One anomaly to explain them all*

Jack Favilukis† and Terry Zhang‡

December 9, 2019

Abstract

We argue that conditional on the existence of momentum, many other asset pricing anomalies are not particularly anomalous. First, empirically, we show that the return to the momentum strategy (i.e. the momentum premium) is persistently higher among some stocks than others, and that the momentum premium of a portfolio negatively predicts the portfolio's unconditional average return. Second, we rationalize this in a standard model to which we add momentum; the intuition is that speculators prefer to buy assets with a higher momentum premium, and bid up the prices of those assets. Third, we find that for many asset pricing anomalies, the momentum premium of the long leg is much lower than the momentum premium of the short leg. Thus, according to our model, the long leg should earn higher unconditional average returns in equilibrium, which "explains" the anomaly. Once accounting for this effect, the average Fama French 3 factor α across 36 prominent anomalies falls by up to 47%. Finally, we show that although the CAPM β is negatively related to the average unconditional return of a large set of portfolios, it is strongly positively related to the portfolios' momentum premia, which helps explain the apparent empirical failure of the CAPM.
JEL classification: G11, G12, G14.

*

†Sauder School of Business, University of British Columbia. Email: jack.favilukis@sauder.ubc.ca
‡Australian National University. Email: terry.zhang@anu.edu.au

1 Introduction

Cross-sectional asset pricing has been described as a zoo (Cochrane (2011), Feng et al. (2019)), with hundreds of proposed factors or anomalies floating around, many perhaps data mined or insignificant out of sample (Harvey et al. (2015), Hou et al. (2018)). We propose a unified explanation for the high excess returns on a large set of anomalies; the explanation relies on the existence of momentum.

Momentum, initially brought to the attention of academics by Levy (1967) and Jegadeesh and Titman (1993), is the empirical observation that firms whose equity returns have been particularly high (low) over the past year or so, continue to outperform (underperform) over the next year or so. Momentum is one of the most well known, most studied, and prevalent asset pricing anomalies. It is present in a variety of asset classes (Asness et al. (2013)) and its performance does not appear to have disappeared after its discovery (McLean and Pontiff (2016)). Various explanations for momentum have been proposed, from rational Johnson (2002), Berk et al. (1999), Hou et al. (2015); to behavioral Barberis et al. (1998), Grinblatt and Han (2005), Frazzini (2006); to institutional Lou (2012), Lou and Polk (2019); to explanations based on liquidity and trading costs (Pastor and Stambaugh (2003), Sadka (2006)). Momentum is often explained as mostly an under-reaction, with some longer term over-reaction.

Our model is closest Hong and Stein (1999), who generate momentum through limited inattention and under-reaction. The main difference is that in our model, some assets under-react by more than others. This leads to differences in the profitability of momentum strategies across different types of assets. While our model endogenously generates momentum, our goal is not to explain momentum, rather it is to show that conditional on its existence, the high observed returns on many anomaly portfolios are to be expected.

Our main empirical finding is described below. We first sort firms into various portfolios based on 36 prominent anomalies identified in the literature, with ten portfolios per anomaly, plus ten industry portfolios for a total of 370 portfolios. For example, firms in the bottom decile of the size distribution, or firms in the third decile of the operating profitability distribution are two such portfolios. Within each portfolio, we compute the return from following a momentum strategy, that is buying past winners and selling past losers only within this portfolio. Momentum profits are systematically different across portfolios. For example, firms in the top idiosyncratic volatility decile persistently earn high momentum returns, while firms in the bottom idiosyncratic volatility decile persistently earn low momentum returns.

Our first key finding is that there is a strong negative relationship between a portfolio's momentum profits, and its average buy-and-hold return. That is, portfolios that tend to systematically offer momentum trading opportunities (MTOs) and have high conditional trading returns, such as high idiosyncratic volatility firms, tend to also have low unconditional or buy-and-hold returns. Figure 2 illustrates this fact.

Our second key finding is that MTOs tend to be present in the short leg, that is the low buy-and-hold return leg, of many anomalies. For example, high market-to-book, low cash flow yield, high investment rate, low operating profitability, high idiosyncratic volatility, and high distress firms all offer great MTOs – a momentum strategy within such firms is very profitable. These are also firms which have low average buy-and-hold returns, as identified by the anomaly literature. Thus, we identify a common feature across many seemingly unrelated anomalies. We argue that when one sorts on some characteristic that the literature has identified as leading to high (anomalous) returns, what one is actually sorting on is MTOs. Controlling for the presence of MTOs, the average long-short anomaly alpha is reduced by between 23% and 47%. Since, as argued by Harvey et al. (2015), McLean and Pontiff (2016), and Hou et al. (2018), many anomaly alphas appear artificially high due to in-sample data mining, our MTO effect potentially explains much more than 47% of the anomaly performance.

Our third key finding is theoretical and underpins the first two. We show that if MTOs are present in a set of firms, then these same firms should have lower average buy-and-hold returns. Thus, through the lens of our model, any anomaly whose short leg offers MTOs should have a positive return spread between its long and short legs. The model's intuition is that active, or speculative traders are more interested in firms which offer active trading opportunities, than those firms that do not. As a result, MTO firms have higher unconditional prices and lower unconditional expected returns, despite offering high returns for traders who follow an active (conditional) strategy.

To provide additional support for this mechanism, we identify times when there are few momentum traders in the market, using the co-momentum measure of Lou and Polk (2019). In these times, consistent with Lou and Polk (2019), momentum profits are highest since there are few traders to arbitrage these profits away. On the other hand, these are times when the buy-and-hold spread between high and low MTO firms is lowest because there are few traders who push up the price of high-MTO assets. Thus, speculative momentum traders are the ones responsible for the buy-and-hold spread between low- and high-MTO assets.

The above findings argue that if any subset of firms, such as high market-to-book firms, offers MTOs, then these firms should have lower expected returns. However, they do not offer an explanation as to why high market-to-book firms should offer MTOs. Our fourth set of results explores which characteristics make a firm more likely to offer MTOs. Our model suggests that information about high MTO firms should be difficult to interpret, making these firms hard to value.[1] Empirically, high MTO firms tend to have less analyst coverage, more analyst disagreement, more volatile fundamentals, more volatile stock returns, and more autocorrelated stock returns. However, their most striking feature is their loading on market risk. As shown in Figure 3, CAPM β almost perfectly lines up with a portfolio's average momentum return. This can potentially explain the failure of the CAPM; indeed, in a regression of a portfolio's average buy-and-hold return on its CAPM β, the slope turns from negative to positive when we control for MTOs.

To our knowledge, we are the first to systematically link an anomaly's excess return to the momentum trading profits of the underlying firms. However, some of our results have been documented in other papers individually. Avramov et al. (2007) and Garlappi and

[1] This is consistent with how much of the literature has thought about momentum. For example Daniel and Titman (1999) write that the "momentum effect is likely to be strongest in those stocks whose valuations require the interpretation of ambiguous information".

Yan (2011) show that momentum profits are stronger for highly distressed firms and firms with low credit rating; the credit distress puzzle is that these firms also have anomalously low returns. Daniel and Titman (1999) and Sagi and Seasholes (2007) show that momentum profits are stronger for growth firms and high growth option firms; the value premium puzzle is that these firms also have anomalously low returns. It has also been shown that momentum profits are stronger in small and low analyst coverage firms (Hong et al. (2000)), high volume firms (Lee and Swaminathan (2000)), and high revenue volatility firms (Sagi and Seasholes (2007)).

Two other closely related papers are Hong and Sraer (2016) and Zhang (2019). Hong and Sraer (2016) argue that disagreement is highest for high beta assets. In the presence of short sale constraints, high disagreement implies that these assets will be overvalued, leading to returns that are too low, which can explain the failure of the CAPM. In our model, it is also the case that speculative behavior leads to prices that are too high, although we focus on speculation through momentum and do not require short sale constraints. In our empirical results, we find that high beta is one of the characteristics associated with high MTO and therefore too low returns. Zhang (2019) studies the portfolio choices and trading behavior of active versus passive institutions. He finds that the firms held by active institutions tend to have low passive buy-and-hold returns, but that the institutions themselves do not earn low returns. This is because such institutions follow active trading strategies and invest in inefficiently priced firms which present trading opportunities. This is analogous to the high MTO, low buy-and-hold firms that we focus on.

Finally, a number of papers have documented momentum profits in well diversified portfolios, as opposed to individual assets. These papers on factor momentum include Grinblatt and Moskowitz (2002), Avramov et al. (2017), Zaremba and Shemer (2018), and Arnott et al. (2019), and Ehsani and Linnainmaa (2019). However, these results are fundamentally different and orthogonal to our study. These studies find that well diversified portfolios that have done well (poorly) in the past tend to continue doing well (poorly). This is an exact analog to a firm level momentum strategy, but at the portfolio level. We study the profitability of a firm level momentum strategy within a well diversified portfolio. Our MTO score measures whether a momentum strategy is profitable within a portfolio, and is unrelated to whether this portfolio has done well in the past. For example, the value portfolio consistently has a low MTO score, however in Avramov et al. (2017) it is sometimes a winner and sometimes a loser, depending on whether it outperformed other anomalies over the previous year.

In the remainder of the paper, we present our empirical results in section 2, the model follows in section 3, and section 4 concludes.

2 Empirical results

2.1 Data

Our empirical tests use data from various standard sources. We obtain information about stock returns, trading volume, market capitalization, and the number of shares outstanding from the Center for Research in Security Prices (CRSP) database. We collect accounting related information from COMPUSTAT. We download returns on the risk-free rate, the Fama-French size, value, and momentum factors from Kenneth French's website.

Our test assets consist of 370 different portfolios of firms. We create these portfolios by sorting firms according to various characteristics. Our purpose is to have a diverse set of portfolios that capture different dimensions of cross-sectional variation in expected returns. Table 1 presents the list of firm characteristics that we consider, and the return spread between the long and short legs of each anomaly. Pooling together this large set of portfolios, we want to assess how much of the cross-sectional variation in the expected returns of these anomalies can be explained by our proposed mechanism.

We choose the sorting characteristics for forming portfolios based on the asset pricing literature's most prevalent anomalies. We start with the anomalies in Table 4 of Hou et al. (2015), which they compute to be the only statistically significant ones in a list of nearly 80 anomalies.[2] We augment these with all anomalies in Stambaugh et al. (2015) and in Fama

[2] There are a total of 35 anomalies on this list. We exclude Analyst Forecast Surprise because it does not span the time period of the remainder of our data. For the other two announcement anomalies (Standard Earning Surprise and Abnormal Announcement Return), Hou et al. (2015) list both the 6 month and 1

and French (2015) that are not already on this list. We also add size, as it is not on any of the above lists. Finally, we add a set of industry portfolios based on the Fama-French 10 industry designation; these are not anomalies, but are a good, independent set of portfolios for testing our explanation. This leads to 37 different characteristics.

For each characteristic, we form portfolios by sorting firms into deciles.[3] This is how we arrive at the total of 370 different portfolios. More specifically, based on each characteristic, we sort firms into portfolios in June of year t from 1966 to 2017. When forming portfolios in year t, we compute accounting characteristics based on the information available in December of year $t-1$. For market-based characteristics that only use information from the CRSP, we measure them in June of year t. Firms in each portfolio are kept constant for the next twelve months with four exceptions: earnings announcement return, consecutive earnings growth, standardized unexpected earnings, and failure probabilities. These anomalies require high rebalancing frequencies. For these anomalies, we re-sort twice a year: June and December. We drop a firm from a portfolio in month t if it is delisted in month $t-1$ or its price drops below \$1 in month $t-1$. Following Fama and French (1992), we form decile portfolios based on NYSE cut-offs.

2.2 Two relevant measures of returns

For each portfolio of firms, we compute two different types of return series. The first type of return is the value-weighted return of a portfolio in excess of risk-free rate from July of year t to June of year $t+1$. We call this return the buy-and-hold return of a portfolio; this buy-and-hold return is the subject of interest in most of the asset pricing literature.

The second type of return is relatively novel to the literature, it measures the profitability

month horizon as separate significant anomalies. We keep only the two 6 months horizon anomalies because we do not want to put too much weight on any single anomaly and these announcement anomalies are highly correlated. We also include only the most standard one (R11-1) of the four different momentum anomalies. This is because our goal is to explain other anomalies conditional on the existence of momentum, not to explain momentum.

[3]For industries, we do not sort into deciles. Rather, each of the 10 industries is a portfolio. For some of the results, to be consistent with the way we report anomalies, we order industries based on their average historical return.

of a momentum strategy *within* each portfolio. Since the momentum strategy is a conditional strategy, this is a conditional, or active, return, rebalanced monthly. To compute this return, we sort the firms in each portfolio into three terciles in the beginning of each month t based on each firm's past performance, which we measure as the cumulative return from month $t-12$ to $t-2$. We then compute the value-weighted return in excess of risk-free rate for firms in the top tercile, which we call the return of winners in a portfolio. Similarly, we compute the value-weighted return in excess of risk-free rate for firms in the bottom tercile, which we call the return of losers in a portfolio. We refer to the firms in the middle tercile as neutral firms. The momentum return of a portfolio is the return of winners minus the return of losers within that portfolio. Our main findings, presented below, is that the momentum return of a portfolio can explain the portfolio's risk-adjusted buy-and-hold return.

2.3 Summary Statistics

As shown in Table 2 Panel A, the average buy-and-hold return of these 370 portfolios is 0.59% per month, with a Sharpe ratio of 0.42. In Panel B, the average return of winners minus losers is 0.56% per month with a Sharpe ratio of 0.34. The difference between the two returns is stark when looking at the CAPM or FF3 alphas, which are, respectively 0.03 and 0.00 for the average portfolio's buy-and-hold return, but 0.63 and 0.78 for the average portfolio's momentum return. This is consistent with the prior literature on momentum, which finds winners have higher returns than losers. Not surprisingly, the Fama-French-Carhart alpha of the momentum return of the average portfolio is close to zero.

Table 2 also shows the cross-sectional variation in the average buy-and-hold return and momentum return. Across the 370 portfolios, there is much more cross-sectional variation in momentum returns than buy-and-hold returns. The the cross-sectional standard deviation of average buy-and-hold returns, CAPM alphas, and FF3 alphas is, respectively, is 0.13%, 0.16%, and 0.16% per month, compared to 0.33%, 0.35%, and 0.36% for the momentum return.

2.4 Portfolio level results

This section documents our main findings. We show that the risk-adjusted buy-and-hold return of a portfolio is negatively related with the portfolio's average momentum return. In other words, if it is especially profitable to implement a conditional momentum strategy for a portfolio of firms, then that same portfolio is likely to have an especially low unconditional or buy-and-hold return. Table 3 and Figure 2 present our main result.

In Panel A of Table 3, we run a cross-sectional regression of a portfolio's average buy-and-hold return on its average momentum return. The relationship is negative and significant, with a t-statistic of -6.15. The slope of -0.14 implies that for every 1% increase in a portfolio's expected momentum return, its expected buy-and-hold return falls by 0.14%.

The results are stronger when we adjust the buy-and-hold return for risk, the t-statistics are between -8.74 and -15.02 for regressions where the dependent variable is the Sharpe ratio, or the CAPM, FF3, or FFC4 alpha. The R^2 is high too, for example the momentum return alone explains 41.6% of the cross-sectional variation in portfolio Sharpe ratios. We redo this with controls (Panel B), and in sub-periods (Panels C and D), with the results largely unchanged. The negative relationship between the momentum return, and various risk adjusted measures of the buy-and-hold return are presented graphically in Figure 2.

For robustness, we also redo the main result but excluding all microcaps (firms in the bottom 20% of the size distribution and those with stock prices below $5), or excluding the financial crisis (July 2008 - June 2011), these results, presented in Appendix Tables A1 and A2 are similar to those in Table 3. This suggests that our results are not driven by illiquid firms or extreme events. Finally, we redo the above analysis but separately for each anomaly. That is, we sort all firms into deciles based on each characteristic, and run a regression of buy-and-hold alphas on momentum returns for each group of 10 portfolios. These results, presented in Appendix Table A4, show that of the 36 different characteristics, the slope is negative for 31, and significantly negative for 25.

A portfolio's momentum profit can explain between 23% and 47% of the within anomaly variation in alpha, this computation is described next. We construct a MTO adjusted alpha in the following way. We first run a cross-sectional regression of a portfolio's average buy-and-hold FF3 alpha on its average momentum return. The slope of this regression is $b^{all} = -0.23$ if we use all 370 portfolios (Panel A of Table 3), and $b^{extr} = -0.41$ if we use only the 74 most extreme portfolios, that is the top and bottom deciles for each anomaly. We then define a MTO adjusted alpha as the residual from this regression: $\alpha^{adj} = \alpha - b \times R^{MOM}$. The spread in the raw and in the adjusted alphas is reported in Table 4 for each anomaly. The average anomaly has an alpha spread of 0.43% per month, this is both a simple average and an average of absolute alphas because the top decile always has a higher return than the bottom. Once adjusted for MTO, the average alpha spread falls from 0.43% to 0.32% (all) or to 0.23% (extreme); the average absolute alpha spread falls from 0.43% to 0.33% (all) or 0.28% (extreme).

2.5 Firm level results

We next carry out a similar analysis, but at the firm level. For each firm, we compute a momentum trading opportunities (MTO) score, which measures whether in the future, the firm is likely to be a good investment for a momentum trader. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. We define portfolios that are ranked in the top 10 percentile as MTO portfolios. Then, for each firm, we define its MTO score as the number of MTO portfolios it belongs to relative to total portfolios it belongs to.

Appendix Table A6 shows the correlations of the MTO score with various firm characteristics. Consistent with the mechanism explained above, the MTO score is positively correlated with high idiosyncratic volatility, high asset growth, high market-to-book, and low gross profit – all characteristics associated with low average returns by various studies in the anomaly literature. The MTO score is also positively correlated with the CAPM beta. This helps explain the failure of the CAPM, since high CAPM beta firms should have high returns, but empirically do not. Our model suggests that having a high MTO score pushes these firms' expected return down, and indeed, as is discussed below, once we control for the

MTO score, the loading of the average return on CAPM beta becomes positive. The MTO score is negatively associated with size, thus it does not help explain the size effect since smaller firms typically have higher returns.

We first sort all firms into ten deciles based on their MTO score. Panel A of Table 5 shows the Fama-French 3 factor alpha, as well as the factor loadings for a momentum trading strategy within each portfolio. As might be expected, the momentum strategy has a Fama-French 3 factor alpha close to zero in low MTO score firms but an alpha of 1.45% per month (4.57 t-stat) in high MTO score firms; the difference between the high and low MTO portfolios is 1.38% per month with a t-statistic of 5.04. Panel B shows the buy-and-hold alpha of the same portfolios. Consistent with the previous set of results, the buy-and-hold alpha has the opposite pattern, with high MTO firms having a negative alpha of -0.69% per month (-6.00 t-stat), compared to a positive 0.20% per month (3.32 t-stat) for low MTO firms; the difference is 0.89% per month with a t-statistic of -6.07.

In Table 6 we run Fama and MacBeth (1973) cross-sectional regressions of a firm's buy-and-hold return on its MTO score. The first column is univariate, showing that MTO scores are negatively related to buy-and-hold returns. The remaining columns show that controlling for MTO reduces the magnitude of the other anomalies. The second and third columns show a regression of buy-and-hold returns on just size, and on size and MTO score; the following two columns show a regression of buy-and-hold returns on just book-to-market, and on book-to-market and MTO score; in the following columns, we do the same for momentum, asset growth, gross profit, idiosyncratic volatility, and CAPM beta. The final two columns show multivariate regressions of buy-and-hold returns on all seven characteristics, and all seven characteristics together with MTO score. The slope on book-to-market falls from 0.32 to 0.23 when the MTO score is added, the slope on asset growth falls from -0.54 to -0.39, the slope on gross profit falls from 0.73 to 0.55, and the slope on idiosyncratic volatility falls from -26.92 to -19.13, thus these anomalies become less strong.

On the other hand, the slope on CAPM beta turns from negative to positive (though insignificant), suggesting that the CAPM works better when controlling for MTO. In a multivariate regression that includes MTOs, the slope on CAPM beta has a positive t-statistic of 2.10. The intuition is similar to the other anomalies. As will be discussed in Section 2.7, high MTO firms tend to also have high CAPM betas. This implies that the average buy-and-hold returns of high beta firms are pushed down relative to what they would be if they did not have MTOs. Once we control for MTOs, the positive relationship between CAPM beta, risk, and return becomes apparent. Section 3.9 shows that the same is true in the model. Because of the strong correlation between CAPM beta and MTOs, a regression of expected return on beta leads to a negative slope, even though without MTOs, the slope is positive and the CAPM works perfectly in the model.

As discussed earlier, the MTO score does not reduce the importance of size, thus, whatever is driving small firms to have high expected returns does not appear to be related to MTOs.[4]. The MTO score also does not reduce the importance of momentum, nor should it since in our model, momentum is the anomaly that drives all others. Appendix Table A7 redoes the same analysis excluding micro stocks, the results do not change significantly, suggesting that the results are not being driven by illiquid firms. Note that at the firm level, MTO is likely a noisy measure of actual momentum trading opportunities. Any noise in the MTO score should bias us towards finding no change in the coefficients on the various anomaly characteristics at the firm level.

2.6 Persistence in a portfolio's momentum return

In this section we show that MTOs are a persistent quantity associated with certain portfolio characteristics. We then show that using this persistence, one could design a profitable trading strategy.[5] Panel A of Table 7 presents a pooled regression of the momentum return of portfolio i in year t on the momentum return of portfolio i in years $t-k$ through $t-1$, where $k \in (1, 3, 5, 10, 15)$. Portfolios with high past momentum returns tend to continue having high momentum returns. The effect becomes stronger as k rises, perhaps because a longer horizon helps average out the year-to-year noise and provides a better signal of

[4]Several papers have come up with rational expectations for the size effect, for example Berk (1995) or Carlson et al. (2004)

[5]Note that we already presented one profitable trading strategy based on firm level MTO score in Table 5.

the portfolio's true MTOs. For comparison, we carry out the same exercise but for the buy-and-hold return in Panel B. There is little evidence of persistence.

Next, we split our sample into two. Using the 1966-1992 data, we compute each portfolio's average momentum return to identify portfolios with the strongest momentum trading opportunities. We then compute the buy-and-hold return, Sharpe ratio, and alphas of each portfolio for the 1992-2018 period. Table 8 shows that there is a strong negative relationship between a portfolio's MTOs in 1966-1992, and its risk adjusted buy-and-hold return in 1992-2018. For example, a 1% higher average momentum return in 1966-1992 is associated with a 0.17 lower Sharpe ratio in 1992-2018. One could form a trading strategy by buying and holding only those portfolios that had low MTOs in the previous period.

Finally, Table 9 presents a similar, but dynamic exercise. In June of each year, we sort all 370 portfolios based on their momentum return over the previous 10 years. We then compute the return of a strategy that buys the lowest 37 MTO portfolios, and sells the highest 37 MTO portfolios. This strategy requires rebalancing every June. The CAPM, FF3, and FFC4 alphas of this strategy are 0.40%, 0.34%, and 0.28% per month, with t-statistics 4.39, 5.45, and 4.54 respectively.

Such a strategy may or may not survive transaction costs. However, our goal is not to identify a strategy that could be exploited by hedge funds. Rather, it is to show that momentum trading creates price pressure, which leads to lower average returns for MTO firms.

2.7 Determinants of a portfolio's momentum return

Up to now, we have argued that any portfolio with high MTOs should have lower expected returns. We have also argued that many of the anomalies identified by the asset pricing literature can be explained, at least in part, by the presence of MTOs in the short leg of the anomaly. However, so far, we have been silent on the fundamental reasons for why a particular firm or characteristic might be associated with MTOs. We explore this question here. Jumping ahead to our model, firms whose information is hard to interpret, or whose fundamentals are not revealed to all market participants at the same time, or over which there is more disagreement have high MTOs. While the exact data analogs to the model are hard to measure, we find that for both fundamentals and stock returns, MTO firms have fewer analysts covering them, more dispersion in analyst forecasts, more volatility of fundamentals and of stock returns, and higher loading on aggregate risk. Since it is reasonable to assume that such firms would be harder to value, we interpret the data to be broadly consistent with the model.

Before we explore the determinants of the MTO score, it is useful to conceptually separate all firms into losers, neutral, and winners. Note that a firm with a high MTO score need not currently be a great investment opportunity. High MTO score firms may be past winners, in which case a momentum trader would want to buy them, or past losers, in which case a momentum trader would want to short them. Alternately, if a high MTO firm had relatively average returns over the past 12 months, it is neither a past winner or past loser, we call such firms neutral firms. A currently neutral MTO firm will become an investment opportunity for a momentum investor in the future, if it receives a series of positive returns. On the other hand, a low MTO firm, whether winner, loser, or neutral, will never be an attractive investment for a momentum investor. Formally, the loser, neutral, and winner firms are defined as being in the bottom, middle, and top terciles of expected returns, conditional on being in the portfolio.

We first consider measures of analyst coverage, presented in Table 10. As we move from low MTO score to high MTO score quintile, the number of analysts covering a firm monotonically declines. On the other hand, the dispersion in analyst forecasts, for both EPS and long term EPS growth, monotonically rises. These results suggest that high MTO firms are harder to value and are subject to more disagreement.

We next consider measures of fundamentals. We sort all firms into quintiles by their MTO score and compute the volatility of quarterly net income growth, sales growth, asset growth, gross profit, ROA, and EPS/Price. Table 11 shows that by any measure, MTO firms have more volatile fundamentals. Appendix Table A8 shows that MTO firms also have stronger loadings on aggregate GDP growth. Thus, based at least on these simple measures, MTO firms appear to have more fundamental risk.

Table 12 is similar to Table 11, but reports the volatility of stock returns rather than fundamentals. This is done both at the portfolio level, where we compute the volatility of the portfolios ranked by MTO, and at the individual firm level, where we use firms' MTO score to sort them into quintiles. As with fundamentals, MTO firms have more volatile stock returns. Another interesting finding is that when firms are ranked by their loser, neutral, or winner status (that is, based on past returns), volatility follows a U-shaped pattern, highest for losers and winners and lowest for neutral firms. This pattern is behind the key mechanism in our model, discussed in the next section. Interestingly, this U-shaped pattern is prominently present only in stock returns, but far less so in fundamentals. Table 13 shows that MTO firms also have far more autocorrelated returns. This is not surprising, since momentum and positive autocorrelation are closely linked.

In Table 14, we regress a portfolio's average momentum return on various portfolio characteristics. The most striking result is the strong loading of the momentum return on the portfolio's market beta, this can be seen visually in Panel A of Figure 3. Thus, assets with MTOs appear risky, at least according to the CAPM. Possibly this implies that those investors that trade these assets pay attention to CAPM risk in setting prices. Sections 2.5 and 3.9 discuss why this positive correlation helps explain the apparent empirical failure of the CAPM. Additionally, MTOs are also positively related to the SMB beta, negatively to the HML beta, and positively to idiosyncratic volatility.

As an additional robustness check, we also redid the main result using post-earnings announcement drift (PEAD) return as the key sorting variables, rather than momentum return. Table A3 in the Appendix shows that these results are largely similar to our main result in Table 3. Assets with stronger PEAD tend to have lower risk-adjusted buy-and-hold returns. This suggests that the key driving force is not necessarily momentum itself, but under-reaction and the possibility for informed investors to profit.

2.8 Relation to co-momentum

Lou and Polk (2019) identify times when there are many momentum traders in the market using a variable they label as co-momentum.[6] We use this as an independent test of our mechanism. As we show below, when there are few momentum traders in the market, the spread in momentum profit between high- and low-MTO assets is very high. This is the key result in Lou and Polk (2019) who argue that in these times, there are too few traders arbitraging away momentum profits. Conversely, when there are few momentum traders in the market, the spread in buy-and-hold returns between low- and high-MTO assets is very low. This is because there are too few momentum traders to push up the price of high-MTO assets. As discussed in section 3.11, this is exactly what happens in our model.

These results are presented in Table 15 and Figure 4. Panel A in both the table and figure present the momentum return of these portfolios. When co-momentum is low (blue line), the spread between high and low MTO portfolios is largest, 1.10% per month, compared to only 0.37% per month when co-momentum is high (green line). Panel B in both the table and the figure show the buy-and-hold CAPM alpha. When co-momentum is low, the spread between high and low MTO portfolios is smallest (in magnitude), -0.17% per month, compared to -0.59% per month when co-momentum is high. Panel C presents a similar result for the Fama and French 3-factor alpha.

3 Model

Our model is similar to Hong and Stein (1999). Like in Hong and Stein (1999), who refer to them as "newswatchers", a fraction of investors observe imprecise and delayed information about asset quality, but do not learn from prices. In Hong and Stein (1999), the remaining investors are "momentum traders" who condition on historical price changes but are only allowed to form simple, univariate strategies of the signal; Hong and Stein (1999) need such

[6] Lou and Polk (2019) define co-momentum as the correlation of stocks within the portfolios used for momentum trading, that is, the correlation within the loser portfolio, and the correlation within the winner portfolio. Their interpretation is that when there are many momentum traders, they buy all winner stocks, increasing their correlations, and sell all loser stocks, decreasing their correlations.

investors to have short term underreaction, and long term overreaction. Unlike Hong and Stein (1999), our remaining investors are fully rational, risk-averse arbitrageurs, therefore there is no long term overreaction in our model. The other important difference between our model and Hong and Stein (1999) is that we have two types of assets which differ in the way their information is revealed to the market. As a result, some assets in our model have momentum trading opportunities (MTOs), and some do not.

3.1 Simple example

Before we describe the full model, it is useful to build intuition with a simple, partial equilibrium example. Consider two assets, each with price 100, and each of which receives the same positive news between $t = 0$ and $t = 1$. The first asset reacts fully to the news, its price rises to 110 at $t = 1$, and then stays at 110 at $t = 2$, because no new information is released. The second asset under-reacts to news, its price rises to 105 at $t = 1$ and then to 110 at $t = 2$. The situation after negative news is analogous, with the price of the first asset falling to 90 and staying there, while the price of the second falling to 95, and then to 90. This is pictured in Figure 1.

Now consider an investor who has a horizon of one period. At $t = 0$, the asset that under-reacts appears more attractive, because it has the same expected return with a lower volatility. Even if the investor has a longer horizon, she may prefer the under-reacting asset because she can sell after negative news at $t = 1$, before realizing the full loss. In general equilibrium, there must be demand for both assets. This implies that the under-reacting asset should have a higher price and a lower return to compensate investors for bearing less risk. The full model described below is dynamic and general equilibrium; assets are long lived and prices clear markets conditional on demand from investors. However the basic intuition remains the same. Investors find assets with momentum trading opportunities more attractive, pushing up their prices and pushing down their returns.

3.2 Assets

There are n infinitely lived, risky assets available in fixed net supply of $1/n$ shares each. There is also a risk free rate asset available in unlimited net supply with a constant interest rate of r^f.

Risky asset i's dividend is $D_t^i = X_t + Y_t^i + Z_t^i$ where X_t is an iid, aggregate component, which is common to all assets; Z_t^i is an iid, asset specific (idiosyncratic) component; and $Y_t^i = \rho Y_{t-1}^i + \epsilon_t^i$ is a persistent, asset specific (idiosyncratic) component.

If all investors know X_t, Y_t^i, and Z_t^i at the time that they are investing, then asset i's ex-dividend price P_t^i may only be a function of Y_t^i, since, conditional on Y_t^i, all assets are identical going forward. More generally, as will be explained below, some investors will have an imprecise estimate of Y_t^i. As a result, the equilibrium ex-dividend price will be a function of more than just Y_t^i, that is $P_t^i = \Psi(Y_t^i, S_t^i)$ where S_t^i is a vector of relevant additional state variables.

We define $Q_{t+1}^i = P_{t+1}^i + D_{t+1}^i$ to be the $t+1$ payout. At t, investors have a belief over Q_{t+1}^i. In particular, they have a belief about the $n \times 1$ vector of expected payouts $\Upsilon_t = E_t[Q_{t+1}]$ and the $n \times n$ variance-covariance matrix Σ_t, whose (i, j) entry is $\Sigma_t^{i,j} = Cov[Q_{t+1}^i, Q_{t+1}^j]$.

A fraction q of assets are 'high information' assets, with the remainder being 'low information' assets. Both high and low information assets have exactly the same dividend process, described above. As will be discussed below, the only difference between the two is how information about each asset's persistent quality Y_t^i becomes known to investors. This will lead to the 'low information' assets having MTOs, while the 'high information' not.

3.3 Investors

There are many overlapping generations of investors and each generation lives for just two periods. At t they receive a constant labor income endowment Π. They do not consume at t but invest their entire endowment in a portfolio of risky assets with portfolio weights w_t^i and a risk free asset with weight $1 - \sum w_t^i$. At $t+1$ they sell their entire portfolio and consume the proceeds of the sale. Thus, their only choice variable is the $n \times 1$ vector of portfolio weights w_t. The investors receive a return $R^p_{t+1} = (1 - \sum w^i_t)(1 + r^f) + \sum w^i_t \frac{Q^i_{t+1}}{P^i_t}$. The investors have mean-variance utility, so that they solve the optimization problem

$$U_t = \max_{w_t} \Pi \times \left(E_t[R^p_{t+1}] - 0.5\gamma V_t[R^p_{t+1}]\right) = \Pi \times \left(1 + r^f + \max_{w_t} \left(w'_t \mu_t - 0.5\gamma w'_t \Omega_t w_t\right)\right) \quad (1)$$

where w_t is an $n \times 1$ vector of risky asset portfolio weights, μ_t is an $n \times 1$ vector of expected excess returns whose ith entry is $\Upsilon^i_t / P^i_t - 1 - r^f$, Ω_t is an $n \times n$ variance-covariance matrix whose (i, j)th entry is $\Sigma^{i,j}_t / (P^i_t P^j_t)$, and γ is approximately equal to relative risk aversion.[7] It is well known that if unconstrained, the portfolio choice of these investors is $w_t = \frac{1}{\gamma} \mu'_t \Omega^{-1}_t$, and their shares demand is $\Pi w_t / P^i_t$ per capita.

There are two types of investors. At time t, informed investors know the Y^i_t of every asset, informed investors are a fraction p of the population.

Uninformed investors are identical to informed investors with one exception. Just like the informed investors, for all high information assets, they know Y^i_t at t. On the other hand, for all low information assets, at t they know Y^i_{t-1} but not Y^i_t.

We assume that both informed and uninformed investors know the true equilibrium pricing function $\Psi(y, s)$. They each use this function, and their belief about Y^i_{t+1} to construct expectations of next period's payout and its variance covariance: $\Upsilon^i_t = E_t[D^i_{t+1} + \Psi(Y^i_{t+1}, S^i_{t+1})]$ and $\Sigma^{i,j}_t = Cov_t[D^i_{t+1} + \Psi(Y^i_{t+1}, S^i_{t+1}), D^j_{t+1} + \Psi(Y^j_{t+1}, S^j_{t+1})]$. Informed investors correctly believe that $Y^i_{t+1} = \rho Y^i_t + \epsilon^i_{t+1}$ and $D^i_{t+1} = X_{t+1} + (\rho Y^i_t + \epsilon^i_{t+1}) + Z^i_{t+1}$. For high information assets, uninformed investors agree. However, for low information assets, uninformed investors do not know Y^i_t and substitute Y^i_{t-1} instead, which they do know. Therefore, they believe that $Y^i_{t+1} = \rho Y^i_{t-1} + \epsilon^i_{t+1}$ and $D^i_{t+1} = X_{t+1} + (\rho Y^i_{t-1} + \epsilon^i_{t+1}) + Z^i_{t+1}$.

Given this belief about Υ_{t+1} and Σ_{t+1}, both investors choose their individual portfolio w_t as a function of observed prices. In particular, they construct the mean and variance covariance of returns as $\mu^i_t = \Upsilon^i_{t+1} / P^i_t$ and $\Omega^{i,j}_t = \Sigma^{i,j} / (P^i_t P^j_t)$. From these they compute their portfolio share $w_t = \frac{1}{\gamma} \mu'_t \Omega^{-1}_t$.

[7] Consider an exponential utility investor facing a portfolio problem with log normal returns: $\max E_t[-e^{-\psi W R_{t+1}}]$. This investor's absolute risk aversion is ψ and her relative risk aversion is approximately $W\psi$. Using Stein's lemma, it can be shown that this problem is equivalent to $\max E_t[R_{t+1}] - 0.5 W \psi V_t[R_{t+1}]$.

Since informed and uninformed investors disagree about the expected payout on some assets, conditional on the same price they will also disagree on the return and will tend to take opposing positions on such assets. Informed investors know all available t information and their behavior is purely rational. In particular, since they know Y_t^i, the price that clears markets is always consistent with their pricing function $P_t^i = \Psi(Y_t^i, S_t^i)$.

Uninformed investors know Y_{t-1}^i but not Y_t^i for low information assets. If we allowed them to learn from prices, as in Grossman and Stiglitz (1980), they could back out Y_t^i from prices, unless we introduced noise traders. Instead, for simplicity, we assume that they do not learn for prices and they 'agree to disagree' with informed investors, for example as in Hong and Stein (1999) and Scheinkman and Xiong (2003). They form expectations about Q_{t+1} by assuming that $Y_t^i = Y_{t-1}^i$, they then choose their portfolio demand w_t similarly to the high information investors.

When $Y_t^i = Y_{t-1}^i$, the price that clears markets is consistent with both the informed and uninformed investors' pricing function. However, when $Y_t^i \neq Y_{t-1}^i$, then the market clearing price is not necessarily equal to the uninformed investors' belief $P_t^i = \Psi(Y_t^i, S_t^i) \neq \Psi(Y_{t-1}^i, S_t^i)$. In this sense, the uninformed investors are irrational. They believe that their pricing function is correct at $t+1$. If the same pricing function is inconsistent with prices at t, they believe that they have identified a good investment opportunity.

While this behavior is clearly suboptimal, we believe that this type of heuristic is reasonable. Furthermore, on average, uninformed investors are correct because, on average $Y_t = Y_{t-1}$. However, after a low information asset has received an unexpected negative (positive) shock to quality, which is observed by the informed but not the uninformed investors, the uninformed investors believe that prices are too low (high). This behavior creates price pressure in the opposite direction of informed investors' trading, and leads to underreaction and momentum, preventing prices from falling (rising) too much after a negative (positive) shock to an asset's quality. Eventually, as the low information's asset quality is revealed, the price further falls (rises) to its fully rational level.

In one version of the model, we also allow for long term investors. These investors own a fraction p^{LT} of the wealth, and they consume their returns each period, thus keeping their wealth constant. They do not rebalance, and own the same portfolio each period. They have a belief about the payout's unconditional mean and variance covariance matrix. Based on this, they form a constant portfolio. Note that all low information assets look identical to the investors, as do all high information assets, therefore the portfolio weight of every low information asset is the same, as is the portfolio weight of every informed asset, but they may not be equal to each other.

3.4 Equilibrium

The equilibrium consists of a pricing function $\Psi(y, s)$ such that given this pricing function, the aggregate demand for each risky asset, whose computation is described above, is equal to the supply of each risky asset $1/n$. Note that the risk free asset is in unlimited supply, therefore we do not require the net bond demand to be zero.[8]

3.5 Calibration

The model is meant to qualitatively illustrate the channel, it is far too simple for its quantitative implications to be taken seriously. Nevertheless, we choose parameters that seem consistent with the data.

The first set of parameters can be directly mapped to the data. The aggregate dividend component X takes one of three values: $(0.88, 1.00, 1.12)$ with equal probability. The asset specific iid component Z takes one of three values: $(-0.1, 0.0, 0.1)$ with equal probability. The asset specific persistent component Y takes one of three values $(-0.5, 0.0, 0.5)$ with a transition probability matrix $(0.88, 0.08, 0.04; 0.06, 0.88, 0.06; 0.04, 0.08, 0.88)$. As explained in the appendix, these values are consistent with the dividends of U.S. industries from Ken French's website. We set the risk free rate to r^f =3% to match the value weighed average price-to-dividend ratio of these industries. We set the net worth to $\Pi = E[D]/r^f = 1/r^f$ so that the net demand for bonds is close to zero, on average.

[8] An alternative model could include the risk free rate as an additional equilibrium quantity, and this risk free rate would clear the bond market. While this is possible, this would require for us to keep track of an additional state variable, which complicates the solution procedure.

We have less guidance from the data for the second set of parameters. In the baseline model we set the number of assets $n = 10$. Because n is relatively small, the asset specific risk is not fully diversified away in the aggregate. We do this because, as documented by Grinblatt and Moskowitz (2002), and more recently Arnott et al. (2019) and Ehsani and Linnainmaa (2019), momentum is prevalent in relatively well diversified portfolios, and once controlling for portfolio momentum, firm level momentum becomes significantly less profitable, but not vice versa.[9] Thus, we interpret each asset as a well diversified portfolio formed on some common characteristics, for example a particular industry. We set the fraction of informed investors $p = 0.50$ and the fraction of high information assets $q = 0.50$. We set risk aversion $\gamma = 30$ in order for the model to have a reasonably high Sharpe ratio; this risk aversion is quite high because there is nothing else in this model to overcome the equity premium puzzle. Below, we explore the sensitivity of our results to n, p, q, and γ.

3.6 Solution Method

We assume that the only additional relevant state variable is $S_t^i = Y_{t-1}^i$. We follow Krussell and Smith (1998) who argue that (Y_{t-1}^i, Y_t^i) is an approximately sufficient state space at t for asset i if, when (Y_{t-1}^i, Y_t^i) is used to predict P_{t+1}^i, the R^2 is sufficiently high.

We start with a guess for the pricing function $P_{t+1}^i = \Psi(Y_t^i, Y_{t+1}^i)$. Conditional on this guess and the law of motion for Y, we compute $\Upsilon(Y_t^i) = E_t[P_{t+1}^i + D_{t+1}^i]$ and $\Sigma(Y_t^i, Y_t^j) = Cov_t[P_{t+1}^i + D_{t+1}^i, P_{t+1}^j + D_{t+1}^j]$ for each i and j and for each investor type. Note that for the low information assets, the uninformed investors use the same function Ψ, but do not know Y_t^i, and substitute Y_{t-1}^i instead.

Once these functions are computed, we simulate the economy for 1000 periods. Each period, we choose a vector of prices P_t^i to clear markets, given the net demand. Note that the portfolio weight of informed investors is $w^I = \frac{1}{\gamma}(\Upsilon(Y_t) * 1_P^{-1})(1_P \Sigma(Y_t)^{-1} 1_P)$ where 1_P is a diagonal matrix with P_i along the diagonal. The portfolio weight of uninformed investors is $w^U = \frac{1}{\gamma}(\Upsilon(\widetilde{Y}_t) * 1_P^{-1})(1_P \Sigma(\widetilde{Y}_t)^{-1} 1_P)$ where $\widetilde{Y}_t^i = Y_t^i$ for high information assets and $\widetilde{Y}_t^i = Y_{t-1}^i$

[9]Another benefit of a smaller n is that the model takes far less time to solve.

for low information assets. The net demand is $\Pi(w^I p + w^U(1-p)1_P^{-1}$. To clear markets, each period we compute the net demand for an initial set of prices, and then increase (decrease) each asset's price if its net demand is higher (lower) than the fixed share supply of $1/n$ per asset.

Using the simulated prices, we compute an updated belief about the pricing function $P_t^i = \Psi(Y_{t-1}^i, Y_t^i)$ and restart the process. We continue until the prices converge. The R^2 of actual prices relative to those predicted by the function Ψ is 0.999997.

3.7 The momentum effect

Our goal is not to explain momentum, rather we argue that if momentum exists, then it is possible to explain other common anomalies. To do so, we build momentum into a standard equilibrium model in the simplest way apparent to us. We conjecture that our main channel would work in most models where momentum is present.

In the baseline model, the momentum strategy of buying assets with an above average past return (winners), and selling assets with a below average past return (losers) leads to an average return of 0.62% per period, with a volatility of 3.53%. For comparison, the aggregate equity premium is 0.28%, with a volatility of 1.44%, and it is slightly negatively correlated with the momentum strategy. The momentum strategy is concentrated among low information assets: the strategy of buying low information winners and selling low information losers leads to an average return of 1.13% with a volatility of 4.80%, doing the same for high information assets leads to zero average return with a volatility of 3.95%. For this reason, we also refer to low information assets as momentum trading opportunity (MTO) assets. This result is similar to the data, shown in Panel A of Table 5, where momentum is concentrated among high MTO firms, and is virtually non-existent among low MTO firms.

In the model, momentum is due to underreaction. Consider a low information asset in the model which just received a positive quality shock (increase in Y). Informed investors observe this increase and correctly interpret that this asset's present value of future dividends increases. Uninformed investors do not observe the increase and incorrectly believe that the present value of future dividends is unchanged. Since uninformed investors cannot learn from prices, the market clearing price is a weighted average of the two investors' belief, which is above the previous price, but below the correct present value of future dividends. The following period the uninformed investors learn the true value of Y and the price rises again. Thus, there is underreaction, continuation, and momentum. The same intuition works after a negative quality shock, thus, there is continuation after both positive and negative price movements. While the momentum effect is stronger when there are limits to arbitrage (n is low or idiosyncratic volatility is high), it never fully goes away even as $n \to \infty$.[10]

3.8 The long-short effect

Our key result is that the average return on high information, low MTO assets is higher than the average return on low information, MTO assets. As will be explained below, this is due to excess demand for low information assets by speculative traders.

Recall that the momentum strategy is profitable only among low information, high MTO assets. This includes MTO winners, MTO losers, as well as MTO neutral assets, which had neither particularly high or low returns in the last period. The momentum strategy leads to zero profits among high information assets or low MTO assets.

A strategy of buying a portfolio of low MTO (high information) assets and selling a

[10] Consider a problem with n uncorrelated assets, where both investor types agree that today's price and tomorrow's price are both P and that the variance of the payout $P + D$ is σ. However, they disagree on the expected dividend: the informed investors believe it is D for all assets, while the uninformed, whose proportion in the population is $1 - p$, believe it is $D + \epsilon$ for all assets. The demand by the informed and uninformed investors is, respectively, $w_F = \frac{\frac{D}{P} - r^f}{\gamma\sigma/P^2}$ and $w_S = \frac{\frac{D+\epsilon}{P} - r^f}{\gamma\sigma/P^2}$. The net demand is $\frac{\frac{D+\epsilon(1-q)}{P} - r^f}{\gamma\sigma/P^2}$ and must equal to the supply $1/n$. Rearranging and solving for P:

$$P(\epsilon) = \frac{D + (1-p)\epsilon}{r^f} \left(\frac{1 + \sqrt{1 - \frac{4r^f\gamma\sigma/n}{(D+(1-p)\epsilon)^2}}}{2} \right) \tag{2}$$

Next, consider what happens as n rises or σ falls. Conditional on ϵ, the price rises as lower risk makes this a safer investment. On the other hand, the difference between the high and low ϵ price, $P(\epsilon) - P(0)$ becomes smaller in magnitude. Thus, lower risk weakens the price effect of disagreement. However, the price effect never fully goes away: $\lim_{n\to\infty} P = \frac{D+(1-p)\epsilon}{r^f}$. Therefore, even with perfect diversification, the price is a weighted average of both the informed and the uninformed investors' beliefs, even if the uninformed investor is wrong.

portfolio of high MTO (low information) assets leads to an average return of 0.35%, with a volatility of 2.80% and a slightly negative correlation with the aggregate return. Low MTO assets have a higher return despite having a lower volatility of 1.85%, compared to 2.16% for a portfolio of MTO assets. Recall that buying a portfolio of momentum assets, which includes winners, losers, and neutral assets, is not what is usually referred to as a momentum strategy, which buys winners and sells losers. In fact, in the model, the time series correlation of a momentum strategy within the MTO assets, and the unconditional portfolio of MTO assets is negative.

The above results are similar to the data. Tables 3, 5, and 6 all show that low-MTO firms have higher average returns than high-MTO firms. Table 12 shows that high MTO firms have more volatile returns.

To understand the intuition, suppose, counterfactually, that both hihg MTO (low information) and low MTO (high information) assets have the same unconditional expected return. The informed and uninformed investors agree that low MTO (high information) assets are fairly priced, as are the MTO (low information) neutral assets. However they disagree about the high MTO (low information) winners and losers.

Consider the investment decision of an informed investor. She believes that high MTO (low information) winners are underpriced, and high MTO (low information) losers are overpriced. Therefore, the informed investor will first of all, wish to take long positions in high MTO winners, and short positions in high MTO losers. Conversely, the uninformed investor believes that high MTO winners are overpriced and that high MTO losers are underpriced. The uninformed investor will first of all, wish to take long positions in high MTO losers, and short positions in high MTO winners. Since both investors are endowed with the same amount of wealth, these positions offset and do not lead to additional price pressure for either high MTO (low information) or low MTO (high information) assets.

However, both investors are subject to limits to arbitrage, in particular, both are risk averse and cannot perfectly diversify the asset specific risk because n is finite. Once their risk capacity for mispriced assets is exhausted, they must choose the next best assets for their portfolio. If the expected return on high MTO (low information) and low MTO (high information) assets is the same, then the next best assets are the high MTO neutral (low information) assets. There are two closely related ways to understand this intuition. First, high MTO neutral assets have the same expected return but lower volatility than the low MTO (high information) assets. The individual volatility of high MTO neutral assets is 3.9%, compared to 9.1% for high MTO winners and losers, and 4.1% for low MTO assets. Why do they have lower volatility? Because momentum is a return continuation and an underreaction. Whether a high MTO neutral asset receives a positive or a negative shock next period, its price will not move by as much as that of a low MTO asset which receives the same shock. Second, high MTO neutral assets provide good future trading opportunities for investors. If the asset receives a positive quality shock, the asset will be underpriced in the future and the investor will continue to hold it, but if it receives a negative quality shock, the investor will have a chance to get out before the price falls all the way to fundamental value. Low MTO assets do not present such opportunities.

Since both low and high information investors have demand for MTO neutral assets, there will be price pressure on such assets. The price of high MTO (low information) assets will rise, and their expected return will fall, relative to low MTO (high information) assets. This is our key result.

3.9 A more realistic model: matching the volatility pattern and the failure of CAPM

Table 12 shows that in the data, as in our baseline model, high MTO assets have more volatile stock returns than low MTO assets, and the volatility pattern within high MTO assets follows a U-shaped pattern, highest for winners and losers, and lowest for neutral assets. However, one unrealistic feature of the baseline model is that the stock return volatility of MTO neutral assets is lower than that of low MTO assets in the model; Table 12 shows that in the data, high MTO neutral assets have a higher stock return volatility than low MTO assets. We believe that the reason for this difference is that, as shown in Table 11, high MTO assets have higher fundamental volatility in the data, while in the baseline model, fundamental volatility is identical for all assets.

We extend the baseline model in one of three ways, all leading to higher fundamental volatility of high MTO (low information) assets. We allow high MTO assets to have a higher volatility of Y, a higher volatility of Z, or higher loading on X. All three make high MTO assets fundamentally riskier, increasing their unconditional return and reducing the unconditional return spread between low MTO and MTO assets. However, for each of the three cases, it is not difficult to find parameter values such that, just as in the data, high MTO neutral assets both have a higher volatility of returns than low MTO assets, and a lower unconditional average return.

This extension also allows us to explicitly showcase why the CAPM fails. Consider the model where high MTO assets have a higher loading on X, thus there is a perfect correlation between β and MTO; Figure 3 shows that in the data this correlation is imperfect but quite high. We set the loadings of dividends on X to be 0.5 for high information (low MTO) assets and 4.0 for low information (hihg MTO) assets. This leads to equity return betas of approximately 0.65 for low MTO and 1.38 for MTO assets. However, due to speculative price pressure, the expected excess returns are higher for low MTO assets: 0.46% compared to 0.32% per period. This implies that the slope of the security market line (SML) is negative. Consider an alternative model where the assets still differ in their loadings on X, however all assets are high information assets ($q = 0$), so there is no momentum. In such a model, the CAPM works perfectly, with the slope of the SML positive and exactly equal to the average equity premium. Section 2.5 and Table 6 explain that in the data, controlling for MTO flips the slope of the SML from negative to positive.

3.10 Parameter sensitivity

As long as there is momentum in our model, and there is heterogeneity in its strength across assets, the key result of higher expected returns on high information (low MTO) assets is present for every parameter combination we have tried. For momentum to exist, Y must be persistent and both types of investors must exist $0 < p < 1$. For heterogeneity, both types of assets must exist $0 < q < 1$. However, depending on the parameters, it may be quantitatively weaker or stronger.

The effect is stronger if Y is more persistent or more volatile, this is because the momentum effect itself is stronger, leading to more disagreement, a stronger underreaction, and a stronger advantage for MTO neutral assets. The effect is stronger when Z is more volatile, or when we introduce a quadratic asset holding cost,[11] this is because limits to arbitrage are stronger. The effect is weaker when we increase the number of assets n or decrease risk aversion γ, this is because limits to arbitrage are weaker. However, in all of these cases, the magnitude of the effect remained roughly the same as the aggregate equity premium. The effect is weaker if we decrease the fraction of low information investors because the momentum effect is weaker. Finally, the effect is weaker if we introduce long horizon investors, who base their holdings on an unconditional belief about the return properties of low and high information assets. This is because long horizon investors cannot and do not time the market, therefore, the high returns of low MTO (high information) assets look very attractive, and this pushes up the price and pushes down the return of these assets.

3.11 Relation to co-momentum

We solve a version of the model where $p = 0.05$, that is there are far fewer informed investors. This corresponds to times when few momentum traders are participating in markets, as measured by low co-momentum in Lou and Polk (2019). With fewer momentum traders, the momentum return is much higher. For all firms, it is 1.59% compared to 0.62% in the baseline model; for high-MTO firms it is 3.01% compared to 1.13% in the baseline model. Consistent with the empirical findings in Lou and Polk (2019), with fewer momentum traders, these momentum profits are being left on the table. However, our key variable of interest, the difference in average returns on high information, non-MTO assets compared to low information, MTO assets is smaller than in the baseline model: 0.24% compared to 0.35%. This is because with few momentum traders, there is little price pressure to push the price

[11] A quadratic holding cost $\lambda \sum w_i^2$ where w_i is each asset's weight in the portfolio, is equivalent to an increase of each asset's volatility by λ.

of MTO assets up. As explained in section 2.8, this is exactly what we see in the data.

3.12 Discussion

In this section, we discuss some of the links between our model and the real world, as well as some alternative explanations.

Our model implies that in order for the price pressure channel to exist, the horizon of the momentum traders must be shorter than the horizon of momentum profits. Indeed, as discussed above, adding long horizon investors to the model reduces the return difference between low and high MTO assets. However, as long as some traders with short horizons exist, the channel will not disappear. Since the horizon of momentum profit appears to be between 6 and 18 months in the real world, it is reasonable to assume that there is a significant number of traders with horizons shorter than this. Another feature of our model is that the uninformed traders do not know the most recent fundamentals about assets Y_t, but do know which assets are low versus high MTO, and understand that high MTO assets under-react. This is not a necessary assumption. Even if uninformed traders did not understand that there are MTO differences across assets and simply allocated randomly between neutral high MTO and low MTO assets, there would still be price pressure on high MTO assets from the informed investors.

Our model is likely not the only one that can match the empirical observations in Section 2. For example, suppose that noise traders preferred certain types of assets, and noise trader demand growth was positively autocorrelated. Then assets preferred by noise traders would exhibit momentum, and their prices would be relatively too high, leading to low average returns. However, this explanation is quite mechanical. The disposition effect has been linked to momentum and is another potential explanation. For example, we could model momentum driven by the disposition effect, as in Grinblatt and Han (2005), rather than by slow release of information, as in Hong and Stein (1999). We conjecture that the difference in returns between high and low MTO firms would exist in such a model because any investors with non-prospect theory preferences would bid up high MTO assets just as they do in our model.

4 Conclusion

We document a new empirical finding. We define a portfolio as having high (low) momentum trading opportunities (MTOs) if a momentum strategy on the portfolio's underlying stocks leads to relatively high (low) profits. We show that portfolios with high MTOs tend to have low buy-and-hold returns. We solve a model with momentum traders which rationalizes this finding, the intuition being that momentum traders push up prices and push down expected returns of MTO assets. We then argue that this finding can explain up to 50% of the returns on anomaly portfolios identified by the financial economics literature. This is because for many anomalies, the short (low buy-and-hold return) leg tends to have high MTOs, and the long (high buy-and-hold return) leg tends to have low MTOs.

References

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang, 2006, The cross-section of volatility and expected returns, *The Journal of Finance* 61, 259–299.

Arnott, Robert, Mark Clements, Vitali Kalesnik, and Juhani Linnainmaa, 2019, Factor momentum, *Working paper* .

Asness, Clifford, Tobias Moskowitz, and Lasse Petersen, 2013, Value and momentum everywhere, *Journal of Finance* 68, 929–985.

Avramov, Doron, Si Cheng, Amnon Schreiber, and Koby Shemer, 2017, Scaling up market anomalies, *Journal of Investing* 26, 89–105.

Avramov, Doron, Tarun Chordia, Gergana Jostova, and Alexander Philipov, 2007, Momentum and credit rating, *Journal of Finance* 62, 2503–2520.

Barberis, Nicholas, Andrei Shleifer, and Robert Vishny, 1998, A model of investor sentiment, *Journal of Financial Economics* 49, 307–343.

Barth, Mary E, John A Elliott, and Mark W Finn, 1999, Market rewards associated with patterns of increasing earnings, *Journal of Accounting Research* 37, 387–413.

Berk, Jonathan, 1995, A critique of size-related anomalies, *Review of Financial Studies* 8, 275–286.

Berk, Jonathan, Richard Green, and Vasant Naik, 1999, Optimal investment, growth options, and security returns, *Journal of Finane* 54, 1553–1607.

Campbell, John Y, Jens Hilscher, and Jan Szilagyi, 2008, In search of distress risk, *The Journal of Finance* 63, 2899–2939.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finane* 59, 2577–2603.

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

Cooper, Michael J, Huseyin Gulen, and Michael J Schill, 2008, Asset growth and the cross-section of stock returns, *The Journal of Finance* 63, 1609–1651.

Daniel, Kent, and Sheridan Titman, 1999, Market efficiency in an irrational world, *Financial Analysts Journal* 55, 28–40.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *The Journal of Finance* 61, 1605–1643.

Davis, James L, Eugene F Fama, and Kenneth R French, 2000, Characteristics, covariances, and average returns: 1929 to 1997, *The Journal of Finance* 55, 389–406.

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

Ehsani, Sina, and Juhani Linnainmaa, 2019, Factor momentum and the momentum factor, *Working paper* .

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

Fama, Eugene, and Kenneth French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427–465.

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

Fama, Eugene, and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–36.

Fama, Eugene F, and Kenneth R French, 2008, Dissecting anomalies, *The Journal of Finance* 63, 1653–1678.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2019, Taming the factor zoo: a test of new factors, *Working paper* .

Foster, George, Chris Olsen, and Terry Shevlin, 1984, Earnings releases, anomalies, and the behavior of security returns, *Accounting Review* 574–603.

Frazzini, Andrea, 2006, The disposition effect and underreaction to news, *Journal of Finance* 61, 2017–2046.

Garlappi, Lorenzo, and Hong Yan, 2011, Financial distress and the cross-section of equity returns, *Journal of Finance* 66.

Green, Jeremiah, John RM Hand, and X Frank Zhang, 2013, The supraview of return predictive signals, *Review of Accounting Studies* 18, 692–730.

Grinblatt, Mark, and Bing Han, 2005, Prospect theory, mental accounting, and momentum, *Journal of Financial Economics* 78, 311–339.

Grinblatt, Mark, and Tobias Moskowitz, 2002, Do industries explain momentum, *Journal of Finance* 54, 1249–1290.

Grossman, Sanford, and Joseph Stiglitz, 1980, On the impossibility of informationally efficient markets, *American Economic Review* 70, 393–408.

Hafzalla, Nader, Russell Lundholm, and E Matthew Van Winkle, 2011, Percent accruals, *The Accounting Review* 86, 209–236.

Harvey, Campbell, Yan Liu, and Caroline Zhu, 2015, ... and the corss-section of expected returns, */revuew if /fubabcuak /stydues* 29, 5–68.

Hirshleifer, David, Kewei Hou, Siew Hong Teoh, and Yinglei Zhang, 2004, Do investors overvalue firms with bloated balance sheets?, *Journal of Accounting and Economics* 38, 297–331.

Hong, Harrison, Terence Lim, and Jeremy Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265–295.

Hong, Harrison, and David Sraer, 2016, Speculative betas, *Journal of Finance* 71, 2095–2144.

Hong, Harrison, and Jeremy Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143–2184.

Hou, Kewei, Chen Xue, and Lu Zhang, 2015, Digesting anomalies: an investment approach, *Review of Financial Studies* 28, 650–705.

Hou, Kewei, Chen Xue, and Lu Zhang, 2018, Replicating anomalies, *Review of Financial Studies* forthcoming.

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

Johnson, Timothy, 2002, Rational momentum effects, *Journal of Finance* 57, 585–608.

Lee, Charles, and Bhaskaran Swaminathan, 2000, Price momentum and trading volume, *Journal of Finance* 55, 2017–2069.

Levy, Robert, 1967, Relative strength as a criterion for investment selection, *Journal of Finance* 22, 595–610.

Lou, Dong, 2012, A flow-based explanation for return predictability, *Review of Financial Studies* 25, 3457–3489.

Lou, Dong, and Christopher Polk, 2019, Comomentum: Inferring arbitrage activity from return correlations, *Working paper* .

McLean, R. David, and Jeffrey Pontiff, 2016, Does academic research destroy stock return predictability, *Journal of Finance* 71, 5–32.

Novy-Marx, Robert, 2010, Operating leverage, *Review of Finance* 15, 103–134.

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

Ohlson, James A, 1980, Financial ratios and the probabilistic prediction of bankruptcy, *Journal of accounting research* 109–131.

Pastor, Lubos, and Robert Stambaugh, 2003, Liquidity risk and expected stock returns, *Journal of Political Economy* 111, 642–685.

Richardson, Scott A, Richard G Sloan, Mark T Soliman, and Irem Tuna, 2005, Accrual reliability, earnings persistence and stock prices, *Journal of accounting and economics* 39, 437–485.

Sadka, Ronnie, 2006, Momentum and post-earnings-announcement drift anomalies: the role of liquidity risk, *Journal of Financial Economics* 80, 309–349.

Sagi, Jacob, and Mark Seasholes, 2007, Firm specific attributes and the cross-section of momentum, *Journal of Financial Economics* 84, 389–434.

Scheinkman, Jose, and Wei Xiong, 2003, Overconfidence and speculative bubbles, *Journal of Political Economy* 111, 1183–1219.

Sloan, Richard G, 1996, Do stock prices fully reflect information in accruals and cash flows about future earnings?, *Accounting review* 289–315.

Stambaugh, Robert, Jianfeng Yu, and Yu Yuan, 2015, Arbitrage asymmetry and the idiosyncratic volatility puzzle, *Journal of Finance* 70, 1903–1948.

Thomas, Jacob K, and Huai Zhang, 2002, Inventory changes and future returns, *Review of Accounting Studies* 7, 163–187.

Titman, Sheridan, KC John Wei, and Feixue Xie, 2004, Capital investments and stock returns, *Journal of financial and Quantitative Analysis* 39, 677–700.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

Zaremba, Adam, and Jacob Shemer, 2018, Is there momentum in factor premia? evidence from international equity markets, *Research in International Business and Finance* 46, 120–130.

Zhang, Terry, 2019, Trading opportunities and the portfolio choices of institutional investors, *Working paper* .

Figure 1: Simple model example

This figure plots a hypothetical price path for an asset that appropriately reacts compared to an asset that under-reacts with a one period delay.

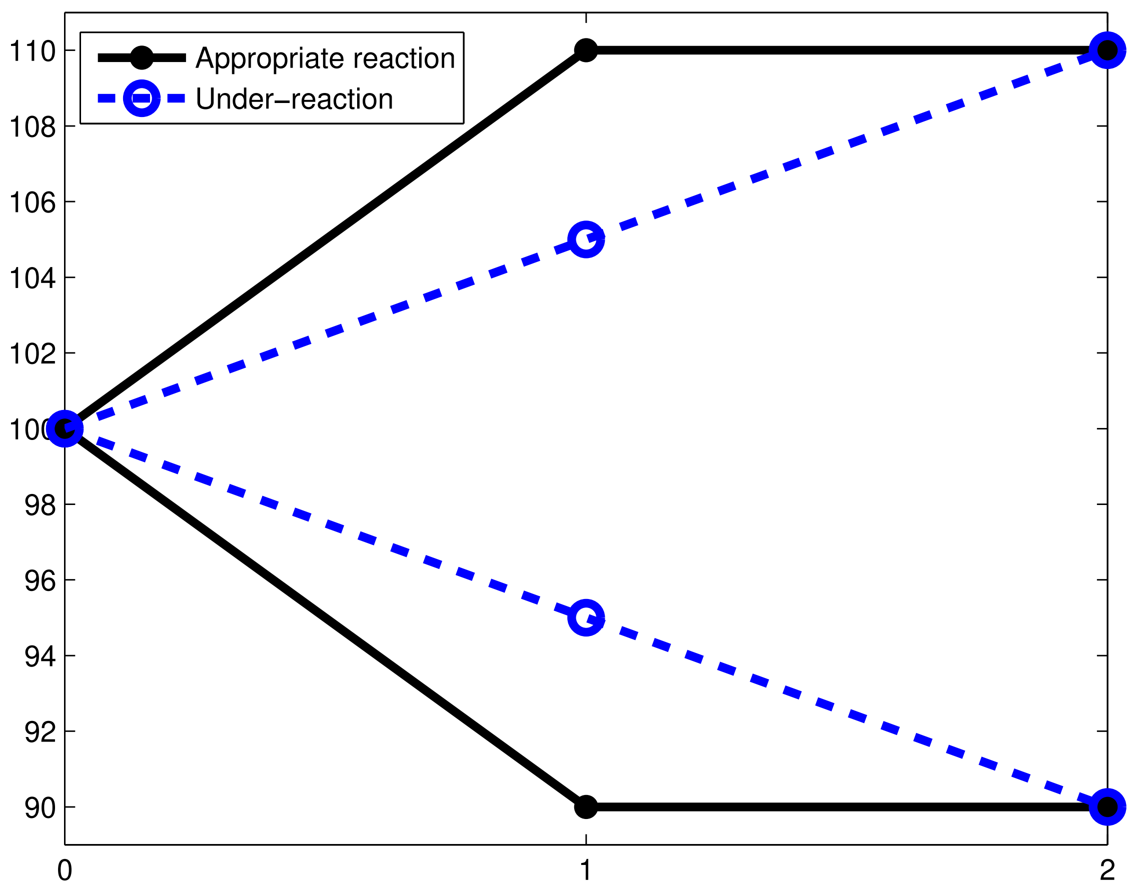

Figure 2: Buy-and-Hold return versus Momentum return

This figure plots the relationship between buy-and-hold return and momentum return of a portfolio. Each point in the scatter plot is a portfolio. Panel A shows the Sharpe ratio of buy-and-hold return vs. the Sharpe ratio of momentum return. Panels B, C, and D show, respectively, the CAPM, FF3, and FFC4 alphas vs. the average momentum return.

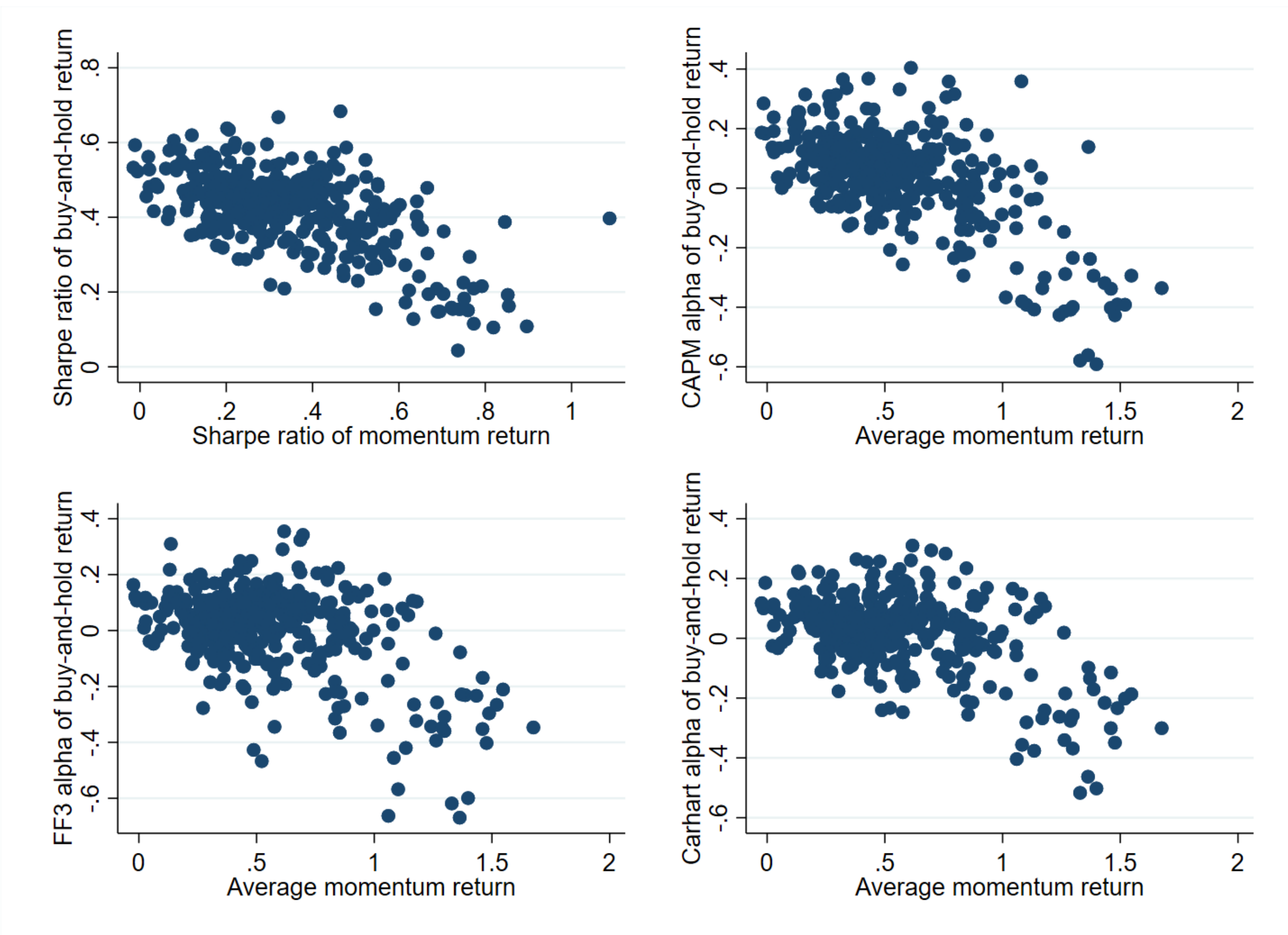

Figure 3: CAPM beta and expected return

This figure plots the relationship between average momentum return and CAPM beta in Panel A and the relationship between average buy-and-hold return and CAPM beta in Panel B. The CAPM beta is measured as the market factor loading of the buy-and-hold return of each portfolio in both panels. There are 370 observations on both panels, corresponding to 370 portfolios. The sample period is from July 1966 to June 2018.

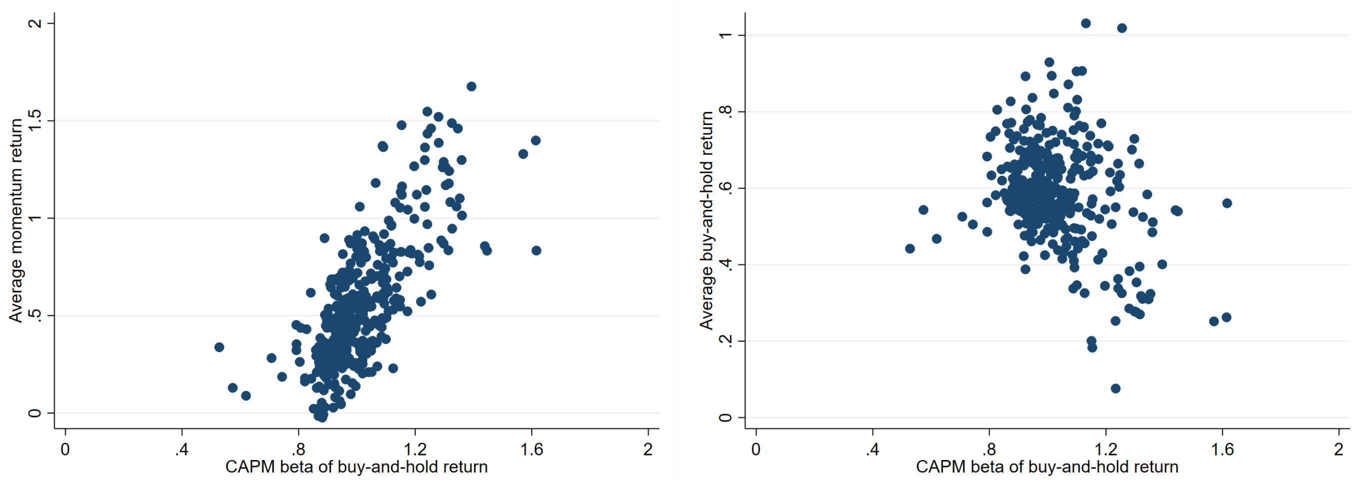

Figure 4: CAPM beta and expected return

This figure plots the key quantities in the paper split by comomentum tercile. We sort our 370 portfolios into ten groups, based on their past momentum return (MTO). We also classify each year as low, mid, or high co-momentum based on the June measure of co-momentum for each year, where co-momentum is defined as in Lou and Polk (2019). We compute the equal weighted average momentum return, and buy-and-hold return of each of the ten portfolio groups, from July of the year, to June of the following year, conditional on that year being low, mid, or high co-momentum. In Panel A, we plot the average momentum return of the ten groups for each sub-sample period. Analogously, in Panel B we plot the CAPM alpha, and in Panel C the FF3 alpha.

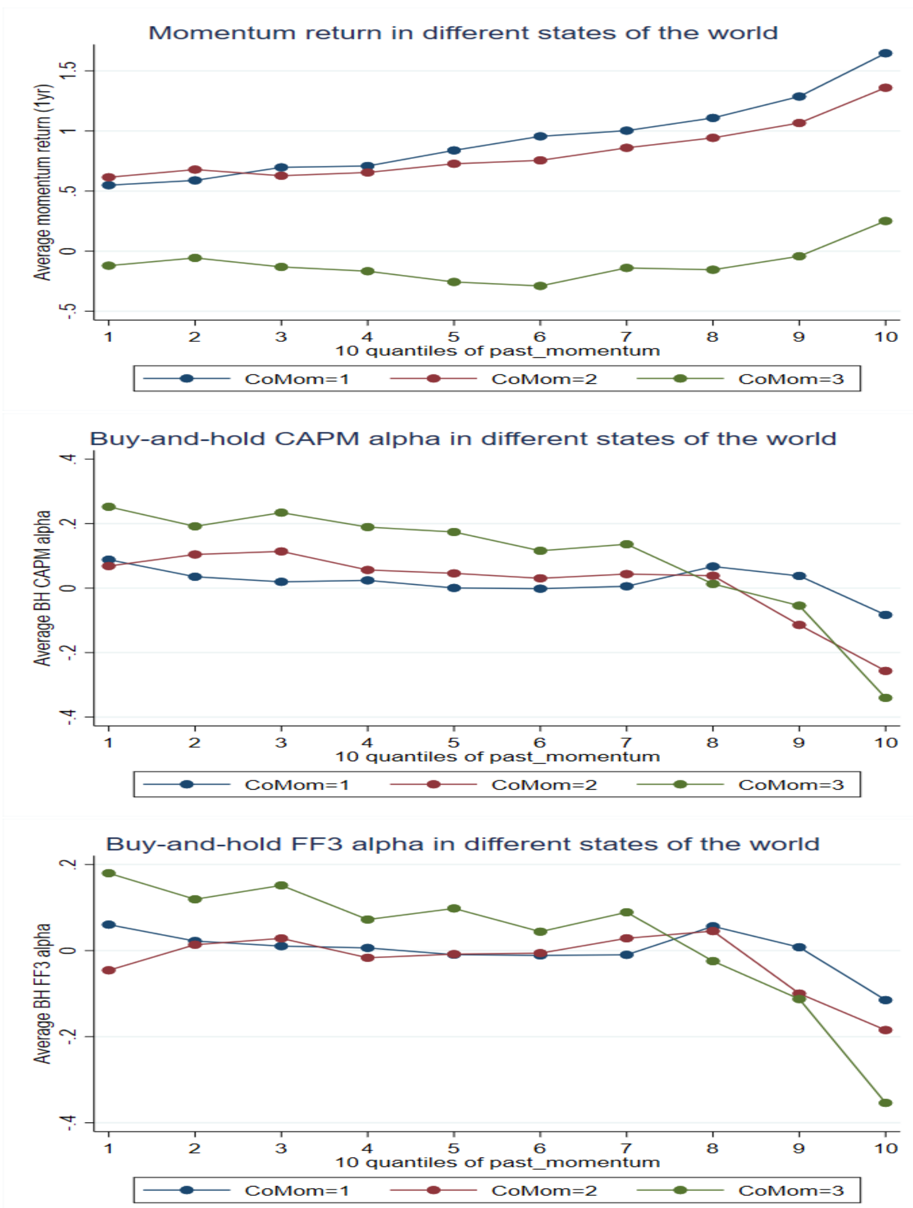

Table 1: Test Assets

This table lists the characteristics that we use to sort firms into portfolios, as well as the buy-and-hold alpha of the top minus bottom decile. Whenever possible, we use NYSE cut-offs to form portfolios. We form portfolios in June of each year t and keep the portfolio composition constant (except for delisting and penny stocks) from July of year t to June of year $t+1$.

Anomaly	Portfolios	α^{CAPM}	t-stat	α^{FF3}	t-stat	α^{FF4}	t-stat
Size	10	-0.15	(-0.82)	0.11	(1.64)	0.13	(1.94)
Book-to-market	10	0.42	(2.30)	-0.20	(-2.04)	-0.19	(-1.92)
Composite issuance	10	-0.77	(-6.65)	-0.61	(-6.08)	-0.51	(-5.05)
Momentum	10	0.40	(1.79)	0.71	(3.44)	0.23	(1.25)
Net issuance	10	-0.81	(-7.27)	-0.71	(-6.68)	-0.64	(-5.92)
CAPM beta	10	-0.53	(-2.67)	-0.45	(-2.70)	-0.24	(-1.42)
Idiosyncratic volatility	10	-0.69	(-3.07)	-0.69	(-4.76)	-0.56	(-3.82)
Total volatility	10	-0.73	(-3.20)	-0.65	(-3.97)	-0.56	(-3.37)
Mispricing score	10	-0.82	(-5.94)	-0.98	(-7.52)	-0.68	(-5.76)
Earnings yield	10	0.76	(4.00)	0.42	(2.63)	0.35	(2.15)
Cash flow yield	10	0.66	(3.48)	0.30	(1.89)	0.23	(1.41)
Net payout yield	10	0.73	(4.48)	0.46	(3.67)	0.43	(3.34)
Ad. expense to market	10	0.57	(2.68)	0.07	(0.44)	0.18	(1.03)
ROA	10	0.37	(2.43)	0.55	(4.24)	0.46	(3.51)
Gross profitability	10	0.37	(2.45)	0.71	(5.84)	0.55	(4.59)
ROE	10	0.28	(1.68)	0.46	(3.38)	0.40	(2.85)
Asset growth	10	-0.57	(-4.47)	-0.30	(-2.76)	-0.21	(-1.89)
Investment to asset	10	-0.54	(-4.58)	-0.35	(-3.20)	-0.29	(-2.68)
Investment growth	10	-0.43	(-4.05)	-0.27	(-2.74)	-0.18	(-1.86)
Inventory change	10	-0.61	(-4.96)	-0.44	(-3.77)	-0.35	(-3.02)
Inventory growth	10	-0.47	(-4.19)	-0.31	(-2.91)	-0.21	(-1.98)
% operating accrual	10	-0.41	(-3.44)	-0.40	(-3.31)	-0.30	(-2.48)
Operating accrual	10	-0.37	(-3.11)	-0.40	(-3.33)	-0.36	(-2.95)
% total accrual	10	-0.43	(-3.51)	-0.24	(-2.22)	-0.27	(-2.50)
Net operating asset	10	-0.52	(-4.55)	-0.47	(-4.18)	-0.43	(-3.77)
Operating leverage	10	0.30	(1.96)	0.49	(3.59)	0.45	(3.20)
Change in P&I	10	-0.51	(-4.13)	-0.27	(-2.53)	-0.24	(-2.14)
O-score	10	-0.34	(-2.08)	-0.50	(-3.75)	-0.45	(-3.30)
Duration	10	-0.50	(-3.04)	-0.04	(-0.35)	0.01	(0.12)
Org. capital to asset	10	0.51	(4.35)	0.44	(3.75)	0.31	(2.67)
R&D expense to market	10	0.08	(0.33)	0.04	(0.18)	0.05	(0.22)
Fama French 10 industry	10	0.47	(2.40)	0.69	(3.63)	0.45	(2.39)
Failure probability	10	-0.40	(-1.59)	-0.99	(-5.37)	-0.62	(-3.71)
SUE	10	0.26	(2.18)	0.35	(2.99)	0.16	(1.37)
Announcement return	10	0.14	(1.12)	0.22	(1.74)	0.02	(0.17)
Consecutive earning growth	10	0.05	(0.40)	0.26	(2.36)	0.12	(1.08)
Systematic volatility	10	-0.41	(-2.08)	-0.41	(-2.05)	-0.19	(-1.01)

Table 2: Summary Statistics

This table reports various summary stats for the stock returns on the 370 portfolios in our sample. We form portfolios in June of each year t and keep the portfolio composition constant (except for delisting and penny stocks) from July of year t to June of year $t+1$. The buy-and-hold return is the value-weighted return of holding all firms in a portfolio and rebalancing each July when the portfolio is formed. The momentum return is the return on a long-short portfolio of past winners minus past losers within each portfolio, which is rebalanced every month. Within each portfolio, past winners are portfolio firms in the top 1/3 of the cumulative return distribution from $t-12$ to $t-2$ for that portfolio, past losers are firms in the bottom 1/3.

Portfolio	Portfolios	mean	p50	sd	min	p25	p75	max
	A: Buy-and-Hold return							
Average	370	0.59	0.58	0.13	0.08	0.53	0.66	1.03
Sharpe ratio	370	0.42	0.43	0.10	0.04	0.37	0.48	0.68
CAPM alpha	370	0.03	0.06	0.16	-0.59	-0.03	0.12	0.40
FF3 alpha	370	0.00	0.03	0.16	-0.67	-0.05	0.10	0.35
Carhart alpha	370	0.02	0.04	0.13	-0.52	-0.03	0.09	0.31
Kurtosis	370	-0.41	-0.41	0.14	-0.91	-0.48	-0.32	0.29
Skewness	370	5.04	4.96	0.62	4.00	4.67	5.32	9.63
Idiosyncratic volatility	370	1.75	1.65	0.45	0.48	1.48	1.90	4.07
	B: Momentum return							
Average	370	0.56	0.49	0.33	-0.02	0.32	0.73	1.68
Sharpe ratio	370	0.34	0.32	0.18	-0.02	0.21	0.44	1.09
CAPM alpha	370	0.63	0.57	0.35	0.00	0.38	0.81	1.75
FF3 alpha	370	0.78	0.71	0.36	0.11	0.53	0.98	1.98
Carhart alpha	370	-0.02	-0.07	0.31	-0.68	-0.25	0.16	0.93
Kurtosis	370	-0.54	-0.52	0.35	-2.06	-0.75	-0.31	0.41
Skewness	370	7.62	7.15	2.41	4.22	5.90	8.53	21.91
Idiosyncratic volatility	370	5.40	5.36	0.68	2.99	4.96	5.85	7.44

Table 3: Momentum profits and Buy-and-Hold returns

This table reports the cross-sectional relationship between a portfolio's average buy-and-hold return (with and without adjusting for risk) and average momentum return. Each observation corresponds to a portfolio. The y-variables are related to the buy-and-hold return of a portfolio, they are: the average buy-and-hold return, its Sharpe ratio, its CAPM alpha, its FF3 alpha, and its FFC4 alpha. The x-variable is the average return of a momentum strategy within a portfolio. Panel A presents the baseline result. Panel B presents the same result, but including the portfolio return's kurtosis, skewness, and idiosyncratic volatility as controls. Panels C and D present sub-sample analysis, for 1966-1992 and 1992-2018 respectively. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}
	A: Baseline result					B: Full sample with controls				
R^{MOM}	-0.14	-0.20	-0.31	-0.23	-0.20	-0.14	-0.18	-0.27	-0.19	-0.18
	(-6.15)	(-15.02)	(-12.33)	(-8.74)	(-8.86)	(-5.55)	(-13.12)	(-11.15)	(-7.84)	(-8.03)
Skewness						0.13	0.08	0.25	0.20	0.15
						(1.82)	(1.68)	(3.20)	(2.90)	(2.84)
Kurtosis						0.03	0.02	0.04	-0.03	-0.02
						(1.65)	(1.57)	(1.62)	(-1.67)	(-1.25)
Ivol						0.01	-0.03	-0.05	-0.07	-0.03
						(0.58)	(-2.57)	(-2.32)	(-3.07)	(-1.72)
Obs.	370	370	370	370	370	370	370	370	370	370
Adj. R^2	0.135	0.416	0.398	0.235	0.250	0.178	0.434	0.444	0.305	0.288
	C: 1966-1992 sub-sample					D: 1992-2018 sub-sample				
R^{MOM}	-0.13	-0.13	-0.21	-0.12	-0.10	-0.09	-0.20	-0.32	-0.27	-0.26
	(-5.24)	(-9.31)	(-8.64)	(-4.39)	(-4.29)	(-3.28)	(-10.24)	(-9.43)	(-8.39)	(-9.51)
Obs.	370	370	370	370	370	370	370	370	370	370
Adj. R^2	0.096	0.231	0.232	0.083	0.081	0.043	0.267	0.294	0.218	0.262

Table 4: Adjusting anomaly alphas for momentum trading opportunities

The first four columns of this table, for each anomaly, show the average momentum return R^{MOM} and the buy-and-hold FF3 alpha $\alpha^{FF3}_{B\&H}$ of the two most extreme decile portfolios (10 and 1). The next two columns present the difference or spread between the momentum return of the top and the bottom deciles, and the buy-and-hold FF3 alpha spread between top and bottom decile. The last two columns present the adjusted for MTOs alpha spread. The adjustment is done in the following way. We regress a portfolio's buy-and-hold FF3 alpha on its momentum return, either for just the 74 extreme portfolios (Extreme) or for all 370 portfolios (All). The slopes of these regressions are, respectively $b = -0.41$ and $b = -0.23$. We then define the adjusted alpha as $\alpha^{adj} = \alpha - b \times R^{MOM}$.

	Low B&H decile		High B&H decile		Low - High spread		Adjusted spread	
	R^{MOM}	$\alpha^{FF3}_{B\&H}$	R^{MOM}	$\alpha^{FF3}_{B\&H}$	R^{MOM}	$\alpha^{FF3}_{B\&H}$	Extreme	All
Operating leverage	0.87	-0.27	0.85	0.22	-0.03	0.49	0.48	0.49
SUE	0.36	-0.17	0.80	0.18	0.44	0.35	0.54	0.46
Ad. expense to market	0.83	-0.05	1.08	0.02	0.25	0.07	0.18	0.13
Mispricing score	1.36	-0.67	0.14	0.31	-1.23	0.98	0.48	0.70
Book-to-market	0.76	-0.09	1.18	0.10	0.42	0.20	0.37	0.29
CAPM beta	0.83	-0.31	0.13	0.14	-0.70	0.45	0.17	0.29
Cash flow yield	1.49	-0.30	0.42	0.00	-1.06	0.30	-0.14	0.05
Composite issuance	1.30	-0.36	0.43	0.25	-0.87	0.61	0.25	0.41
Earning announcement return	1.06	-0.18	0.87	0.04	-0.19	0.22	0.15	0.18
Duration	1.46	-0.17	0.77	-0.13	-0.69	0.04	-0.24	-0.12
Earnings yield	1.46	-0.35	0.77	0.07	-0.69	0.42	0.14	0.26
Failure probability	1.06	-0.66	0.69	0.32	-0.37	0.99	0.83	0.90
Gross profitability	0.85	-0.37	0.70	0.34	-0.16	0.71	0.64	0.67
Investment growth	1.24	-0.34	0.81	-0.07	-0.44	0.27	0.09	0.17
Fama French 10 industry	0.52	-0.47	0.45	0.22	-0.07	0.69	0.66	0.67
Asset growth	1.52	-0.27	0.67	0.04	-0.85	0.30	-0.05	0.10
Investment to asset	1.43	-0.23	0.74	0.11	-0.69	0.35	0.06	0.19
Inventory change	1.39	-0.23	0.80	0.21	-0.59	0.44	0.20	0.30
Inventory growth	1.27	-0.26	0.72	0.05	-0.55	0.31	0.08	0.18
Idiosyncratic volatility	1.33	-0.62	0.19	0.07	-1.14	0.69	0.22	0.43
Size	1.36	-0.08	0.24	0.03	-1.12	0.11	-0.35	-0.15
Momentum	1.10	-0.57	0.97	0.14	-0.13	0.71	0.66	0.68
Consecutive earning growth	0.23	-0.12	0.89	0.11	0.66	0.23	0.51	0.39
Net payout yield	1.26	-0.39	0.29	0.07	-0.97	0.46	0.06	0.24
Net operating asset	1.48	-0.40	0.64	0.07	-0.83	0.47	0.13	0.28
Operating accrual	1.29	-0.35	1.14	0.05	-0.14	0.40	0.34	0.37
Percentage operating accrual	1.08	-0.46	0.58	-0.06	-0.50	0.40	0.19	0.28
Org. capital to asset	1.37	-0.23	0.69	0.21	-0.68	0.44	0.16	0.28
O-score	1.68	-0.35	0.43	0.15	-1.24	0.50	-0.01	0.22
Change in P&I	1.55	-0.21	0.72	0.06	-0.83	0.27	-0.06	0.08
R&D expense to market	0.73	0.08	0.61	0.13	-0.12	0.05	0.00	0.02
ROA	1.18	-0.32	0.68	0.23	-0.50	0.55	0.34	0.43
ROE	1.30	-0.31	0.88	0.16	-0.42	0.46	0.29	0.37
Net issuance	1.13	-0.42	0.61	0.29	-0.52	0.71	0.50	0.59
Systematic volatility	1.01	-0.34	0.99	0.07	-0.03	0.41	0.40	0.40
Percentage total accrual	1.17	-0.26	0.97	-0.03	-0.20	0.24	0.15	0.19
Total volatility	1.40	-0.60	0.09	0.05	-1.31	0.65	0.11	0.35
Avg.						0.43	0.23	0.32
Avg. Abs.						0.43	0.28	0.33

Table 5: Momentum profits and Buy-and-Hold returns at firm level: Portfolio sort

This table reports the abnormal return of firms sorted into deciles based on their MTO score. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. Panel A reports the FF3 alpha and the corresponding factor loadings of the momentum returns of these deciles, where momentum return is computed similarly as previous tables as the difference between winner return and loser return within a decile. Panel B reports the FF3 alpha and the corresponding factor loadings of the buy-and-hold returns of these deciles. The sample period of returns is from July 1976 to June 2018. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	L	2	3	4	5	6	7	8	9	H	H-L
	A: Momentum return, Fama French 3 factor model										
α	0.07	-0.05	0.17	0.58	0.73	0.63	0.98	1.24	1.08	1.45	1.38
	(0.41)	(-0.27)	(0.91)	(2.92)	(3.35)	(2.78)	(4.15)	(4.55)	(3.84)	(4.57)	(5.04)
β^{MKT}	-0.04	-0.04	-0.02	-0.20	-0.17	-0.10	-0.14	-0.24	-0.19	-0.23	-0.19
	(-1.00)	(-0.91)	(-0.42)	(-4.32)	(-3.37)	(-1.77)	(-2.49)	(-3.65)	(-2.78)	(-3.06)	(-2.90)
β^{HML}	-0.32	-0.22	-0.21	-0.46	-0.49	-0.30	-0.42	-0.61	-0.37	-0.50	-0.18
	(-4.78)	(-3.32)	(-3.11)	(-6.39)	(-6.15)	(-3.59)	(-4.89)	(-6.11)	(-3.59)	(-4.27)	(-1.79)
β^{SMB}	-0.03	0.12	0.07	0.13	0.03	-0.04	0.18	0.22	0.01	-0.03	0.00
	(-0.48)	(1.94)	(1.12)	(1.88)	(0.46)	(-0.46)	(2.22)	(2.31)	(0.10)	(-0.27)	(0.00)
Obs.	504	504	504	504	504	504	504	504	504	504	504
Adj. R^2	0.039	0.030	0.020	0.095	0.074	0.021	0.060	0.089	0.027	0.038	0.013
	B: Buy-and-Hold return, Fama French 3 factor model										
α	0.20	0.21	0.19	0.03	-0.08	0.03	0.00	-0.12	-0.17	-0.69	-0.89
	(3.32)	(3.30)	(2.63)	(0.42)	(-1.03)	(0.44)	(0.01)	(-1.52)	(-1.93)	(-6.00)	(-6.07)
β^{MKT}	0.91	0.89	0.90	0.93	1.04	1.02	1.05	1.14	1.16	1.34	0.43
	(62.74)	(59.86)	(53.09)	(51.38)	(58.37)	(56.32)	(65.34)	(58.42)	(54.96)	(48.93)	(12.31)
β^{HML}	0.23	0.19	0.15	0.19	0.21	0.07	0.12	0.00	-0.18	-0.45	-0.68
	(10.41)	(8.24)	(5.81)	(7.02)	(7.76)	(2.59)	(4.80)	(0.15)	(-5.59)	(-10.72)	(-12.70)
β^{SMB}	-0.20	-0.17	-0.14	-0.15	-0.15	-0.03	0.08	0.22	0.35	0.57	0.77
	(-9.53)	(-7.76)	(-5.60)	(-5.58)	(-5.87)	(-1.26)	(3.42)	(7.84)	(11.46)	(14.47)	(15.27)
Obs.	504	504	504	504	504	504	504	504	504	504	504
Adj. R^2	0.887	0.879	0.852	0.842	0.874	0.872	0.904	0.891	0.891	0.884	0.644

Table 6: Momentum profits and Buy-and-Hold returns at firm level: Fama-MacBeth

This table reports the relationship between the stock return and the MTO score at the firm level. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. We then run Fama-MacBeth regressions of stock return on its MTO score and other control variables. All t-statistics, shown in parentheses, are based on Fama-MacBeth standard errors. The number of observations is between 2,333,382 and 1,765,803; it is omitted to conserve space.

	Monthly return																
MTO	-2.49		-2.81		-2.43		-2.43		-2.06		-2.53		-1.52		-2.12		-1.47
	(-3.92)		(-4.44)		(-3.53)		(-4.05)		(-2.85)		(-3.55)		(-3.60)		(-3.65)		(-4.02)
Size		-0.02	-0.09													-0.15	-0.18
		(-0.55)	(-3.14)													(-5.99)	(-7.39)
BM				0.32	0.23											0.16	0.13
				(4.02)	(3.19)											(2.76)	(2.26)
Momntm.						0.64	0.58									0.54	0.54
						(4.47)	(4.34)									(5.04)	(4.96)
Asset gr.								-0.54	-0.39							-0.42	-0.29
								(-7.23)	(-4.81)							(-7.06)	(-4.66)
Gross prof.										0.73	0.55					0.52	0.41
										(6.14)	(4.54)					(4.56)	(3.65)
Ivol												-26.92	-19.13			-28.98	-24.70
												(-5.46)	(-4.81)			(-7.96)	(-7.62)
β^{CAPM}														-0.02	0.09	0.17	0.22
														(-0.16)	(0.72)	(1.57)	(2.10)
Adj. R^2	0.008	0.019	0.006	0.019	0.009	0.020	0.004	0.016	0.003	0.017	0.016	0.021	0.017	0.026	0.026	0.046	0.048

Table 7: Persistence of momentum and buy-and-hold returns

This table reports the persistence of a portfolio's momentum return in Panel A and buy-and-hold return in Panel B. Each observation corresponds to a portfolio-year pair. In Panel A, from July of year t to June of year $t+1$, we compute the average momentum return of a portfolio as the y-variable in all five columns. Then, we compute the average momentum return over the past 1, 3, 5, 10, or 15 years as the x-variable in columns 1 to 5, respectively. We run Fama-MacBeth regressions and omit the constant in reporting. We perform the procedures to buy-and-hold returns in Panel B. All t-statistics, shown in parentheses, are based on Fama and MacBeth (1973) standard errors.

	A: Persistence of Momentum R_t^{MOM}						B: Persistence of Buy-and-Hold $R_t^{B\&H}$				
R_{t-1}^{MOM}	0.08					$R_{t-1}^{B\&H}$	0.02				
	(3.36)						(0.36)				
$R_{t-3,t-1}^{MOM}$		0.22				$R_{t-3,t-1}^{B\&H}$		0.02			
		(4.56)						(0.31)			
$R_{t-5,t-1}^{MOM}$			0.29			$R_{t-5,t-1}^{B\&H}$			-0.01		
			(4.87)						(-0.12)		
$R_{t-10,t-1}^{MOM}$				0.40		$R_{t-10,t-1}^{B\&H}$				0.08	
				(4.64)						(0.85)	
$R_{t-15,t-1}^{MOM}$					0.48	$R_{t-15,t-1}^{B\&H}$					0.05
					(4.66)						(0.37)
Obs.	18,538	17,798	17,058	15,208	13,358		18,538	17,798	17,058	15,208	13,358
Adj. R^2	0.026	0.047	0.056	0.070	0.084		0.126	0.093	0.056	0.042	0.045

Table 8: Risk-adjusted buy-and-hold vs. momentum returns in an out-of-sample test

This table reports the cross-sectional relationship between a portfolio's average buy-and-hold return (with or without adjusting for risk) and its past average momentum return with or without other control variables. Each observation corresponds to a portfolio. In both Panel A and B, the y-variables are related to the buy-and-hold return of a portfolio. They are the average, Sharpe ratio, CAPM alpha, FF3 alpha, and FFC4 alpha of buy-and-hold return measured from July 1992 to June 2018 in columns 1 to 5 respectively. The x-variable in Panel A is the average momentum return of a portfolio measured from July 1966 to June 1992. The x-variables in Panel B are the average momentum return, the idiosyncratic volatility of the buy-and-hold return (measured against the FF3 model), the skewness of the buy-and-hold return, and the kurtosis of the buy-and-hold return measured from July 1966 to June 1992. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}
	A: No controls				
R^{Mom}	-0.06	-0.17	-0.25	-0.21	-0.18
	(-2.66)	(-10.36)	(-9.27)	(-7.49)	(-7.41)
Obs.	370	370	370	370	370
Adj. R^2	0.028	0.286	0.277	0.196	0.190
	B: Controls				
R^{Mom}	-0.05	-0.14	-0.22	-0.18	-0.16
	(-2.24)	(-8.42)	(-8.23)	(-6.44)	(-6.70)
Kurtosis	0.05	0.08	0.12	0.08	0.04
	(1.48)	(2.48)	(2.34)	(1.53)	(0.93)
Skewness	0.03	0.02	0.03	0.01	0.01
	(2.40)	(2.03)	(2.37)	(0.54)	(0.49)
Ivol	0.00	-0.07	-0.05	-0.07	-0.03
	(0.00)	(-3.71)	(-1.39)	(-1.90)	(-1.06)
Obs.	370	370	370	370	370
Adj. R^2	0.036	0.319	0.290	0.208	0.188

Table 9: Trading strategy based on past momentum return

This table reports the abnormal returns of a trading strategy based on past average momentum returns of a portfolio. In June of each year t, we sort our 370 portfolios into 10 groups based on each portfolio's average momentum return measured over the previous 10 years. The first three columns correspond to the low (bottom decile) group. The next three columns correspond to the high (top decile) group. The final three columns correspond to the difference between high and low. From July of year t to June of year $t+1$, we take the simple average of the buy-and-hold returns of the portfolios in each group and report their CAPM alpha, FF3 alpha, and FFC4 alpha as well as corresponding factor loadings (e.g. market, value, size, and momentum). The sample period for returns is from July 1976 to June 2018. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	Low MTO portfolios			High MTO portfolios			High minus Low		
α	0.14	0.08	0.09	-0.26	-0.25	-0.19	-0.40	-0.34	-0.28
	(3.75)	(2.83)	(3.05)	(-4.01)	(-5.75)	(-4.46)	(-4.39)	(-5.45)	(-4.54)
β^{MKT}	0.89	0.93	0.93	1.22	1.15	1.13	0.33	0.22	0.21
	(107.35)	(136.93)	(134.64)	(82.87)	(108.94)	(112.20)	(16.23)	(14.99)	(14.22)
β^{HML}		0.16	0.16		-0.12	-0.15		-0.28	-0.31
		(15.88)	(15.10)		(-7.35)	(-9.44)		(-12.64)	(-13.69)
β^{SMB}		-0.06	-0.06		0.33	0.33		0.39	0.39
		(-5.98)	(-5.88)		(21.49)	(23.22)		(18.22)	(18.89)
β^{UMD}			-0.01			-0.08			-0.07
			(-1.42)			(-7.90)			(-4.81)
Obs.	504	504	504	504	504	504	504	504	504
Adj. R^2	0.958	0.975	0.975	0.932	0.969	0.972	0.343	0.702	0.714

Table 10: Analyst coverage

This table reports measures how MTO score is related to analyst coverage. We sort all firms into five quintiles by MTO score. To compute a firm's MTO score in June of year t, we first rank our 360 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. Then, for each MTO quintile, we compute the average number of analysts covering each firm, the standard deviation of earnings per share scaled by price, and the standard deviation in long-term EPS growth forecasts. Data on analyst forecasts is from I/B/E/S and sample period is from 1981 to 2018.

	Low MTO	2	3	4	High MTO
Number of analysts	6.0	5.5	4.4	3.2	2.0
Vol of EPS/P	0.17%	0.21%	0.28%	0.48%	0.86%
Vol of Δ EPS	3.0%	3.3%	3.7%	4.4%	6.0%

Table 11: Volatility of fundamentals

This table reports measures of volatility for several firm fundamentals across firms with different MTO scores. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. We then sort all firms into quintiles based on the MTO score, and report the average standard deviation across all firms in a quintile. The above procedure was done for all firms in a portfolio, but we also do this for just the losers, just the winners, and just the losers in a portfolio. The loser, neutral, and winner firms are defined as being in the bottom, middle, and top terciles of expected returns, conditional on being in the portfolio.

	Low MTO	2	3	4	High MTO	Low MTO	2	3	4	High MTO
	Net Income growth					Sales growth				
All	2.40	2.58	2.90	3.69	4.99	0.35	0.28	0.31	0.41	0.77
Loser	3.02	3.21	3.61	4.58	5.98	0.39	0.28	0.31	0.40	0.78
Neutral	1.95	2.16	2.54	3.34	4.73	0.27	0.25	0.28	0.38	0.70
Winner	2.14	2.28	2.45	3.10	4.46	0.38	0.30	0.33	0.45	0.81
	Asset growth					Gross profit				
All	0.28	0.24	0.26	0.32	0.53	0.07	0.07	0.07	0.08	0.10
Loser	0.27	0.23	0.23	0.28	0.44	0.07	0.07	0.07	0.08	0.10
Neutral	0.23	0.21	0.22	0.28	0.45	0.06	0.07	0.07	0.07	0.09
Winner	0.34	0.28	0.30	0.38	0.62	0.07	0.07	0.07	0.08	0.10
	ROA					EPS/Price				
All	0.04	0.03	0.03	0.04	0.08	0.05	0.05	0.06	0.08	0.11
Loser	0.05	0.03	0.03	0.05	0.09	0.07	0.06	0.08	0.11	0.15
Neutral	0.02	0.02	0.02	0.03	0.08	0.03	0.04	0.05	0.06	0.10
Winner	0.04	0.02	0.03	0.03	0.08	0.04	0.03	0.04	0.05	0.08

Table 12: Volatility of returns

This table shows the volatility of stock returns for firms and portfolios with different MTO scores. At the portfolio level, we first compute the standard deviation of the returns of each of the 370 portfolios, we also compute the standard deviation of the residual, after controlling for the FF3 factors. We then sort portfolios into MTO quintiles based on their average momentum return over their entire history. We report the average standard deviation across all 74 portfolios in each quintile. At the firm level, for each firm we compute an MTO score. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. We then sort all firms into quintiles based on the MTO score, and report the average standard deviation of all firms in a quintile. The above procedure was done for all firms in a portfolio, but we also do this for just the losers, just the winners, and just the neutral firms in a portfolio. The loser, neutral, and winner firms are defined as being in the bottom, middle, and top terciles of expected returns, conditional on being in the portfolio.

	Low MTO	2	3	4	High MTO	Low MTO	2	3	4	High MTO
	Portfolio level vol					Firm level vol				
All	16.4	17.1	17.7	19.0	22.8	44.6	45.0	45.5	66.6	83.1
Loser	23.0	23.9	25.2	26.0	30.3	52.2	43.3	49.6	83.3	93.1
Neutral	16.9	17.2	17.9	19.0	22.5	35.7	51.4	39.8	52.1	73.8
Winner	18.2	19.0	19.9	21.5	25.1	44.4	39.4	46.5	60.5	81.2
	Portfolio level, residual vol					Firm level, residual vol				
All	5.7	5.8	6.1	6.6	7.8	30.9	41.8	40.8	61.5	78.5
Loser	12.5	13.2	13.8	13.8	16.4	32.5	38.0	44.4	77.8	80.8
Neutral	8.2	8.2	8.5	8.9	10.2	25.3	47.4	35.2	45.6	69.1
Winner	9.2	9.5	9.9	10.1	11.3	34.0	39.5	42.2	56.7	84.5

Table 13: Autocorrelation of returns

This table shows the autocorrelation of stock returns with different MTO scores. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. We then sort all firms into quintiles based on the MTO score. For each quintile, we report the regression coefficient of regressing its value weighted return on its one month lagged value weighted return.

	R_t				
	Low Mom	2	3	4	High mom
R_{t-1}	0.01	0.03	0.09	0.10	0.13
	(0.31)	(0.75)	(2.03)	(2.32)	(2.87)
Obs.	503	503	503	503	503
Adj. R^2	-0.002	-0.001	0.006	0.009	0.014

Table 14: Determinants of momentum return portfolio characteristics

This table reports some of the characteristics that may explain the average momentum return. Each observation corresponds to a portfolio. The y-variable in every column is the average momentum return of a portfolio over the entire sample period from July 1966 to June 2018. The x-variables are the estimated market, HML, SMB, and UMD betas of the portfolio's buy-and-hold return, the portfolio's idiosyncratic volatility (controlling for the FF3 factor model), and the buy-and-hold return's skewness and kurtosis. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	Average Momentum Return				
$\beta_{B\&H}^{MKT}$	1.71	1.00	1.07	1.60	0.79
	(14.83)	(5.21)	(5.32)	(13.49)	(4.41)
$\beta_{B\&H}^{SMB}$		0.61	0.60		0.58
		(7.77)	(7.64)		(7.43)
$\beta_{B\&H}^{HML}$		-0.56	-0.55		-0.59
		(-10.37)	(-9.87)		(-9.43)
$\beta_{B\&H}^{UMD}$			0.18		
			(0.93)		
Ivol				0.09	0.12
				(2.74)	(4.32)
B&H skewness					-0.21
					(-2.31)
B&H kurtosis					0.02
					(0.95)
Obs.	370	370	370	370	370
Adj. R^2	0.544	0.595	0.595	0.554	0.617

Table 15: Relation to co-momentum

This table reports the key quantities in the paper split by comomentum tercile. We sort our 370 portfolios into ten groups, based on their past momentum return (MTO). We also classify each year as low, mid, or high co-momentum based on the June measure of co-momentum for each year, where co-momentum is defined as in Lou and Polk (2019). We compute the equal weighted average momentum return, and buy-and-hold return of each of the ten portfolio groups, from July of the year, to June of the following year, conditional on that year being low, mid, or high co-momentum. In Panel A, we report the average momentum return of the top and bottom group of portfolios, as well as their difference, in each sub-sample period. Analogously, in Panel B we report the CAPM alpha, and in Panel C the FF3 alpha.

	Low MTO	High MTO	Difference
	Average Momentum Return		
Low Comomentum	0.55	1.65	1.10
	(2.37)	(5.51)	(6.49)
Medium Comomenum	0.61	1.36	0.75
	(2.28)	(3.91)	(4.70)
High Comomentum	-0.12	0.25	0.37
	(-0.33)	(0.49)	(1.49)
	Buy-and-hold CAPM alpha		
Low Comomentum	0.09	-0.08	-0.17
	(1.63)	(-0.91)	(-1.30)
Medium Comomenum	0.07	-0.26	-0.33
	(0.86)	(-1.89)	(-1.64)
High Comomentum	0.25	-0.34	-0.59
	(4.18)	(-3.07)	(-4.07)
	Buy-and-hold FF3 alpha		
Low Comomentum	0.06	-0.11	-0.18
	(1.43)	(-1.72)	(-1.79)
Medium Comomenum	-0.05	-0.18	-0.14
	(-0.83)	(-2.80)	(-1.46)
High Comomentum	0.18	-0.35	-0.53
	(3.40)	(-3.75)	(-4.29)

A Appendix: creating anomaly portfolios

A.1 Size

At the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on their June-end market capitalization. Portfolio returns are the value-weighted return of stocks in each decile from July of year t to June of year $t+1$.

A.2 Book-to-market

At the end of June of each year t, we use NYSE breakpoints to sort stocks into deciles based on their book-to-market ratio, which is book equity for the fiscal year ending in calendar year $t-1$ divided by the market capitalization at the end of December of $t-1$. Following Davis et al. (2000), we measure book equity as stockholder's equity, plus balance-sheet deferred taxes and investment tax credit, minus the book value of preferred stock. If stockholders' equity (COMPUSTAT item SEQ) is not available, we use common equity (COMPUSTAT item CEQ) plus the par value of preferred stocks instead. If CEQ is not available, we use total asset (COMPUSTAT item AT) minus total liabilities (COMPUSTAT item LT).

A.3 Composite issuance

Daniel and Titman (2006) show that stocks with more composte equity issuance underperform. We compute composite equity issuance as the difference between the growth rate in market capitalization and cumulative stock return. At the end of June of each year t, we measure the growth rate of market capitalization from the end of June in year $t-1$ to June in year t. We also measure the cumulative stock return from the end of June in year $t-1$ to June in year t. We sort stocks into deciles at the end of June of each year t based on their composite equity issuance using NYSE breakpoints.

A.4 Momentum

Jegadeesh and Titman (1993) find that stocks with high momentum have higher returns. At the end of June of each year t, we measure each stock's 11-month total return from June of year $t-1$ to May of year t as their momentum. We then sort stocks into deciles based on the momentum using NYSE breakpoints. We require a stock to have at least six months of non-missing returns during the period that we measure momentum.

A.5 Net issuance

Following Fama and French (2008), we measure a stock's net issuance as the growth rate of split-adjusted number of shares outstanding (CRSP item CSHO times adjustment factor CFACSHR) from June of year $t-1$ to June of year t. We then sort stocks into deciles based on the net issuance at the end of June of year t using NYSE breakpoints.

A.6 CAPM beta

At the end of June of year t, we regress a stock's excess monthly return on the excess market return over the past five years. We require a stock to have at least 36 non-missing observations in the past five years. We then sort stocks into deciles based on their estimated market beta using NYSE breakpoints.

A.7 Idiosyncratic volatility

Ang et al. (2006) find that stocks that are highly volatile tend to have lower returns. At the end of June of year t, we measure a stock's idiosyncratic volatility as the residual volatility from the Fama-French 3 factor model over the past three months using daily returns. We then sort stocks into deciles based on their idiosyncratic volatility using NYSE breakpoints. We require a stock to have at least 20 non-missing observations over the three-month period.

A.8 Systematic volatility

Following Ang et al. (2006), we estimate a stock's exposure to shocks to systematic volatility. Specifically, we estimate the following equation using daily returns

$$r_{it} = a + \beta_{i,mkt} r_{mkt,t} + \beta_{i,svol} \Delta VXO_t + \epsilon_{it} \tag{3}$$

where $r_{mkt,t}$ is market excess return and ΔVXO_t is the daily change in the Chicago Board Options Exchange S&P 100 volatility index (VXO). We estimate this equation for the three month period before June of each year t. We then sort stocks into decile based on their $\beta_{i,svol}$ using NYSE breakpoints at the end of June of each year t. Due to data limitation, our first month of portfoilo return is in July 1986.

A.9 Total volatility

At the end of June of year t, we sort stocks into deciles based on their daily return volatility measured over the past three months. We use NYSE breakpoints and require a stock to have at least 20 non-missing observations over the three-month period.

A.10 Overpricing score

We sort stocks into deciles based on their overpricing score at the end of June of year t using NYSE breakpoints. Overpricing score is created by Stambaugh et al. (2015), which is an average of a stock's characteristic score over 11 different anomalies.

A.11 Earnings yield

Following Hou et al. (2015), we measure earnings yield as income before extraordinary items (COMPUSTAT item IB) for the fiscal year ending in year $t-1$ divied by the market capitalization at the end of December of year $t-1$. We then sort stocks into deciles based on their earnings yield at the end of June of year t using NYSE breakpoints.

A.12 Cash flow yield

Following Hou et al. (2015), we measure a stock's cash flow as income before extraordinary items (Compustat annual item IB), plus equity's share of depreciation (item DP), plus deferred taxes (if available, item TXDI). The equity's share is defined as market capitalization divided by total assets minus the book equity plus market capitalization. We then measure cash flow yield as the ratio of cash flow to the stock's December market capitalization. We sort stocks into deciles based on their cash flow yield using NYSE breakpoints at the end of June of year t.

A.13 Net payout yield

Following Hou et al. (2015), we measure a stock's net payout as dividends on common stock plus repurchases minus the sale of common and preferred stocks in the fiscal year ending in year $t-1$. We then divide net payout by the stock's December-end market capitalization as net payout yield. We sort stocks into deciles based on the net payout yield using NYSE breakpoints at the end of June of year t.

A.14 Advertising expense to market

Following Hou et al. (2015), we measure a stock's advertisement expense to its market capitalization by dividing its advertising expenses (COMPUSTAT item XAD) for the fiscal year ending in year $t-1$ by the stock's market capitalization in December year $t-1$. We then sort stocks into deciles based on their advertising expense to market using NYSE breakpoints at the end of June of year t. Due to data limitations, our first portfolio return is in July 1973.

A.15 ROA

We measure a stock's return on asset (ROA) as income before extraordinary items (Compustat annual item IB) divided by total assets (COMPUSTAT item AT) for the fiscal year ending in year $t-1$. We then sort stocks into deciles based on their ROA using NYSE breakpoints at the end of June of year t.

A.16 ROE

We measure a stock's return on equity (ROE) as income before extraordinary items (Compustat annual item IB) divided by its book equity for the fiscal year ending in year $t-1$. We then sort stocks into deciles based on their ROE using NYSE breakpoints at the end of June of year t.

A.17 Gross profitability

Novy-Marx (2013) finds that stocks with higher gross profit have higher returns. We measure gross profitability as total revenue (COMPUSTAT item REVT) minus the cost of goods sold (COMPUSTAT item COGS), then divide it by total assets (COMPUSTAT item AT). We then sort stocks into deciles based on their gross profitability using NYSE breakpoints at the end of June of year t.

A.18 Asset growth rate

Cooper et al. (2008) show that companies that grow their asset rapidly have lower subsequent returns. We measure asset growth rate as the growth rate of total asset (COMPUSTAT item AT) from fiscal year ending in year $t-2$ to fiscal year ending in year $t-1$. We then sort stocks into deciles based on their asset growth rate using NYSE breakpoints at the end of June of year t.

A.19 Investment to asset

Titman et al. (2004) find that companies that invest excessively have lower returns. We measure investment to asset as the sum of changes in gross property, plant, and equipment (Compustat annual item PPEGT) and changes in inventory (item INVT) in fiscal year ending in year $t-1$ divided by total assets in fiscal year ending in year $t-2$. We then sort stocks into deciles based on their investment to asset using NYSE breakpoints at the end of June of year t.

A.20 Investment growth

We measure investment growth as the growth rate of capital expenditure (COMPUSTAT item CAPX) from fiscal year ending in year $t-2$ to fiscal year ending in year $t-1$. We then sort stocks into deciles based on their investment growth rate using NYSE breakpoints at the end of June of year t.

A.21 Inventory change

Following Thomas and Zhang (2002) and Hou et al. (2015), we measure inventory change as change in inventory (COMPUSTAT item INVT) from fiscal year $t-2$ to $t-1$ scaled by the average total asset of fiscal year $t-2$ and $t-1$. We then sort stocks into deciles based on their inventory change using NYSE breakpoints at the end of June of year t.

A.22 Inventory growth

Following Hou et al. (2015), we measure inventory growth as the growth rate of inventory from fiscal year ending in year $t-2$ to fiscal year ending in year $t-1$. We then sort stocks into deciles based on their inventory growth rate using NYSE breakpoints at the end of June of year t.

A.23 Operating accrual

Sloan (1996) finds that stocks with more accruals have lower returns. Following Hou et al. (2015), we use the balance-sheet approach of Sloan (1996) to measure operating accrual prior to 1988 and use cash flow statement to measure operating accrual after 1988. Specifically, for the balance-sheet approach, we measure opearting accrual as changes in noncash working capital minus depreciation (COMPUSTAT item DP). Noncash working capital is the difference between noncash current asset (item ACT minus item CHE) and non-debt current liability, which is current liabilities (item LCT) less current debt (item DLC) and tax payable (item TXP). After 1988, we measure operating accrual as net income (item NI) minus net cash flow from operations (item OANCF). We scale operating acrrual by the lagged total asset. Then, at the end of June of year t, we sort stocks into deciles based on their operating accrual using NYSE breakpoints.

A.24 Percentage operating accrual

Hafzalla et al. (2011) show that operating accrual scaled by net income also predicts stock return. We measure percentage accrual as operating accrual scaled by the absolute value of net income in the fiscal year ending in year $t-1$. Then, at the end of June of year t, we sort stocks into deciles based on their operating accrual using NYSE breakpoints.

A.25 Percentage total accrual

Richardson et al. (2005) show that total accrual predicts stock return. Following Hou et al. (2015), prior to 1988, we use balance-sheet method to measure total accrual. Specifically, total accrual is change in noncash working capital plus change in net non-current operating asset plus change in net financial assets. Noncash working capital is the difference between noncash current asset (item ACT minus item CHE) and non-debt current liability, which is current liabilities (item LCT) less current debt (item DLC) and tax payable (item TXP). Non-current operating asset is non-current operating assets (item AT minus item ACT minus item IVAO) minus non-current operating liabilities (item LT minus item LCT minus DLTT). Net financial assets is financial assets (item IVST plus item IVAO) minus long-term debt (DLTT) and debt in minus current liabilities (DLC) plus preferred stocks (item PSTK). We scale total accrual by the absolute value of net income in the fiscal year ending in year $t-1$ as percent total accrual. Then, at the end of June of year t, we sort stocks into deciles based on percent total accrual using NYSE breakpoints.

A.26 Net operating assets

Hirshleifer et al. (2004) show that net operating assets scaled by total assets can predict a stock's return. We measure net operating assets as operating assets (defined as item AT minus item CHE) minus operating liabilities, which equal total assets (item AT) minus debt in current liabilities (item DLC), minus long-term debt (item DLTT), minus common equity (item CEQ), minus minority interest (item MIB), and minus preferred stocks (item PSTK). We then scale net operating assets by the lagged total assets. At the end of June of year t, we sort stocks into deciles based on their net operating assets using NYSE breakpoints.

A.27 Operating leverage

Following Novy-Marx (2010), we define operating leverage as operating cost, which is cost of goods sold (item COGS) plus selling, general, and administrative expenses (item XSGA), scaled by the total asset of fiscal year ending in year $t-1$. We then sort stocks into deciles based on their operating leverage using NYSE breakpoints at the end of June of year t.

A.28 Change in P&I

Change in P&I is the sum of changes in property, plant and equipments (COMPUSTAT item PPEGT) and inventory (item INVT) scaled by the lagged total assets (item AT). At the end of June of year t, we sort stocks into deciles based on their change in P&I using NYSE breakpoints.

A.29 R&D expense to market

We measure a stock's R&D expense (COMPUSTAT item XRD) as a percentage of its December-end market capitalization for the fiscal year ending in year $t-1$. We then sort stocks into deciles based on their R&D expense to market using NYSE breakpoints at the end of June of year t. Due to data limitations, our first month of R&D portfolio return is from July 1977.

A.30 Industry

We use Fama-French 10 industries as our definition of industry.

A.31 Duration

Following Dechow et al. (2004) and Hou et al. (2015), we measure equity duration as

$$Duration = \frac{\sum_{t=1}^{T} t \times CD_t/(1+r)^t}{ME} + \left(T + \frac{1+r}{r}\right) \frac{ME - \sum_{t=1}^{T} CD_t/(1+r)^t}{ME}, \quad (4)$$

where CD_t is the net cash distribution in year t, ME is market equity, T is the length of forecasting period (assume to be 10 years), and r is the cost of equity (assume to be 0.12). Net cash distribution is defined as $CD_t = BE_{t-1}(ROE_t - g_t)$, where BE is book-equity, ROE is return on equity, and g_t is the book equity growth rate. When forecasting ROE, the starting year ROE is income before extraordinary items (COMPUSTAT item IB) divided by lagged book equity (item CEQ). Then, we assume ROE follows a first-order autoregressive process with an autocorrelation of 0.57 and long-run mean of 0.12. We also model the growth rate in book equity as a first order autoregressive process with an autocorrelation coefficient of 0.24 and a long run-mean of 0.06. At the end of June of year t, we sort stocks into deciles based on their equity duration using NYSE breakpoints

A.32 Organizational capital

Following Eisfeldt and Papanikolaou (2013) and Hou et al. (2015), we measure organization capital OC as

$$OC_{it} = (1-\delta)OC_{it-1} + SG\&A_{it}/CPI_t \quad (5)$$

where the initial stock of organizational capital is $OC_{i0} = SG\&A_{i0}/(g+\delta)$, $SG\&A_{i0}$ (item XSGA) is the first valid observation of SG&A expense for the firm, g is the long-term growth rate of SG&A (assumed to be 10%), and δ is the depreciation rate of organizational capital (assumed to be 15%). Missing SG&A after the initial year are treated as zero. We scale organizational capital by lagged total asset. At the end of June of year t, we sort stocks into deciles based on their organizational capital using NYSE breakpoints

A.33 O-score

Based on Ohlson (1980) and Hou et al. (2015), we measure O-score as

$$\begin{aligned} OScore = &-1.32 - 0.407\log(TA) + 6.03TLTA - 1.43WCTA + 0.076CLCA \\ &- 1.72OENEG - 2.37NITA - 1.83FUTL + 0.285INTWO - 0.521CHIN \end{aligned} \quad (6)$$

where TA is total asset. $TLTA$ is the leverage ratio defined as the book value of debt (item DLC plus item DLTT) divided by total assets. $WCTA$ is working capital divided by total assets, where working capital is item ACT minus item LCT. $CLCA$ is current liabilities (item LCT) divided by current assets (item (ACT). $OENEG$ is 1 if total liabilities exceeds total asset and is zero otherwise. $NITA$ is net income divided by the total assets. $FUTL$ is the fund provided by operations (item PI) divided by the total liabilities (item LT). $INTWO$ equals to 1 if net income is negative for the prior 2 years, and zero otherwise. $CHIN$ is change in net income scaled by the average absolute value of net income over two years. At the end of June of year t, we sort stocks into deciles based on their equity duration using NYSE breakpoints.

A.34 Failure probability

Following Campbell et al. (2008) and Hou et al. (2015), we measure failure probability as

$$\begin{aligned} FP_t = &-9.164 - 20.264NIMTAAVG_t + 1.416TLMTA_t - 7.129EXRETAVG_t \\ &+ 1.411SIGMA_t - 0.045RSIZE_t - 2.132CASHMTA_t + 0.075MB_t - 0.058PRICE_t \end{aligned} \quad (7)$$

where

$$NIMTAAVG_{t-1,t-12} = \frac{1-\phi^3}{1-\phi^{12}}(NIMTA_{t-1,t-3} + ... + \phi^9 NIMTA_{t-10,t-12}) \tag{8}$$

$$EXRETAVG_{t-1,t-12} = \frac{1-\phi}{1-\phi^{12}}(EXRET_{t-1} + ... + \phi^{11} EXRET_{t-12}) \tag{9}$$

and $\phi = 2^{-1/3}$. $NIMTA$ is net income (COMPUSTAT quarterly item NIQ) divided by the sum of market equity and total liabilities. $EXRET = log(1+R_{it}) - log(1+R_{S\&P500,t})$ is the log excess return of a company relative to the S&P 500 index. $TLMTA$ is the ratio of total liabilities divided by the sum of market equity and total liabilities. $SIGMA$ is the annualized daily stock volatility estimated using a 3-month roling window. $RSIZE$ is the log ratio of the firm's market capitalization to the size of S&P 500 index. $CASHMTA$ is the liquidity position of firm, measured as the ratio of cash and short-term investment (COMPUSTAT quarterly item CHEQ) divided by the sum of market equity and total liabilities (item LTQ). MB is the market-to-book ratio. Following Campbell et al, we add 10% of the difference between the market capitalization and book equity to the book equity to reduce measurement errors. For firms with negative adjusted book equity, we replace it to \$1. $PRICE$ is the firm's log stock price, truncated above at \$15. We winsorize all right-hand-side variables of equation (7) at the 5th and 95th percentiles.

We sort stocks into deciles based on their failure probability twice at the end of June and December of each year. We impose a four-month gap between the date that failure probability is measured and the portfolio sorting date. We use NYSE breakpoints. Due to data limitations, our first monthly return is from January 1977.

A.35 Announcement return

We measure the four day cumulative abnormal return of a stock during its quarterly announcement, which is meant to capture earnings surprise. Timing of the return is from $t-2$ to $t+1$, where announcement day is at $t=0$. We subtract the market return from the stock's return to measure its abnormal return. At the end of June of year t and at the end of December of year t, we sort stocks into deciles based on their announcement return in the most recent quarter. We require the quarter-end day to be within 10 months, but more than 4 months ahead of the portfolio sorting day.

A.36 Standardized unexpected earnings (SUE)

Following Foster et al. (1984), We scale each stock's quarterly earnings per share (COMPUSTAT quarterly item EPSPXQ) by its standard deviation measured over the prior 8 quarters as standardized unexpected earnings (SUE). At the end of June of year t and at the end of December of year t, we sort stocks into deciles based on their SUE in the most recent quarter using NYSE breakpoints. We require the quarter-end day to be within 10 months, but more than 4 months ahead of the portfolio sorting day.

A.37 Consecutive earning growth

We follow Barth et al. (1999), Green et al. (2013), and Hou et al. (2015). We count the number of quarters that a stock experiences positive earnings (COMPUSTAT quarterly item IBQ) change relative to the same quarter in the prior year. At the end of June of year t and at the end of December of year t, we sort stocks into deciles based on how many quarters they have positive earnings change using NYSE breakpoints. We require the most recent quarter-end day to be within 10 months, but more than 4 months ahead of the portfolio sorting day.

Table A1: Momentum profits and Buy-and-Hold returns, excluding microcaps

This table reports the cross-sectional relationship between a portfolio's average buy-and-hold return (with and without adjusting for risk) and average momentum return. It is identical to Table 3 but excludes firms in the bottom 20% of the size distribution, and firms with a stock price below $5. Each observation corresponds to a portfolio. The y-variables are related to the buy-and-hold return of a portfolio, they are: the average buy-and-hold return, its Sharpe ratio, its CAPM alpha, its FF3 alpha, and its FFC4 alpha. The x-variable is the average return of a momentum strategy within a portfolio. Panel A presents the baseline result. Panel B presents the same result, but including the portfolio return's kurtosis, skewness, and idiosyncratic volatility as controls. Panels C and D present sub-sample analysis, for 1966-1992 and 1992-2018 respectively. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}
	A: Baseline result					B: Full sample with controls				
R^{MOM}	-0.1577	-0.2299	-0.3534	-0.2417	-0.2112	-0.1574	-0.1976	-0.3096	-0.1741	-0.1671
	(-6.06)	(-14.92)	(-12.61)	(-6.80)	(-6.94)	(-5.61)	(-11.88)	(-11.33)	(-5.96)	(-6.28)
Skewness						0.1206	0.1132	0.2875	0.3479	0.2633
						(1.97)	(2.48)	(4.19)	(4.31)	(4.03)
Kurtosis						0.0297	0.0241	0.0372	-0.0189	-0.0146
						(1.62)	(1.86)	(1.87)	(-0.85)	(-0.83)
Ivol						0.0201	-0.0265	-0.0232	-0.0703	-0.0435
						(1.03)	(-2.28)	(-1.23)	(-3.48)	(-2.52)
Obs.	368	368	368	368	368	368	368	368	368	368
Adj. R^2	0.117	0.387	0.385	0.180	0.191	0.136	0.404	0.421	0.282	0.265
	C: 1966-1992 sub-sample					D: 1992-2018 sub-sample				
R^{MOM}	-0.1414	-0.1479	-0.2278	-0.0933	-0.0880	-0.0716	-0.2094	-0.3364	-0.2764	-0.2722
	(-4.57)	(-8.13)	(-7.26)	(-2.77)	(-2.88)	(-2.54)	(-10.19)	(-9.74)	(-7.39)	(-8.39)
Obs.	368	368	368	368	368	368	368	368	368	368
Adj. R^2	0.091	0.217	0.207	0.036	0.043	0.021	0.236	0.271	0.179	0.216

Table A2: Momentum profits and Buy-and-Hold returns, excluding financial crisis

This table reports the cross-sectional relationship between a portfolio's average buy-and-hold return (with and without adjusting for risk) and average momentum return. It is identical to Table 3 but excludes the financial crisis from July 2008 to June 2011. Each observation corresponds to a portfolio. The y-variables are related to the buy-and-hold return of a portfolio, they are: the average buy-and-hold return, its Sharpe ratio, its CAPM alpha, its FF3 alpha, and its FFC4 alpha. The x-variable is the average return of a momentum strategy within a portfolio. Panel A presents the baseline result. Panel B presents the same result, but including the portfolio return's kurtosis, skewness, and idiosyncratic volatility as controls. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}
	A: Baseline result					B: Full sample with controls				
R^{MOM}	-0.14	-0.20	-0.30	-0.22	-0.17	-0.13	-0.17	-0.25	-0.16	-0.16
	(-6.33)	(-15.44)	(-12.71)	(-8.59)	(-8.75)	(-5.52)	(-13.17)	(-10.92)	(-6.57)	(-7.50)
Skewness						0.14	0.10	0.29	0.18	0.10
						(2.86)	(2.96)	(5.21)	(3.38)	(2.38)
Kurtosis						0.06	0.04	0.06	0.01	-0.00
						(4.48)	(5.11)	(4.78)	(0.38)	(-0.12)
Ivol						0.02	-0.02	-0.04	-0.08	-0.03
						(0.74)	(-1.80)	(-1.93)	(-2.93)	(-1.36)
Obs.	370	370	370	370	370	370	370	370	370	370
Adj. R^2	0.140	0.426	0.403	0.226	0.240	0.205	0.472	0.477	0.272	0.250

Table A3: PEAD profits and Buy-and-Hold returns

This table reports the cross-sectional relationship between a portfolio's average buy-and-hold return (with and without adjusting for risk) and its average post-earnings announcement drift (PEAD). It is identical to Table 3 but uses PEAD instead of momentum to sort portfolios. Each observation corresponds to a portfolio. The y-variables are related to the buy-and-hold return of a portfolio, they are: the average buy-and-hold return, its Sharpe ratio, its CAPM alpha, its FF3 alpha, and its FFC4 alpha. The x-variable is the average PEAD return of the portfolio, computed as the difference in return between stocks in the top tercile and bottom tercile of each portfolio ranked by their earnings surprises in the past quarter. We measure earning surprise as the difference between realized EPS and average analyst forecast, all scaled by stock price. Panel A presents the baseline result. Panel B presents the same result, but including the portfolio return's kurtosis, skewness, and idiosyncratic volatility as controls. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}	$R^{B\&H}$	SR	α^{CAPM}	α^{FF3}	α^{FFC4}
	A: Baseline result					B: Full sample with controls				
R^{MOM}	-0.01	-0.17	-0.27	-0.19	-0.13	-0.00	-0.12	-0.21	-0.12	-0.09
	(-0.36)	(-7.30)	(-5.53)	(-4.47)	(-3.76)	(-0.01)	(-5.74)	(-5.26)	(-3.31)	(-3.09)
Skewness						0.23	0.18	0.31	0.32	0.29
						(3.31)	(3.34)	(3.04)	(3.23)	(3.57)
Kurtosis						0.01	0.00	0.01	-0.01	-0.01
						(0.67)	(0.39)	(0.62)	(-0.65)	(-0.46)
Ivol						-0.04	-0.10	-0.12	-0.14	-0.09
						(-2.70)	(-8.53)	(-4.86)	(-5.49)	(-4.61)
Obs.	370	370	370	370	370	370	370	370	370	370
Adj. R^2	-0.002	0.161	0.151	0.076	0.052	0.073	0.323	0.253	0.228	0.182

Table A4: CAPM buy-and-hold alpha vs. momentum return for each anomaly

This table reports the relationship between a portfolio's CAPM buy-and-hold alpha and average momentum return for different sets of anomaly characteristics. Each column corresponds to a particular characteristic, which is used to sort all firms into deciles. Thus, each regression has 10 observations. We sort firms into portfolios in June of year t and measure each portfolio's buy-and-hold and momentum returns in July of year t to June of year $t+1$. Then, we use the average CAPM alpha of buy-and-hold returns as the y-variable and the average of momentum returns as the x-variable. The sample period is from July 1966 to June 2018. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	Size	BM	Compos. issue	Momntm	Net issue	β^{CAPM}	Idios. Vol.	Total Vol.	Mispricing score
R^{MOM}	0.09	-0.19	-0.36	-0.32	-0.65	-0.54	-0.49	-0.55	-0.64
	(2.47)	(-1.44)	(-3.00)	(-2.10)	(-2.92)	(-4.20)	(-3.97)	(-6.71)	(-7.06)
Adj. R^2	0.363	0.107	0.471	0.276	0.455	0.650	0.621	0.830	0.844

	Earn. yield	CF yield	Net pay yield	Adv. exp. to market	ROA	Gross profit	ROE	Asset growth	Inv. to asset
R^{MOM}	-0.38	-0.32	-0.36	-0.07	-0.32	0.07	-0.16	-0.33	-0.32
	(-3.16)	(-3.22)	(-3.52)	(-0.36)	(-4.64)	(0.28)	(-1.93)	(-4.45)	(-3.24)
Adj. R^2	0.499	0.510	0.558	-0.108	0.695	-0.114	0.232	0.677	0.513

	Invest. growth	Invntry change	Invntry growth	% operat. accrual	Operat. accural	% tot. accrual	Net operat. asset	Operat. leverage	Change P&I
R^{MOM}	-0.38	-0.31	-0.30	-0.52	-0.37	-0.26	-0.36	0.03	-0.29
	(-3.43)	(-2.03)	(-3.23)	(-4.58)	(-3.73)	(-2.00)	(-3.89)	(0.19)	(-3.71)
Adj. R^2	0.545	0.258	0.511	0.690	0.589	0.251	0.610	-0.120	0.587

	O-score	Duration	Org. cap. to asset	R&D to market	System. vol.	Failure prob.	SUE	Earn. ann. return	Consec. earn. gr.
R^{MOM}	-0.29	-0.39	-0.27	-0.12	-0.19	-0.30	0.18	-0.12	0.05
	(-5.01)	(-4.16)	(-1.99)	(-0.64)	(-1.44)	(-2.24)	(1.04)	(-1.41)	(0.53)
Adj. R^2	0.728	0.645	0.247	-0.071	0.106	0.308	0.009	0.100	-0.086

Table A5: Twenty portfolios with highest and lowest MTO

This table lists top twenty portfolios with the highest and lowest momentum returns. Momentum return is the monthly difference in return between winners and losers within a portfolio. Category is the name of the firm characteristics used to sort firms into portfolios. Portfolio number indicates the level of the characteristics, 1 being the lowest and 10 being the highest. The sample period is from July 1966 to June 2018.

Top 20 momentum profit			Bottom 20 momentum profit		
Anomaly	Decile	Average momentum return	Anomaly	Decile	Average momentum return
O-score	10	1.68	Operating accrual	4	-0.02
Change in P&I	10	1.55	Earnings yield	8	-0.02
Asset growth	10	1.52	Composite issuance	4	-0.01
Cash flow yield	1	1.49	Cash flow yield	8	0.02
Net operating asset	10	1.48	Duration	5	0.03
Duration	10	1.46	Composite issuance	3	0.03
Earnings yield	1	1.46	Asset growth	5	0.03
Investment to asset	10	1.43	ROE	6	0.05
Total volatility	10	1.40	Inventory growth	4	0.05
Inventory change	10	1.39	Systematic volatility	5	0.06
Org. capital to asset	1	1.37	Net payout yield	6	0.08
Size	1	1.36	Total volatility	1	0.09
Mispricing score	10	1.36	Cash flow yield	5	0.10
Idiosyncratic volatility	10	1.33	Ad. expense to market	8	0.11
ROE	1	1.30	Earnings yield	7	0.12
Composite issuance	10	1.30	Failure probability	5	0.12
Operating accrual	10	1.29	Net payout yield	8	0.13
Inventory growth	10	1.27	CAPM beta	1	0.13
Net payout yield	1	1.26	Ad. expense to market	6	0.13
Net payout yield	2	1.26	Mispricing score	1	0.14

Table A6: Correlation between MTO score and other firm characteristics

This table presents correlations between our measure of a how informative its prices are, which we refer to as MTO, and other firm characteristics. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to.

	MTO	Ivol	Asset growth	Book-to-market	Size	CAPM beta	Gross profit
MTO	1.00						
Ivol	0.50	1.00					
Asset growth	0.34	0.14	1.00				
Book-to-market	-0.10	0.03	-0.14	1.00			
Size	-0.33	-0.41	-0.01	-0.33	1.00		
β^{CAPM}	0.21	0.14	0.06	-0.08	0.08	1.00	
Gross profit	-0.10	0.02	-0.08	-0.10	-0.08	0.06	1.00

Table A7: Momentum profits and Buy-and-Hold at firm level: Fama-MacBeth

This table is identical to Table 6 but excludes microcaps. It reports the relationship between the stock return and the MTO score at the firm level. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. We then run Fama-MacBeth regressions of stock return on its MTO score and other control variables. All t-statistics, shown in parentheses, are based on Fama-MacBeth standard errors. The number of observations is between 972,818 and 826,866; it is omitted to conserve space.

	Monthly return																
MTO	-1.61		-1.80		-1.47		-1.93		-1.29		-1.57		-1.09		-1.67		-0.92
	(-2.13)		(-2.42)		(-1.89)		(-2.78)		(-1.55)		(-1.93)		(-2.27)		(-2.87)		(-2.33)
Size		-0.06	-0.11													-0.13	-0.14
		(-1.62)	(-3.20)													(-4.26)	(-4.51)
BM				0.15	0.06											0.09	0.08
				(1.25)	(0.62)											(1.05)	(0.95)
Momntm						0.53	0.54									0.51	0.50
						(3.20)	(3.47)									(3.69)	(3.70)
Asset gr.								-0.28	-0.15							-0.24	-0.15
								(-2.73)	(-1.88)							(-3.48)	(-2.13)
Gross prof.										0.45	0.33					0.41	0.34
										(2.99)	(2.23)					(2.95)	(2.47)
Ivol												-23.90	-16.81			-28.16	-24.93
												(-2.96)	(-2.52)			(-5.38)	(-5.24)
β^{CAPM}														0.01	0.13	0.10	0.13
														(0.07)	(0.88)	(0.80)	(1.05)
R^2	0.022	0.009	0.029	0.011	0.030	0.018	0.037	0.009	0.026	0.007	0.030	0.029	0.036	0.034	0.042	0.079	0.082

Table A8: Loading of fundamentals on GDP

This table presents the loadings of various quantities related to a firm's fundamentals on GDP growth. We sort firms into quintiles based on their MTO score. To compute a firm's MTO score in June of year t, we first rank our 370 portfolios as listed in Table 1 based on each portfolio's average momentum return in the past 10 years. Then, for each firm, we define its MTO score as the number of top 10% MTO portfolios it belongs to relative to total portfolios it belongs to. We then run a pooled OLS regression of all firms in a quintile on GDP growth at the quarterly frequency. All t-statistics, shown in parentheses, are based on robust standard errors of White (1980).

	Low MTO	2	3	4	High MTO	Low MTO	2	3	4	High MTO
	Net Income growth					Sales growth				
GDP gr.	8.69	13.99	18.03	23.15	23.88	2.50	2.42	2.27	2.66	2.77
	(8.05)	(13.21)	(15.97)	(17.17)	(14.52)	(16.48)	(22.27)	(19.93)	(19.20)	(13.60)
Obs.	61,189	72,741	88,536	111,438	145,285	67,748	79,904	99,826	137,074	249,462
Adj. R^2	0.001	0.002	0.003	0.003	0.001	0.004	0.006	0.004	0.003	0.001
	Asset growth					Gross profit				
GDP gr.	1.00	0.87	0.99	1.68	3.70	-0.46	0.02	0.72	0.53	0.96
	(8.20)	(9.14)	(10.51)	(15.48)	(26.60)	(-15.45)	(0.90)	(27.32)	(20.99)	(37.01)
Obs.	67,470	79,663	99,637	136,864	255,050	68,465	79,849	99,983	138,249	258,629
Adj. R^2	0.001	0.001	0.001	0.002	0.003	0.003	-0.000	0.007	0.003	0.005
	ROA					EPS/Price				
GDP gr.	0.02	0.11	0.23	0.35	1.10	0.65	0.63	0.74	0.91	1.66
	(1.17)	(11.00)	(22.04)	(27.45)	(50.41)	(29.82)	(35.60)	(34.52)	(35.58)	(55.99)
Obs.	69,097	80,460	100,780	139,451	261,172	69,642	80,733	101,387	140,497	261,960
Adj. R^2	0.000	0.001	0.005	0.005	0.010	0.013	0.015	0.012	0.009	0.012