story plays out turn out to be quite different. The glamour story argues that naive investors become overoptimistic about the stocks that are in-favor due to a series of good news or good past performance. They naively extrapolate the performance of these firms, and if arbitrage by rational market participants are incomplete, this leads to overpricing of these firms. Moreover, the sell-side analysts have a tendency to recommend these glamourous stocks (Jegadeesh et al. (2004)), leading to overoptimism in their earnings forecast.

In the case of profitability, the overoptimism does not come from good past performance. Rather the unprofitable firms have worse past return and earnings, and they also tend to be newer firms and are more likely to be in financial distress. Investors seem to be betting on these bad performers recovering but in general, they do not.

4.1 Expectation Error

Table 6 displays the expectation error across the portfolios. For each firm-year in the sample, I use the last mean analyst forecast given prior to portfolio formation in July. The expectation error is computed as the difference between the mean "street" forecast earning and the actual "street" earning, both reported by IBES. I winsorize the extreme outliers at the 1% level and report the average error in expectation for each portfolio. The errors across all the profitability portfolios turn out to be positive, indicating that on average the analysts tend to give overly optimistic forecast. This is consistent with previous studies such as Dreman and Berry (1995) that shows analysts usually produce upwardly biased forecasts. This optimism can result from errors in processing earnings-related information or a rational response to their economic incentives. Easterwood and Nutt (1999) argue that this can also be due to analysts underreacting to negative information and overreacting to positive information.

More interestingly for our study, from the first row one sees a monotonically decreasing relationship as one moves from the most unprofitable portfolio to the most profitable portfolio. For the firms in the lowest profitability decile, the average expectation error is 0.175. This is more than five times the average expectation error in the highest profitability decile. Analysts' forecast for low profitability firms are biased much more in the positive direction than for unprofitable firms.

[Table 6 about here.]

Panel A of Table 6 also shows the expectation errors across different states of the economy. If the premium is indeed partly driven by the overoptimism of the investors, then how the overoptimism changes might yield some insight into the cyclical behavior of the premium. Specifically I consider states of the economy as given by recessions and expansions, as well as times of high sentiment and times of low sentiment. The recession and expansion periods are taken from NBER. Baker and Wurgler (2006) have argued that investor sentiments have an effect on the cross-section of stock returns. I use the sentiment index that they constructed, and define periods of high sentiment as those in which the sentiment index is above the

75 percentile, and periods of low sentiment to be times when the sentiment index is below the 25 percentile. The results show that analysts' positive bias increases during times of recession and times of high sentiment. Hribar and McInnis (2012) provide evidence that high sentiment leads to more optimistic forecast for "uncertain" and "difficult-to-value" firms. In this case, these firms are also exactly the unprofitable firms. During times of high sentiment, the upward bias increases more for the firms with low profitability. This might lead to further overpricing of these firms, and thus lower subsequent returns.

Panel B of the table displays the long term growth forecast of earnings given by the analysts, and the dispersion of growth forecast, earnings forecast, and forecast error as given by the standard deviation. I use the last growth forecast given by the analyst prior to portfolio formation. The lowest profitability decile has the highest average growth forecast of 25%. This suggests that the analysts are unconditionally most optimistic about these unprofitable firms. The dispersion of growth forecast, earnings forecast, and forecast errors are meant to capture the uncertainty regarding the performance of the firms. The result shows that dispersions of all the three variables tend to higher for the unprofitable firms. These unprofitable firms tend to be worse performing and therefore have higher uncertainty about their future prospects. Diether et al. (2002) provide evidence that stocks with higher dispersion in analysts' earnings forecast earn lower subsequent returns. While their results are consistent with mine, their explanation cannot be the driving force behind the profitability premium. They argue that when there is more uncertainty surrounding a stock, the prices will reflect the view of the more optimistic investors as those with the lowest valuation face a short-sell constraint. They are directly relating the uncertainty to future stock return, but it does not speak to why there should be higher systematic overoptimism for the firms with more dispersion *as given by analysts*. These analysts are not subject to any buying or selling constraints imposed by the structure of the market as they are just giving their estimate of the firm's performance.

Ball et al. (2015) show that profitability can predict returns up to five years into the future. From that, they argue that the driver of the premium must be risk-based. Mispricings tend to be corrected once market participants discover them and these opportunities cannot last for extended period of time. Therefore, the fact that profitability has persistent power in predicting future returns is an indication that it is a true risk premium. This argument, however, is predicated on the fact that market participants understand the mapping between profitability and subsequent return. The new measures of profitability that have been the interest of recent papers are different from the return on equity that investors used to focus on. It is very possible that market participants have not been paying attention to these ratios prior to the study. This, along, with the high persistence of the profitability measure itself, can contribute to its ability to forecast returns up to several years into the future.

Figure 2 shows the persistence of profitability. I examine the firms that were in each of the profitability deciles five years prior, as shown in the horizontal axis. For the firms in each decile, I plot the average decile portfolio that they are in in the current year. The figure shows that out of the firms in decile 1 of profitability five years before, their average profitability decile is about 3.8. Out of the firms that are in the highest decile of profitability, their average current decile is 8. There is a monotonically increasing relationship between profitability deciles five years prior and today. Thus profitability itself tends to persist into the future. A firm that is profitable this year will tend to still be profitable five years from now.

[Figure 6 about here.]

In order for the mispricing story to be consistent with persistent forecasting power of profitability, the market participants should not understand the mapping between profitability and future performance completely. If they do, then even if profitability itself is persistent, they learn after the observing their mistakes in the first year and correct their expectation for subsequent years. Figure 3 explores whether investors learn from their past errors. The horizontal axis displays the portfolio deciles, while the vertical axis shows the average analyst forecast errors. I investigate the expectation error for portfolio formed from year $t-1$ to year $t-5$.

[Figure 7 about here.]

The results show that the expectation error can persist for up to five years into the future. This is perhaps a little bit surprising. The blue line shows the expectation error for the portfolio formed just one year prior. This is the same as the results shown in Table 6, and one sees a clear monotonically decreasing trend in forecast error as the profitability decile increases. For the portfolios formed from $t-2$ to $t-5$, one still sees a monotonically decreasing relationship. The only exception is the lowest profitability decile 1. This is because the firms that miss their forecast by the most tend to drop out of the sample, so the ones that are still in the sample tend to have done better and mechanically less forecast error. From decile

2 to decile 10, the analyst forecast error decreases, even for the portfolios that are formed five years before. The investors' errorenous perception of the mapping between profitability and expectation errors actually persist. This evidence casts some doubt on the notion that investors actually understand the mapping between profitability and performance. In fact, it seems that they are not paying attention to this measure, and the expectation errors can last years into the future.

4.2 What Drives the Expectation Error?

Given the wedge between forecasted and actual performance, there is a misunderstanding of the mapping from profitability to future stock return. This might be due to investors previously not focusing on the new measure of economic profit that is the subject of recent research. But why is the bias systematically decreasing in the profitability? There are two possible reasons for what leads to the over-optimism of investors. The first possibility, suggested by LSV, is that investors naively extrapolate the past good performance of firms too far into the future. They believe that firms should continue to do well in the future and therefore over-estimate the earnings. If this hypothesis is true, then unprofitable firms should have better prior performance, perhaps using different measure of performance other than profitability, than profitable firms. Alternatively, the over-optimism might result from failing to anticipate how persistent the bad performance of unprofitable firms can be. This hypothesis would suggest that unprofitable firms have done poorly prior to portfolio formation, and they continue to do poorly just after portfolio formation. Investors naively believe that these underperforming firms will mean-revert and recover, but their expectations fail to materialize. This section tries to distinguish between the two possibilities.

I start off by examining some firm characteristics across the profitability portfolios as shown in Panel A of Table 7. Consistent with prior work, high profitability firms tend to be larger in size and they tend to have low book-to-market. This negative correlation means that profitability can be used as a hedge for value strategies and they work especially well jointly. The lowest profitability decile portfolio has the smallest median firm age of 5.6, suggesting that they tend to be younger firms. The highest profitability decile portfolio has a median age of 7.3 while across the portfolios, the highest median age is 11. So while the youngest firms are not profitable, the most mature firms are not profitable either because they are already past their fastest-growing phase. The most profitable firms tend to be somewhere in between.

[Table 7 about here.]

The unprofitable firms tend to have the highest net stock issuance. This is mostly concentrated in the lowest profitability decile, with the average net issuance more than twice that of any other portfolios. Previous studies such as Pontiff and Woodgate (2008) have shown that firms tend to underperformance after seasoned equity offerings. This negative performance is often attributed to market timing by the corporate insiders. When they feel that the stock is overpriced, they issue shares to cash in on the over-valuation. In the case of the profitability premium, this explanation is difficult to reconcile with the fact that the unprofitable firms have been performing badly over the past few years. Indeed, Loughran and Ritter (1995) showed that firms tend to issue new equities following a run-up in their prices. The high net issuance of the unprofitable firms might instead be because these are young firms that need capital. Looking at earnings yield and asset growth also point to the same direction. Unprofitable firms have negative earnings on average, again mostly concentrated in the lowest profitability decile. The earnings yield does not exhibit a strong pattern across the other profitability portfolios. The most unprofitable firms also have negative asset growth and the asset growth increases across the profitability portfolios. Both suggest that these are not the glamorous firms that have been quickly growing or expanding. Instead, they seem to be struggling and shedding assets.

One of the anomalies that the profitability factor is able to explain is the return premium associated with financial distress. Firms in higher distress (higher probability of bankruptcy) tend to earn lower returns. Panel B of Table 7 studies several measures of distress across the profitability portfolios. EDF is the Expected Default Frequency given by the Merton (1974) model that computes default probability for individual firms at the monthly frequency. The model recognizes the equity of the firm as a call option on the firm's unobserved underlying value with a strike price equal to the book value of the firm's debt. It then proceeds to apply classic option pricing theory to obtain the firm value and volatility. The firm's probability of default is backed out as the probability that the firm's value will drop below the face value of its debt. This model has appeared widely in academic papers such as Vassalou and Xing (2004) and Bharath and Shumway (2008). The details of the derivation and implementation of the iterative process can be found in Bharath and Shumway (2008). I also examine accounting based measures of distress as given by the Ohlson (1980) O-score, label "Ohlson

O-Probability" is 0.43 percent. This probability is calculated via a logistic transformation of the Ohlson O-score[2]. Finally I obtain credit rating and default data from Moody's Default and Recovery Database (DRD) and match it to the firms in the sample.

More profitable firms have lower distress risk than less profitable firms by all of these measures. Moreover, most of the differences are driven by the high distress of low profitability portfolio. The O-Probability ranges from 0.06 percent to 0.15 percent for Portfolios 2 to 10, but jumps more than ten times to 1.69 for the least profitable Portfolio 1. The EDF for portfolio 1 to 10 are monotonically decreasing with the difference between Decile 1 and Decile 2 significantly bigger than the difference between any two consecutive portfolios. The proportion of default for the lowest decile is 0.44% and it decreases to merely 0.03% for the highest decile. The overall conclusion is that the unprofitable firms tend to be performing poorly based on other metrics and they have higher financial distress, with the distress risk mostly concentrated in the lowest decile of profitability.

Table 8 shows the past and future stocks returns as well as sales and earnings growth across the profitability portfolios. The average past year return of firms in the lowest profitability decile is 5.5% while the average return for the firms in the highest profitability decile is 25.2%. There is a monotonically increasing relationship between profitability and past 1 year return. The same relationship holds for the past 5-year cumulative returns in the next row. Contrary to the glamour story, it is in fact the profitable firms that have done well in the past. Over the year subsequent to portfolio formation, the firms that are profitable continue to do well, while the firms that are unprofitable continue to underperform. Over the next 5 years, however, the returns equalizes somewhat and the clear monotonic relationship disappear, even though the cumulative return for the least profitable portfolio still lags that of the other portfolios. It is almost as if the investors expect the performance to mean-revert with the

[2]The O-score is defined as

$$\begin{aligned} O-score \quad = \quad & -1.32 - 0.407 log(total\ assets) + 6.03\left(\frac{total\ liabilities}{total\ assets}\right) - 1.43\left(\frac{working\ capital}{total\ assets}\right) \\ & +0.076\left(\frac{current\ liabilities}{current\ assets}\right) - 1.72\boldsymbol{I}(total\ liabilities > total\ assets) - 2.37\left(\frac{net\ income}{total\ assets}\right) \\ & -1.83\left(\frac{funds\ from\ operation}{total\ liabilities}\right) + 0.285\boldsymbol{I}(net\ loss\ for\ last\ two\ years) \\ & -0.521\left(\frac{net\ income_t - net\ income_{t-1}}{|net\ income_t| + |net\ income_{t-1}|}\right) \end{aligned}$$

where $\boldsymbol{I}(\cdot)$ is an indicator function that equals 1 if the condition is satisfied and 0 otherwise.

high profitability firms underperforming while the low profitability firms outperforming but the mean-reversion does not happen as quickly as the investors think. It takes more than 5 years for the reversion in stock return to materialize, and over the next year the profitable firms continue its trend of outperformance.

[Table 8 about here.]

The table also examines accounting measures of performance as given by the sales and earnings per share (EPS) growth. I compute the annualized growth rates by fitting a least squares growth line to the logarithms of the six annual observations. Thus the growth rate for year t is computed as the least squares growth line from year $t-5$ through year t. This follows the same construction procedure used by IBES and Dechow and Sloan (1997). I only keep the variable if there are at least 5 years of data available.

The pattern of short term trend and long term reversal is more pronounced through sales and EPS growth. Over the past 5 years, the low profitable firms have slower sales growth than the profitable firms. Over the next 5 years, however, the sales growth becomes similar across the portfolios, suggesting that the sales mean-revert over the subsequent 5 year period. The EPS growth for the lowest profitability decile is -0.0184 as compared to 0.0244 for the highest profitability decile. Again, unprofitable firms have lower earnings growth so they are not really the glamour firms in the sense of having good past performances. Interestingly over the next 5 years, the EPS growth pattern reverses itself. The lowest profitability decile has a growth rate of 0.0306 while the highest profitability decile has a negative growth rate of -0.0105. Mean-reversion in performance does happen, but it happens over the next 5 years.

Finally, I carry out a portfolio double sort to further understand the relationship between expectation and profitability. The main proxy of expectation that I use is the median growth forecast of the companies issued by the analysts. A higher forecasted growth rate indicates a more positive expectation for the performance of the firm. The goal of the exercise is to investigate whether the difference in expectation captured by profitability is reflected in stock returns, and if so, what would that tell us about the the reason behind the expectation error.

[Table 9 about here.]

Table 9 displays the results with value-weighted returns. The spread in profitability premium is mainly concentrated among the firms with the highest forecast of growth. The average monthly percentage return differential between the high profit and low profit firms in Quintile 1 of forecasted growth is 0.29, and the differentials become -0.2 and 0 in quintile 2 and 3. For Quintile 4 the differential is slightly higher at 0.13. The return differential is 1.11 in the highest estimated growth quintile, more than three times higher than any of the other return differentials. The striking result is that the lowest portfolio return by far is the low-profitability, high expectation portfolio (0.19%). This is consistent with the expectation error hypothesis. Those with the high forecast growth are exactly the firms that investors are most optimistic about, and the low profitability firms are the firms that are more likely to not do well. Thus one would expect these firms to miss their forecast by the most, and therefore displays the lowest return.

The table also displays the average forecast error, forecast growth, and profitability of each of the 25 portfolios. These numbers provide additional insight about exactly what the investors are wrong about. First look at profitability across the portfolios. For profitability quintiles 2 to 5, there is virtually no difference across the forecast growth. The profitability of the firms with the lowest growth forecast is roughly the same as the profitability of the firm with the highest growth forecast. Something more interesting is going on with Quintile 1. The profitability of those with the highest growth forecast is much lower than the rest. This means that among these unprofitable firms, the analysts are expecting the highest growth for the firms that are extremely unprofitable. They are expecting the performance of these firms to mean-revert back up. Yet these firm did not as they again tend to miss their earning forecast on the low side. This contribute to the extreme small return observed for this portfolio. An alternative way to look at this is through the growth forecast of each of the portfolios. The highest growth forecast is given by the portfolio in Quintile 5 of growth forecast and Quintile 1 of profitability. Thus, among those companies with high forecast growth, analysts are most optimistic about the extremely unprofitable firms as they expect poor performers to recover.

5 Conclusion

Profitability has been shown to be an significant and robust determinant of the cross-section of stock returns. It works exceptionally well in combination with book-to-market in generating high spreads in return premium. More importantly, it has been shown to be able to explain most of the existing asset pricing anomalies. Profitability's importance has led to it being incorporated into a new benchmark factor model. Given its exceptional empirical performance, the natural question is to investigate whether the economic mechanism behind its power is based on systematic risk or behavioral mispricing.

This paper finds that the time series pattern of the premium is not consistent with systematic risk. The profitability premium is quite consistent through time. Firms that are profitable have higher returns and less volatility than the unprofitable firms, leading to a significantly higher Sharpe ratio. The profitable firms also have lower drawdown. More fundamentally, the premium is shown to be countercyclical, with the premium increasing during bad times. Thus the low profit firms perform even worse during recessions when the marginal utility of wealth is low. Unprofitable firms do not provide a hedge for bad times.

I investigate the channels for mispricing and show that the analysts tend to be much more optimistic and overestimate the earnings of the firms that are unprofitable relative to the profitable firms. This forecast error is surprisingly persistent. Porfolios that are formed based on profitability deciles five year prior still display this monotonic relationship in forecast errors. I argue that the forecast error is not due to the glamour effect, in which market participants extrapolate the future performance of the firms based on past performance. Indeed the unprofitable firms are not the firms that have performed well in the past. They tend to be younger firms without consistent positive cash flow, with higher probability of being in financial distress.

While one cannot conclusively reject the possibility that profitability is related to systematic risk, any risk-based theoretical model that attempts to explain it in the future must take into consideration its correlational structure with the macro business cycle.

References

Baker, M., and J. Wurgler. 2006. "Investor Sentiment and the Cross-Section of Stock Returns." *Journal of Finance*, 61(4): 1645–1680.

Ball, R., and P. Brown. 1968. "An Empirical Evaluation of Accounting Income Numbers." *Journal of Accounting Research*, 6(2): 159–178.

Ball, R., J. Gerakos, J. Linnainmaa, and V. Nikolaev. 2015. "Deflating Profitability." *Journal of Financial Economics*, 117(2): 225–248.

Barberis, N., and R. Thaler. 2003. *A Survey of Behavioral Finance.* Chap. 18 1053–1128: North-Holland.

Bharath, S., and T. Shumway. 2008. "Forecasting Default with the KMV-Merton Model." *Review of Financial Studies*, 21 1339–1369.

Campbell, J. 1987. "Stock Returns and the Term Structure." *Journal of Financial Economics*, 18 373–399.

Campbell, J., and R. Shiller. 1988. "The Dividend-Price Ratio and Expectation of Future Dividends and Discount Factors." *Review of Financial Studies*, 1 195–228.

Campbell, J., and T. Vuolteenaho. 2004. "Bad Beta, Good Beta." *American Economic Review*, 94(5): 1249–1275.

Chen, L., R. Petkova, and L. Zhang. 2008. "The Expected Value Premium." *Journal of Financial Economics*, 87 269–280.

Chen, L., and X. Zhao. 2009. "Return Decomposition." *Review of Financial Studies.*

Dechow, P., and R. Sloan. 1997. "Returns to Contrarian Investment Strategies: Test of Naive Expectation Hypothesis." *Journal of Financial Economics*(43): 3–27.

Diether, K., C. Malloy, and A. Scherbina. 2002. "Differences of Opinion and the Cross Section of Stock Returns." *Journal of Finance*, 57(5): 2113–2141.

Ding, Z. 2012. "An Implementation of Markov Regime Switching Model with Time Varying Transition Probabilities in Matlab." SSRN elibrary.

Dreman, D., and M. Berry. 1995. "Analyst Forecasting Errors and their Implications for Security Analysis." *Financial Analyst Journal*, 51 30–40.

Easterwood, J, and S. Nutt. 1999. "Inefficiency in Analysts' Earnings Forecast: Systematic Misreaction or Systematic Optimism?" *Journal of Finance*, 54(5): 1777–1797.

Engelberg, J., R. McLean, and J. Pontiff. 2015. "Anomalies and News." *Working.*

Fama, E. 1981. "Stock Returns, Real Activity, Inflation, and Money." *American Economic Review*, 71 545–565.

Fama, E., and K. French. 1988. "Dividend Yields and Expected Stock Returns." *Journal of Financial Economics*, 22 3–25.

Fama, E., and K. French. 1989. "Business Conditions and Expected Returns on Stocks and Bonds." *Journal of Financial Economics*, 25 23–49.

Fama, E., and K. French. 1992. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 427–465.

Fama, E., and K. French. 1996. "Multifactor Explanations of Asset Pricing Anomalies." *Journal of Finance*, 51 55–84.

Fama, E., and K. French. 2006. "Profitability, Investment and Average Returns." *Journal of Financial Economics*, 82 491–518.

Fama, E., and K. French. 2014. "A Five-Factor Asset Pricing Model." *Working Paper*.

Fama, E., and G. Schwert. 1977. "Asset Returns and Inflation." *Journal of Financial Economics*, 5 115–146.

Filardo, A. 1994. "Business Cycle Phases and Their Transition Dynamics." *Journal of Business and Economic Statistics*, 12 299–308.

Gray, S. 1996. "Modeling the Conditional Distribution of Interest Rates as a Regime Switching Process." *Journal of Financial Economics*, 42 24–62.

Gulen, H., Y. Xing, and L. Zhang. 2011. "Value versus Growth: Time-Varying Expected Stock Returns." *Financial Management*, 40 381–407.

Hamilton, J. 1989. "A New Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle." *Econometrica*, 57 357–384.

Hamilton, J., and G. Lin. 1996. "Stock Market Volatility and the Business Cycle." *Journal of Applied Econometrics*, 11 573–593.

Hodrick, R. 1992. "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement." *Review of Financial Studies*, 5 357–386.

Hou, K., C. Xue, and L. Zhang. 2014. "Digesting Anomalies: An Investment Approach." *Review of Financial Studies*, doi: 10.1093/rfs/hhu068.

Hribar, P., and J. McInnis. 2012. "Investor Sentiment and Analysts' Earnings Forecast Errors." *Management Science*, 58(2): 293–307.

Jagannathan, R., and Z. Wang. 1996. "The Conditional CAPM and the Cross-Section of Expected Returns." *Journal of Finance*, 51 3–53.

Jegadeesh, N., J. Kim, S. Krische, and C. Lee. 2004. "Analyzing the Analysts: When Do Recommendations Add Value?" *Journal of Finance*(59): 1083–1124.

Kandel, S., and R. Stambaugh. 1990. "Expectation and Volatility of Consumption and Asset Returns." *Review of Financial Studies*, 3 207–232.

Keim, D., and R. Stambaugh. 1986. "Predicting Returns in the Stock and Bond Markets." *Journal of Financial Economics*, 17 357–390.

La Porta, R. 1996. "EExpectation and the Cross-Section of Stock Returns." *Journal of Finance*, 51(5): 1715–1742.

Lakonishok, J, A. Shleifer, and R. Vishny. 1994. "Contrarian Investment, Extrapolation, and Risk." *Journal of Finance*, 49(5): 1541–1578.

Loughran, T, and J. Ritter. 1995. "The New Issues Puzzle." *Journal of Finance*, 50(1): 23–51.

Merton, R. 1974. "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates"." *Journal of Finance*, 29 449–470.

Novy-Marx, R. 2013. "The Other Side of Value: The Gross Profitability Premium." *Journal of Financial Economics*, 108 1–28.

Ohlson, J. 1980. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research*, 18 109–131.

Perez-Quiros, G., and A. Timmermann. 2000. "Firm Size and Cyclical Variations in Stock Returns." *Journal of Finance*, 55 1229–1262.

Perlin, M. 2014. "MS_Regress The MATLAB Package for Markov Regime Switching Models." SSRN elibrary.

Pontiff, J., and A. Woodgate. 2008. "Share Issuance and Cross-Sectional Returns." *Journal of Finance*, 63(2): 921–945.

Schwert, G. 1989. "Why Does Stock Market Volatility Change Over Time?" *Journal of Finance*, 29 1115–1153.

Stock, J, and M. Watson. 1989. *NBER Macroeconomics Annual.* Chap. New Indexes of Coincident and Leading Economic Indicators 351–409: MIT Press.

Tang, Y., and R. Whitelaw. 2011. "Time-Varying Sharpe Ratios and Market Timing." *Quarterly Journal of Finance*, 1 465–493.

Vassalou, M., and Y. Xing. 2004. "Default Risk in Equity Returns." *Journal of Finance*, 59 831–868.

Wang, H., and J. Yu. 2013. "Dissecting the Profitability Premium." *Working Paper*.

Whitelaw, R. 1994. "Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns." *Journal of Finance*, 49 515–541.

Table 1: **Risk Across Portfolios**

This table reports various measures of risk being applied to the ten operating profitability portfolios. Panel A displays the average holding period returns for 1 month, 1 year, 3 year and 5 years. It also shows the volatility and Sharpe ratio of the portfolios, as well as the worse drawdown for periods ranging from 3 months to 2 year. Panel B decomposes the beta of the portfolios with respect to the market into two components: one associated with discount rate news and one associated with cash flow news. The sample excludes financial firms (those with one-digit standard industrial classification code of six) and goes from July 1963 to December 2013.

	Profitability Portfolios									
	Low	2	3	4	5	6	7	8	9	High
Panel A: Return-based Measures										
Monthly Return	0.00163	0.004	0.00515	0.00516	0.0059	0.00586	0.00514	0.00671	0.00565	0.00565
1-Year Holding Return	0.075	0.1	0.114	0.113	0.123	0.12	0.113	0.132	0.118	0.118
3-Year Holding Return	0.197	0.304	0.359	0.35	0.391	0.387	0.363	0.426	0.377	0.385
5-Year Holding Return	0.3	0.537	0.66	0.644	0.708	0.732	0.683	0.8	0.693	0.717
Volatility	0.236	0.175	0.167	0.166	0.167	0.162	0.165	0.170	0.162	0.169
Sharpe Ratio	0.084	0.281	0.380	0.385	0.439	0.448	0.386	0.491	0.431	0.413
Worst 3-Month Drawdown	-0.51	-0.33	-0.37	-0.36	-0.39	-0.26	-0.34	-0.37	-0.31	-0.33
Worst 6-Month Drawdown	-0.58	-0.51	-0.43	-0.47	-0.48	-0.35	-0.44	-0.41	-0.34	-0.34
Worst 1 Year Drawdown	-0.74	-0.54	-0.44	-0.49	-0.49	-0.4	-0.44	-0.44	-0.42	-0.45
Worst 2 Year Drawdown	-0.86	-0.6	-0.49	-0.46	-0.5	-0.42	-0.45	-0.41	-0.45	-0.5
Panel B: Beta Decomposition										
Discount Rate Beta	0.676	0.526	0.522	0.529	0.482	0.483	0.543	0.571	0.571	0.585
Cash Flow Beta	0.792	0.584	0.521	0.556	0.552	0.519	0.522	0.556	0.498	0.499

Table 2: Summary Statistics for Value, Size and Profitability Premium

This table reports descriptive statistics of value-weighted portfolios based on size, book to market, and operating profitability. For each variable, the firms are sorted into ten deciles based NYSE breakpoints. The premium is calculated as the difference between the top decile and the bottom decile. The market return is given by the value-weighted CRSP index. Recession periods are defined by the NBER. The sample excludes financial firms (those with one-digit standard industrial classification code of six). The sampling period is from July 1963 to December 2013

	Premium	Recession Premium	Expansion Premium	Correlation with Market Return
Value Premium	0.67	0.76	0.66	-0.0493
Size Premium	0.37	1.03	0.27	0.321
Profitability Premium	0.40	0.69	0.36	-0.334

Table 3: Long Horizon Regression of Profitability Portfolios on Aggregate Variables

This table reports long horizon regression results for each of the ten profitability deciles. The independent variables include one-month U.S. treasury rate, the aggregate default spread as given by the yield difference between Moody's Aaa and Baa bonds, and the aggregate dividend yield. Regression results are reported for one-month, two-month, six-month, and twelve-month cumulative returns. The Newey-West standard errors correcting for overlapping returns are reported in parentheses. The sample excludes financial firms (those with one-digit standard industrial classification code of six). The sampling period is from July 1963 to December 2013

	Profitability Deciles									
	Low Profitability	2	3	4	5	6	7	8	9	High Profitability
1-Month Return										
Treasury 1 Month	-6.032***	-2.249*	-2.890**	-2.843**	-3.046**	-2.705**	-2.668**	-3.128***	-3.083***	-2.988**
	(1.858)	(1.354)	(1.169)	(1.271)	(1.338)	(1.197)	(1.164)	(1.201)	(1.154)	(1.194)
Default Spread	0.665	0.217	0.756	0.512	0.506	0.495	0.379	0.634	0.423	0.313
	(0.913)	(0.743)	(0.577)	(0.631)	(0.604)	(0.558)	(0.582)	(0.540)	(0.521)	(0.499)
Div Yield	0.486***	0.203**	0.233***	0.245**	0.231**	0.231**	0.284***	0.223**	0.231**	0.227**
	(0.159)	(0.1000)	(0.0897)	(0.0981)	(0.113)	(0.103)	(0.0997)	(0.104)	(0.0935)	(0.113)
2-Month Return										
Treasury 1 Month	-11.89***	-4.197*	-5.407***	-5.144**	-5.746***	-4.553**	-4.832**	-5.683***	-5.605***	-5.376***
	(3.248)	(2.351)	(1.934)	(2.249)	(2.204)	(1.980)	(1.988)	(1.931)	(1.804)	(1.846)
Default Spread	1.389	0.438	1.446	0.988	1.035	0.992	0.735	1.155	0.675	0.514
	(1.703)	(1.442)	(1.080)	(1.254)	(1.154)	(1.057)	(1.106)	(0.983)	(0.987)	(0.914)
Div Yield	0.979***	0.402**	0.461***	0.482***	0.449**	0.423**	0.547***	0.437***	0.454***	0.445***
	(0.264)	(0.165)	(0.154)	(0.166)	(0.177)	(0.164)	(0.163)	(0.158)	(0.143)	(0.170)
6-Month Return										
Treasury 1 Month	-32.38***	-10.56*	-13.00***	-12.32**	-14.79***	-9.961**	-11.72**	-13.81***	-13.25***	-11.59**
	(8.041)	(5.392)	(4.806)	(5.321)	(4.850)	(4.720)	(4.955)	(4.855)	(4.872)	(5.367)
Default Spread	6.056*	3.494	5.155**	4.389*	4.761**	4.576**	3.515	4.682**	2.849	2.500
	(3.621)	(2.673)	(2.375)	(2.453)	(2.381)	(2.217)	(2.219)	(2.282)	(2.212)	(2.274)
Div Yield	2.683***	0.989**	1.229***	1.263***	1.175***	1.034***	1.473***	1.135***	1.198***	1.100**
	(0.759)	(0.494)	(0.397)	(0.399)	(0.382)	(0.371)	(0.410)	(0.394)	(0.391)	(0.505)
12-Month Return										
Treasury 1 Month	-54.47***	-19.43*	-22.71***	-19.33*	-25.70***	-15.01	-17.99*	-21.17**	-18.87*	-15.46
	(14.86)	(10.02)	(8.640)	(10.01)	(8.676)	(9.394)	(10.05)	(9.794)	(10.61)	(11.09)
Default Spread	9.402*	5.500	8.248**	5.976*	9.238***	7.784**	4.705	7.103**	2.773	2.203
	(5.078)	(3.735)	(3.456)	(3.495)	(3.429)	(3.285)	(3.290)	(3.552)	(3.350)	(3.686)
Div Yield	4.670***	1.908*	2.382***	2.436***	2.202***	1.937***	2.882***	2.057***	2.164***	1.881*
	(1.524)	(1.096)	(0.704)	(0.707)	(0.732)	(0.667)	(0.789)	(0.730)	(0.735)	(1.061)

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: **Coefficient Estimates in the Markov Regime Switching Model**

This table reports coefficient estimates for the two-state Markov Regime Switching model for each of the ten profitability deciles. The dependent variables are the portfolio excess returns. The independent variables include one-month U.S. treasury rate, the aggregate default spread as given by the yield difference between Moody's Aaa and Baa bonds, and the aggregate dividend yield. Standard errors are reported in parentheses. The sample excludes financial firms (those with one-digit standard industrial classification code of six). The sampling period is from July 1963 to December 2013

	Profitability Decile									
	Low	2	3	4	5	6	7	8	9	High
Mean Parameters										
Constant, State 1	-0.064**	-0.0089	-0.032*	-0.05**	-0.01	-0.045	-0.043**	-0.034*	-0.04*	-0.036
	(0.025)	(0.024)	(0.019)	(0.023)	(0.017)	(0.023)	(0.021)	(0.02)	(0.021)	(0.022)
Constant, State 2	-0.029	-0.017	-0.06	0.014	-0.09	0.0076	-0.083	-0.041	-0.093	-0.077
	(0.18)	(0.05)	(0.11)	(0.057)	(0.085)	(0.05)	(0.072)	(0.096)	(0.12)	(0.061)
Interest Rate, State 1	-7.2***	-3.09***	-3.9***	-3.7***	-4.05***	-3.8***	-3.62***	-4.11***	-4.05***	-4.16***
	(1.29)	(1.14)	(1.08)	(1.16)	(0.89)	(1.06)	(1.11)	(1.08)	(1.13)	(1.13)
Interest Rate, State 2	7.26	3.02	3.8	3.55	3.94	3.49	-3.56	4	3.99	3.96
	(12.87)	(3.37)	(8.08)	(3.25)	(5.49)	(3.58)	(7.39)	(9.55)	(6.65)	(8.16)
Default Premium, State 1	1.04*	0.55	1.18**	1.08**	1.33***	1.53***	0.98*	1.1**	1.17**	0.99*
	(0.57)	(0.56)	(0.54)	(0.54)	(0.43)	(0.53)	(0.50)	(0.46)	(0.52)	(0.51)
Default Premium, State 2	-0.58	0.041	-0.61	-0.19	-0.24	-0.25	-0.26	-0.49	-0.32	-0.27
	(4.72)	(1.1)	(2.35)	(1.16)	(2.05)	(1.23)	(2.12)	(3.2)	(2.15)	(2.73)
Dividend Yield, State 1	0.31**	0.081	0.16*	0.22**	0.081	0.19**	0.2**	0.17*	0.19**	0.17*
	(0.11)	(0.098)	(0.084)	(0.094)	(0.071)	(0.096)	(0.088)	(0.088)	(0.088)	(0.092)
Dividend Yield, State 2	-0.05	0.011	0.13	-0.12	0.26	-0.09	0.21	0.052	0.2	0.24
	(0.88)	(0.23)	(0.53)	(0.26)	(0.38)	(0.24)	(0.35)	(0.48)	(0.49)	(0.34)
Variance Parameters										
σ, State 1	0.0022***	0.001***	0.0014***	0.001***	0.0011***	0.0012***	0.0014***	0.0015***	0.0015***	0.0015***
	(0.00016)	(0.0001)	(0.00012)	(0.00012)	(0.00091)	(0.00012)	(0.00013)	(0.00013)	(0.00014)	(0.00012)
σ, State 2	0.021***	0.0051***	0.0062***	0.0051***	0.0072***	0.0046***	0.0058***	0.0073***	0.0057***	0.006***
	(0.0048)	(0.00052)	(0.0011)	(0.00079)	(0.001)	(0.00078)	(0.0011)	(0.0018)	(0.0015)	(0.0013)
Log Likelihood	860.6	1004.42	1021.8	1027.85	1052.98	1031.78	1024.71	1009.73	1021.64	1006.85

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: **Out of Sample Trading Exercise: September 1988 to December 2013**

This table reports the results from an out-of-sample trading experiment using the Markov-Regime Switching model. The dependent variables are the portfolio excess returns. The independent variables include one-month U.S. treasury rate, the aggregate default spread as given by the yield difference between Moody's Aaa and Baa bonds, and the aggregate dividend yield. Trading results compares the returns of T-Bill with the lowest and highest decile of portfolios sorted on profitability. The buy-and-hold strategy simply reinvest all the funds in the relevant portfolio. Switching portfolios take a long position if the recursively estimated mean is positive, otherwise it will invest in one-month T-Bill. All variables are annualized.

		Lowest Profitability		Highest Profitability	
	T-Bills	Buy-and-Hold	Switching Portfolio	Buy-and-Hold	Switching Portfolio
Full Sample					
Mean Return	3.27	4.6	9.98	13.44	12
SD of Return	0.66	26.92	13.1	17.15	11.41
Sharpe Ratio		0.048	0.48	0.55	0.72
Recession					
Mean Return	2.8	-21.54	-4.11	-0.87	4.95
SD of Return	0.68	45.04	19.3	24.6	15.79
Sharpe Ratio		-0.59	-0.55	-0.15	0.18
Expansions					
Mean Return	3.37	8.79	12.17	15.56	13.01
SD of Return	0.65	23.24	11.98	15.83	10.68
Sharpe Ratio		0.22	0.98	0.71	1.08

Table 6: **Expectation Errors Across Portfolios**

Panel A of this table reports the analyst forecast error across the operating profitability portfolios. The error is calculated as difference between the forecast earnings per share and the actual earnings per share reported by IBES, winsorized at the 1% level. The table also reports the forecast error across recessions and expansions, and for times of low sentiment and times of high sentiment. Sentiment index is taken from Jeffrey Wurgler's website and high sentiment is defined as periods in which the index is above the 75 percentile, and low sentiment is defined as periods in which the index is below the 25 percentile. Panel B of the table reports the growth forecast by the analyst as well as the standard deviation of the growth forecast, earnings forecast and forecast errors. The sample period is from July 1975 to December 2013.

	Profitability Portfolios vs Mis-Expectation									
	Low	2	3	4	5	6	7	8	9	High
Panel A: Expectation Error										
All Periods	0.175	0.161	0.13	0.108	0.101	0.091	0.084	0.06	0.053	0.034
Expansion	0.166	0.153	0.13	0.104	0.1	0.093	0.086	0.061	0.055	0.034
Recession	0.255	0.235	0.133	0.145	0.113	0.08	0.071	0.046	0.038	0.043
Low Sentiment	0.121	0.098	0.053	0.065	0.051	0.037	0.029	0.013	0.018	0.012
High Sentiment	0.206	0.198	0.173	0.134	0.133	0.124	0.118	0.088	0.074	0.046
Panel B: Growth Forecast and Dispersions										
Median Growth Forecast	0.25	0.17	0.15	0.15	0.16	0.16	0.16	0.17	0.17	0.2
Dispersion of Growth Forecast	0.81	0.52	0.44	0.41	0.42	0.41	0.4	0.39	0.39	0.45
Dispersion of Forecast	0.168	0.218	0.094	0.087	0.084	0.081	0.075	0.071	0.06	0.052
Dispersion of Forecast Error	0.64	0.65	0.58	0.55	0.53	0.51	0.47	0.4	0.39	0.3

Table 7: **Characteristics Across Portfolios**

This table reports the firm characteristics for the ten decile portfolios sorted on operating profitability. Earnings yield is defined as the earnings per share divided by price. Net issuance is defined as the change in the natural logarithm of number of shares outstanding. Asset growth is defined as the change in the natural logarithm of assets per share outstanding. The sample excludes financial firms (those with one-digit standard industrial classification code of six). The sample period is from July 1963 to December 2013.

	Profitability Portfolios									
	Low	2	3	4	5	6	7	8	9	High
Panel A: Firm Characteristics										
Net Issuance	0.179	0.086	0.068	0.067	0.06	0.056	0.059	0.069	0.057	0.066
Size	285	1112	875	958	1139	1348	1614	1910	2254	2909
Book-to-Market	1.076	1.25	1.14	1.04	0.952	0.853	0.757	0.671	0.577	0.416
Asset Growth	-0.021	0.054	0.065	0.064	0.071	0.08	0.085	0.081	0.103	0.14
Earnings Yield	-0.387	-0.023	0.026	0.048	0.051	0.065	0.073	0.061	0.071	0.06
Mean Firm Age	9.57	15.36	16.78	16.62	16.58	16	15.62	15.17	14.2	11.56
Median Firm Age	5.6	10	11	11	11	10.6	10.5	10.3	9.6	7.3
Panel B: Distress Measures										
Ohlson O-score	1.69	0.15	0.13	0.11	0.09	0.08	0.06	0.06	0.06	0.11
EDF	0.197	0.138	0.111	0.089	0.078	0.066	0.056	0.048	0.042	0.038
Downgrade	1.81	1.53	1.49	1.31	1.32	1.07	1.13	1.04	0.96	1.06
Default	0.44	0.18	0.11	0.1	0.04	0.03	0.04	0.05	0.03	0.03

Table 8: **Sales and Earnings Pattern**

This table reports the stock return, sales and earnings trend across the profitability portfolios. Panel A shows the past one year, past five year, future one year, and the future five year return of the ten decile portfolios sorted on operating profitability. Panel B shows the trend in past five year and future five year sales growth and earnings per share growth. The trend is calculated by fitting a least squares growth line to the logarithms of the six annual observations from year t-5 to year t. Only observations with at least 5 years of data are kept. The sample excludes financial firms (those with one-digit standard industrial classification code of six). The sample period is from July 1963 to December 2013.

	Profitability Low	2	3	4	5	6	7	8	9	High
Panel A: Return										
Past 1 Year Return	0.055	0.124	0.154	0.159	0.174	0.178	0.198	0.206	0.21	0.252
Past 5 Year Return	0.158	0.503	0.664	0.833	0.94	1.06	1.2	1.386	1.64	2.277
Future 1 Year Return	0.074	0.145	0.168	0.164	0.167	0.17	0.161	0.17	0.168	0.185
Future 5 Year Return	0.828	1.06	1.052	1.058	0.99	1.02	0.99	1	1.04	1.07
Dispersion of Future 1 Year Return	0.95	0.68	0.72	0.63	0.58	0.61	0.56	0.61	0.6	0.74
Dispersion of Future 5 Year Return	3.04	2.68	2.53	2.51	2.17	2.25	2.26	2.69	3.2	3.36
Panel B: Sales and EPS Growth										
Past 5 Year Sales Growth	0.0213	0.0312	0.0348	0.0315	0.0324	0.0322	0.033	0.034	0.034	0.0405
Future 5 Year Sales Growth	0.0269	0.0202	0.0228	0.0221	0.0206	0.0206	0.0221	0.0224	0.0232	0.0247
Dispersion of Future Sales Growth	0.27	0.16	0.14	0.13	0.12	0.12	0.12	0.12	0.12	0.15
Past 5 Year EPS Growth	-0.0184	-0.0121	-0.0052	-0.00028	0.0059	0.0134	0.0153	0.0154	0.02	0.0244
Future 5 Year EPS Growth	0.0306	0.0237	0.0116	0.0032	-0.0004	-0.0023	-0.0029	-0.0059	-0.0082	-0.0105
Dispersion of EPS Growth	0.3	0.25	0.21	0.21	0.22	0.2	0.2	0.2	0.19	0.19

Table 9: Portfolio Double Sorts

This table reports descriptive statistics of portfolio double sorts based on operating profitability and analyst growth forecast. The portfolios returns are value-weighted and sorted based on NYSE breakpoints. It shows the monthly return in percentage, the operating profitability, the median growth forecast, and the forecast errors for each of the 25 double-sorted portfolios. The sample excludes financial firms (those with one-digit standard industrial classification code of six). The period is from July 1975 to December 2013.

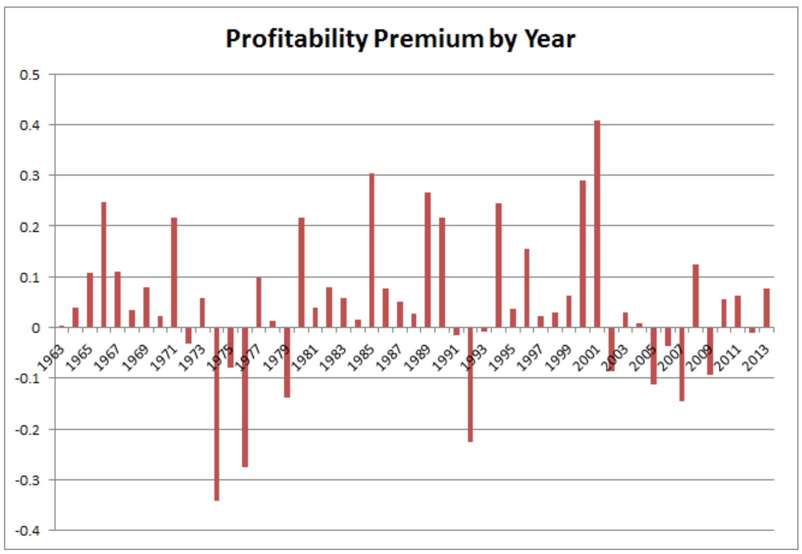
	Profitability					
	1	2	3	4	5	Difference
Median Forecast Growth			**Returns**			
1	0.91	0.97	0.97	1.2	1.2	0.29
2	1.2	1.1	1	1.2	1	-0.2
3	0.98	0.85	1.2	1.1	0.98	0
4	0.87	1.2	0.71	0.98	1	0.13
5	0.19	0.94	0.74	0.99	1.3	1.11
Difference	0.72	0.03	0.23	0.21	-0.1	

	Profitability					
	1	2	3	4	5	Difference
Median Forecast Growth			**Forecast Error**			
1	0.33	0.13	0.14	0.23	0.09	-0.24
2	0.27	0.17	0.12	0.08	0.07	-0.2
3	0.31	0.13	0.11	0.08	0.03	-0.28
4	0.06	0.13	0.07	0.07	0.03	-0.03
5	0.15	0.09	0.07	0.04	0.03	-0.12
Difference	0.18	0.04	0.07	0.19	0.06	

	Profitability				
	1	2	3	4	5
Median Forecast Growth			**Forecast Growth**		
1	0.04	0.05	0.05	0.06	0.05
2	0.1	0.1	0.1	0.1	0.1
3	0.13	0.13	0.13	0.13	0.13
4	0.16	0.16	0.16	0.16	0.16
5	0.34	0.27	0.26	0.25	0.26

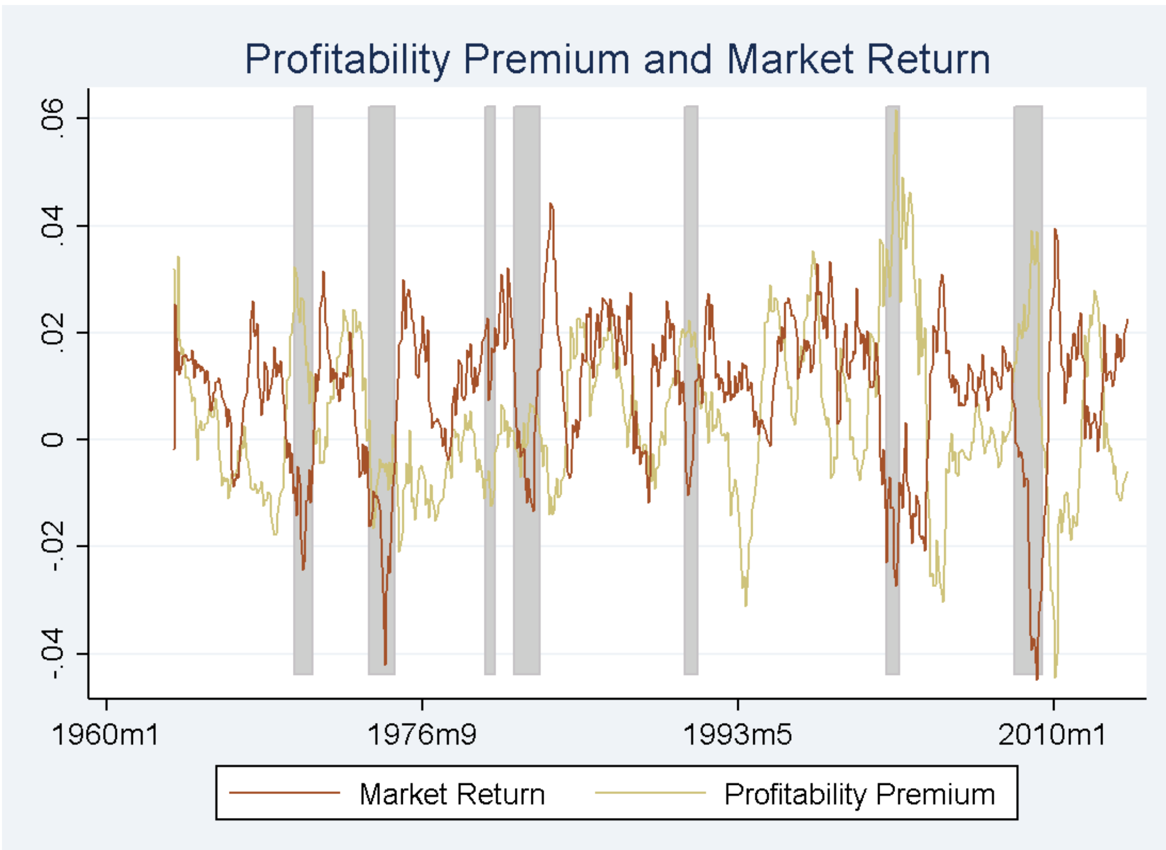
	Operating Profitability				
	1	2	3	4	5
Median Forecast Growth			**Profitability**		
1	0.047	0.114	0.147	0.188	0.293
2	0.043	0.113	0.148	0.188	0.28
3	0.054	0.115	0.149	0.191	0.286
4	0.036	0.114	0.148	0.192	0.3
5	-0.026	0.114	0.148	0.191	0.321

Figure 1: Profitability Premium 1963 to 2013



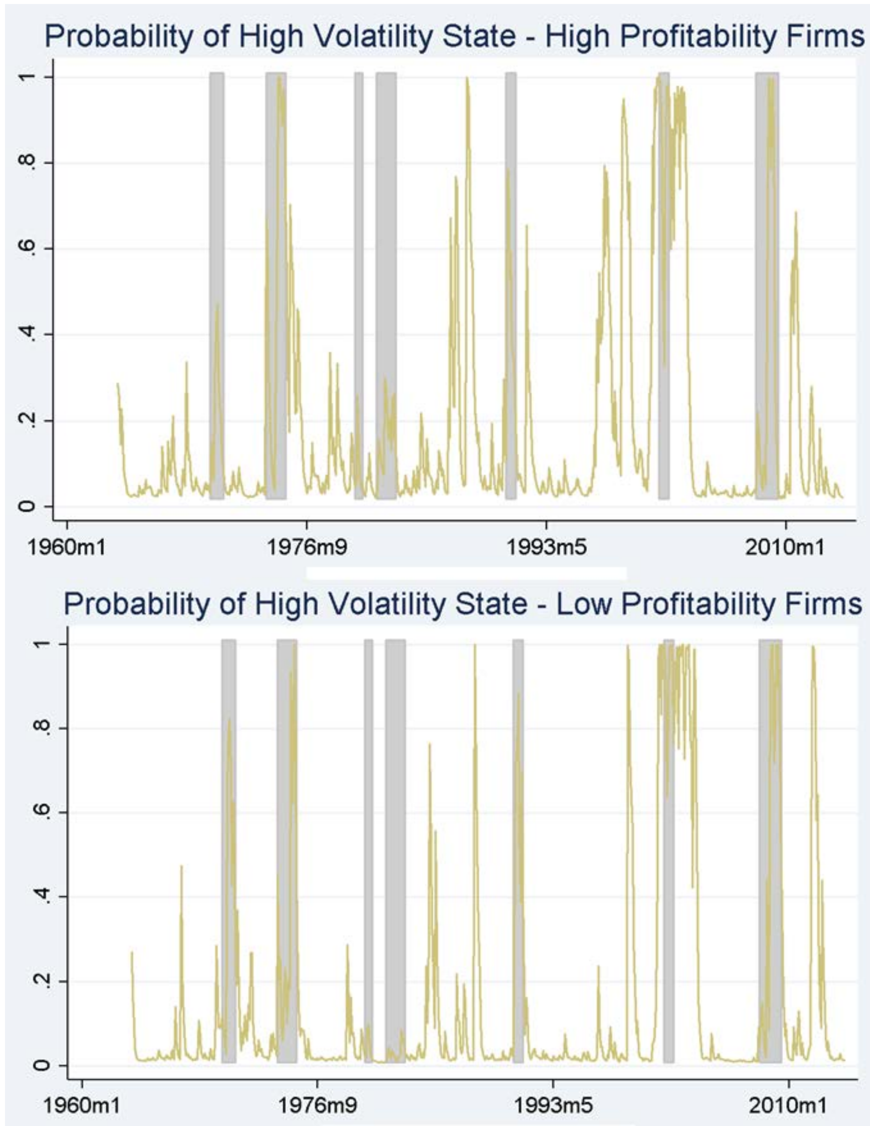
All the companies in the sample are sorted into ten portfolios based on their operating profitability, excluding the financial firms (those with one-digit standard industrial classification code of size). The profitability premium is calculated as the annual return of highest decile profitability portfolio minus the lowest decile profitability portfolio. The sample period is from July 1963 to December 2013, so the first and last year only include six months of return.

Figure 2: Profitability Premium and Market Return



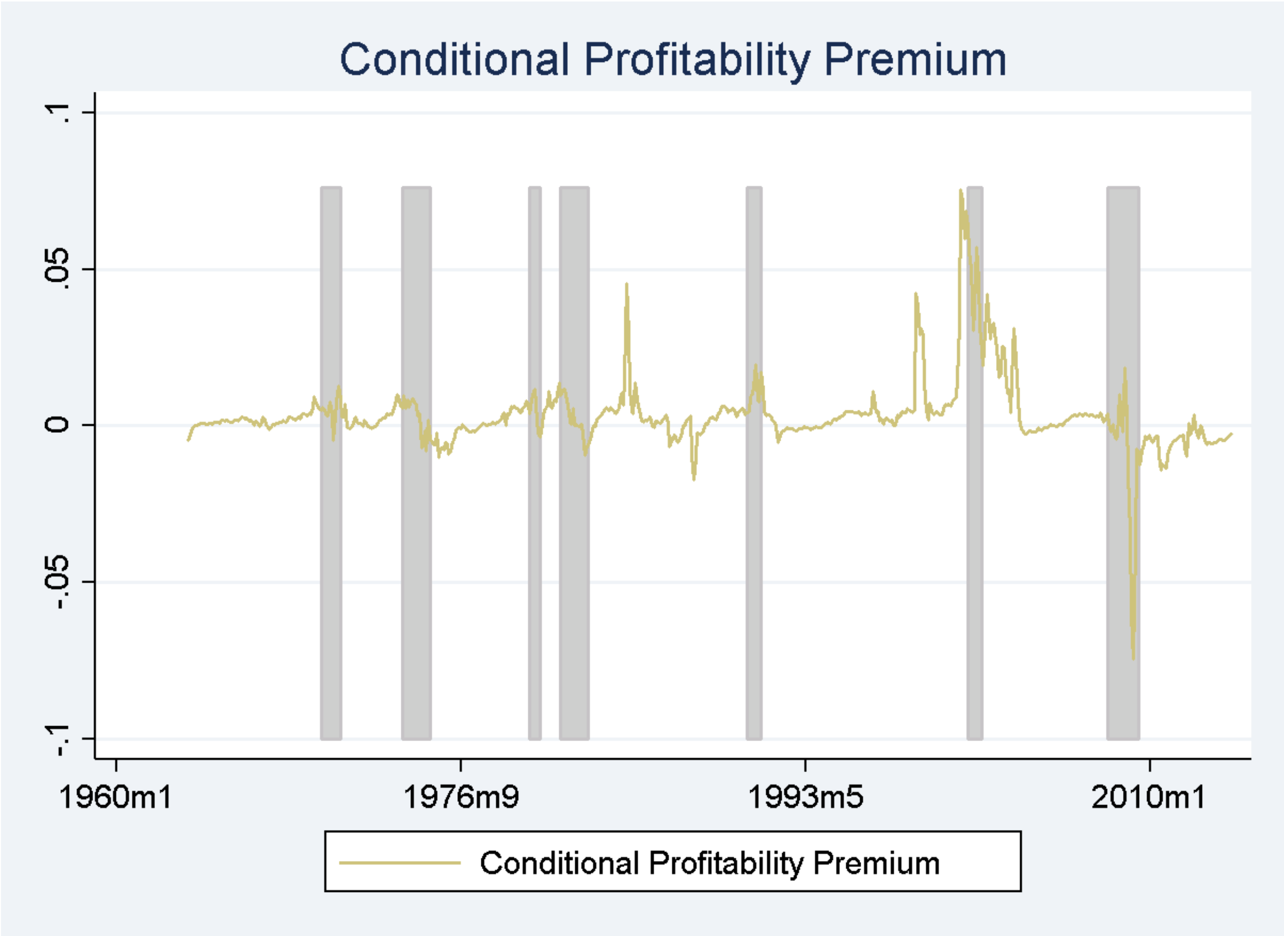
This figure plots the return of highest decile profitability portfolio minus the lowest profitability portfolio. The premium is smoothed over 12 months. For comparison, the figure also plots the 12-month smoothed aggregate market return as given by the CRSP value-weighted index. The NBER recessions are shaded. The sample period is from July 1963 to December 2013.

Figure 3: Probability of High Variance States



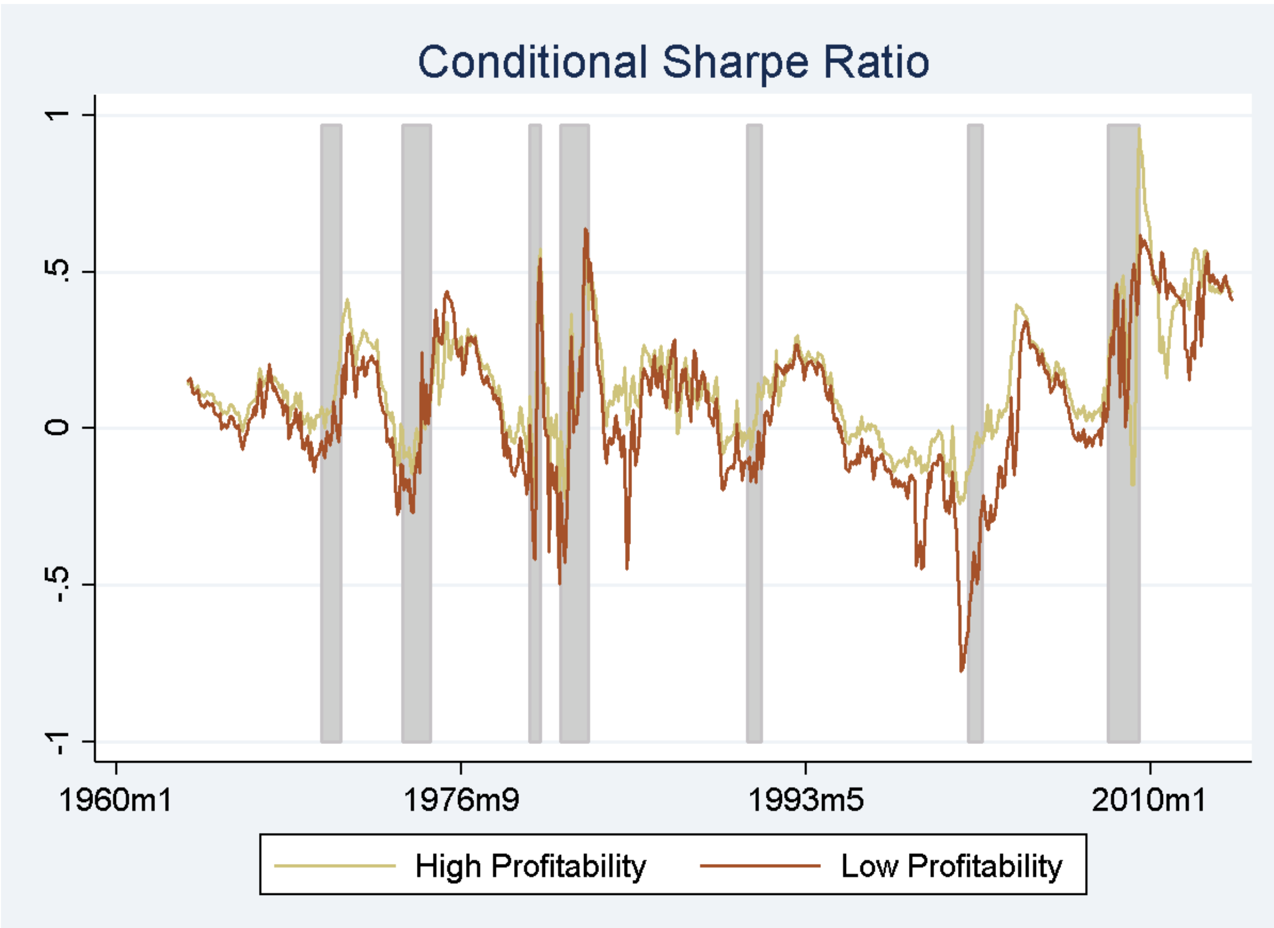
This figure plots the probability of being in a high volatility state as given by the two-state Markov Regime-Switching model. The top plot shows the probability given by the model fitted to the high profitability portfolio and the bottom plot shows the probability given by the model fitted to the low profitability portfolio. The NBER recessions are shaded. The sample period is from July 1963 to December 2013.

Figure 4: Conditional Profitability Premium



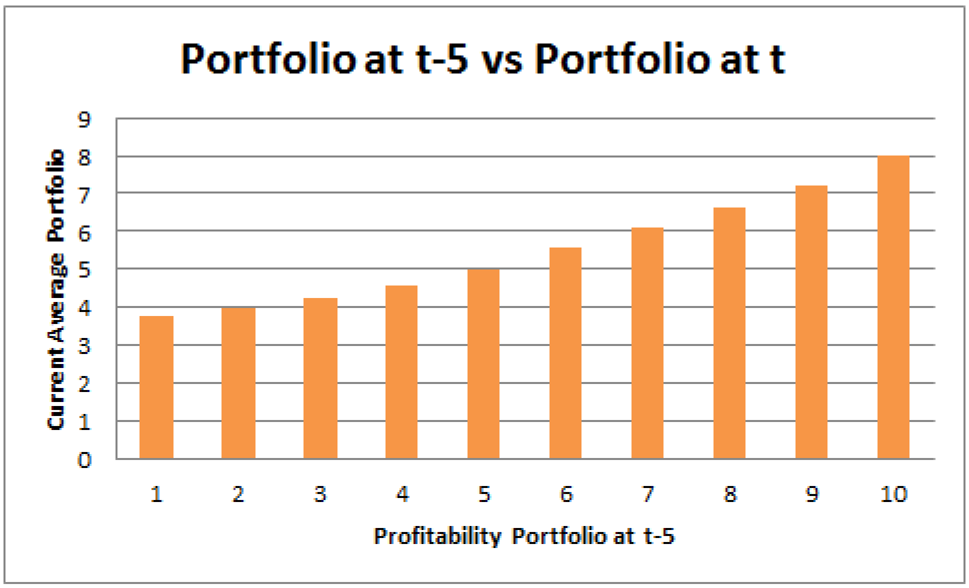
This figure plots the expected profitability premium at time t conditional on information at time t-1, as given by the two-state Markov Regime-Switching model. The premium is calculated as the expected excess return of the top profitability decile minus the expected excess return of the bottom profitability decile. The NBER recessions are shaded. The sample period is from July 1963 to December 2013.

Figure 5: Expected Sharpe Ratio



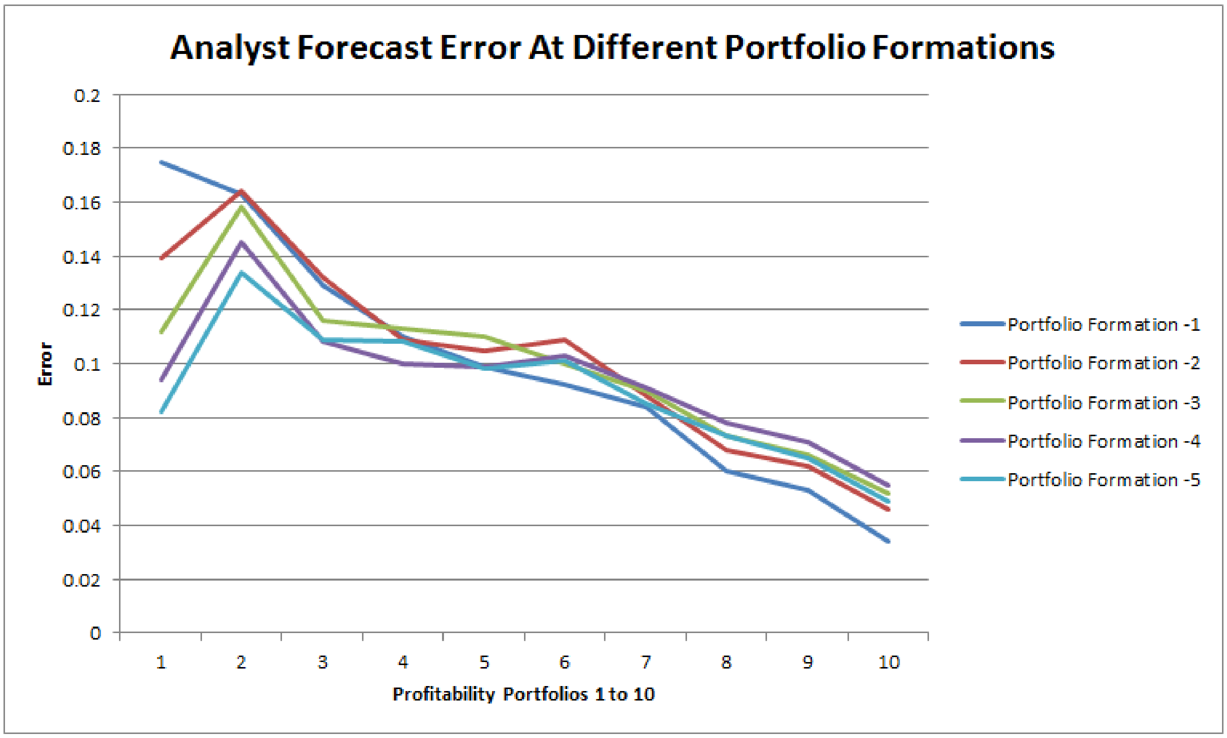
This figure plots the expected Sharpe Ratio at time t conditional on information at time t-1, as given by the two-state Markov Regime-Switching model. The ratio is calculated as the expected excess return fitted by the model divided by the expected volatility fitted by the model. The yellow line shows the conditional Sharpe ratio for the high profitability portfolio while the red line shows the conditional Sharpe Ratio for the low profitability portfolio. The NBER recessions are shaded. The sample period is from July 1963 to December 2013.

Figure 6: Profitability Persistence



This figure shows the persistence of profitability portfolios. The horizontal axis displays the ten decile portfolios of profitability from year t-5. For the firms in each decile from five years ago, the vertical axis displays the average portfolio the firms are in at time t. The sample period is from July 1963 to December 2013.

Figure 7: Expectation Error from t to t-5



This figure shows the persistence of expectation errors. It plots the analyst forecast error in the vertical axis across the ten operating profitability portfolios in the horizontal axis. The error is calculated as difference between the forecast earnings per share and the actual earnings per share reported by IBES, winsorized at the 1% level. The portfolio formation range from year t-1 up to year t-5. The sample period is from July 1975 to December 2013.