Expected investment and the cross-section of stock returns

Qi Lin, Xi Lin

PII: S0165-1765(18)30319-7
DOI: https://doi.org/10.1016/j.econlet.2018.08.012
Reference: ECOLET 8162

To appear in: *Economics Letters*

Received date : 2 July 2018
Accepted date : 8 August 2018

Expected Investment and the Cross-Section of Stock Returns

Qi Lin[a,1,*], Xi Lin[b,2]

[a]*Zhejiang University of Finance & Economics*
[b]*Fudan University*

Abstract

In this paper, we show that the relation between expected investment and future stock returns (i.e., the expected investment-return relation) is negative and inconsistent with the multiperiod q theory. Further analysis reveals that the expected investment change measure of Hou et al. (2018a) is a poor proxy for future investment because of the mismatch of investment characteristics and the incorrect constraint imposed on the regression.

Keywords: Expected investment, Cross-sectional returns, Expected investment change, Multiperiod q theory

JEL classification: G11, G12, G17

*Corresponding author

Email address: `qlin_sf@zufe.edu.cn` (Qi Lin)

[1]School of Finance, Zhejiang University of Finance & Economics, Hangzhou 310018, China. Tel: +86 571 8755 7108

[2]School of Economics, Fudan University, Shanghai 200433, China.

1. Introduction

In a multiperiod q theory framework, the first principle of investment implies that the investment return can be expressed as

$$R^I_{i,t+1} = \frac{(1-\tau_{t+1})\frac{\Pi_{i,t+1}}{A_{i,t+1}} + \frac{1}{2}(1-\tau_{t+1})a_{t+1}(\frac{I_{i,t+1}}{A_{i,t+1}})^2 + \delta_{i,t+1}\tau_{t+1} + (1-\delta_{i,t+1})\left[1+(1-\tau_{t+1})a_{t+1}(\frac{I_{i,t+1}}{A_{i,t+1}})\right]}{1+(1-\tau_t)a_t(\frac{I_{i,t}}{A_{i,t}})}, \tag{1}$$

where $R^I_{i,t}$ is the investment return; $\Pi_{i,t}$, $I_{i,t}$, and $A_{i,t}$ are the operating profits, investment, and productive assets of firm i in time t; $\delta_{i,t}$ is the exogenous depreciation rate of capital and is firm-specific and time-varying, τ_t is the corporate tax rate and is time-varying, and $a_t > 0$.[3] This equation predicts that, all else equal, stocks with high expected investment should earn higher expected returns than stocks with low expected investment.

Hou et al. (2018a,b) argue that decomposing Equation (1) leads to two different components: a "dividend yield", $\frac{(1-\tau_{t+1})\Pi_{i,t+1}/A_{i,t+1}+(1/2)(1-\tau_{t+1})a_{t+1}(I_{i,t+1}/A_{i,t+1})^2+\delta_{i,t+1}\tau_{t+1}}{1+a_t(I_{i,t}/A_{i,t})}$, that represents two well-documented determinants of the expected returns (the profitability and investment-to-assets (I/A) variables) and a "capital gain", $\frac{(1-\delta_{i,t+1})[1+(1-\tau_{t+1})a_{t+1}(I_{i,t+1}/A_{i,t+1})]}{1+a_t(I_{i,t}/A_{i,t})}$, that represents a third return determinant. They also find that their I/A growth proxy positively relates to future stock returns.

In this paper, we focus on the expected investment-return relation, which is directly predicted by the multiperiod q theory, and call attention to the fact that Hou et al.'s decomposition is misleading. By construction, the I/A growth component contains two different anomalies in firm characteristics (the expected and current investment anomalies), where the latter is the same as that in the "dividend yield" component. Therefore, the I/A growth component is not a single but rather a composite anomaly that combines the well-documented asset growth effect of Cooper et al. (2008) with a new expected investment effect.[4] In addition, although the results of Hou et al. (2018a,b) appear to be consistent with the multiperiod q theory, they tell us little about whether the documented positive linkage originates from the positive relation between the expected investment and expected returns, as suggested by the q theory, the inverse of the asset growth effect that is difficult to isolate with their proxy, or the mix of both. More importantly, by using the "capital gain" component as a determinant,

[3] See Liu et al. (2009) and Hou et al. (2018a) for detailed derivations of the equivalence between the stock and the investment returns. In a 2-period investment model (Hou et al., 2015), it is natural to assume that the investment friction parameter a is time-invariant. Since firms usually face a lower degree of investment friction over expansionary periods, in a multiperiod framework, it would be more appropriate to assume that a is time-varying.

[4] Following Hou et al. (2018a), we can also treat the profitability and investment anomalies in the "dividend yield" component as one determinant in order to form a profit/investment factor instead of including them as two different factors in the q-factor model of Hou et al. (2015). However, as shown in the paper, simply combining two different anomalies to generate a powerful return predictor can lead to misleading inference because different characteristics may contain different information about future returns.

Hou et al. (2018a,b) simply assume that the relation between expected investment and future stock returns is positive and consistent with the multiperiod q theory, which relies on certain theoretical assumptions that may or may not be valid, whereas the expected investment-return relation remains an open question.

As such, we first construct a direct expected investment measure, denoted by $E_t[I/A]$, and form an implied expected investment component, denoted by $E_t[\varepsilon_{I/A}]$, by adding current I/A back to the Hou et al.'s (2018a) expected I/A changes ($E_t[d^1 I/A]$), and then conduct both parametric and nonparametric tests to assess their predictive power. Our results reveal that $E_t[d^1 I/A]$ positively relates to future returns, which appears to be consistent with the multiperiod q theory. In contrast, we find opposite evidence that both $E_t[\varepsilon_{I/A}]$ and $E_t[I/A]$ are negative and statistically significant return predictors. The high-minus-low strategy earns -131 basis points per month on average with a Newey-West t-statistic of -7.977 and a Sharpe ratio of -1.192, which is greater than the average return of -0.817% (t = -5.210; Sharpe ratio = -0.841) in $E_t[\varepsilon_{I/A}]$. After controlling for the well-known market, size, book-to-market, momentum, profitability, and investment factors discussed by Carhart (1997), Fama and French (1993, 2015), and Hou et al. (2015), we find that the difference between the returns on the decile portfolios with the highest and lowest $E_t[I/A]$ ($E_t[\varepsilon_{I/A}]$) remains negative and highly significant, and the predictability of $E_t[I/A]$ is consistently stronger than that of $E_t[\varepsilon_{I/A}]$.

There are two potential reasons why the two expected investment measures and the expected investment change measure produce contrasting conclusions. First, we find that the return spread between the highest and lowest deciles on $E_t[d^1 I/A]$ earns positive profits largely because stocks in the lowest (highest) decile are those with high (low) current I/A instead of future I/A, which suggests that $E_t[d^1 I/A]$ is a poor proxy for future investment. In contrast, $E_t[I/A]$ is an appropriate proxy because it consistently aligns with firms' future investment. Stocks with high $E_t[I/A]$ (decile 10) are those that invest aggressively in the future (0.264), whereas stocks with low $E_t[I/A]$ (decile 1) are those that invest conservatively (0.023). Furthermore, we show that the poor linkage between $E_t[d^1 I/A]$ and future investment and the weaker predictability of $E_t[\varepsilon_{I/A}]$ are due to the biased parameter estimates, which is constructed based on the incorrect constraint imposed in the regression. Running a regression using $E_t[d^1 I/A]$ as the dependent variable is equivalent to running a constrained regression of $E_t[I/A]$ and imposing the constraint that the slope of I/A must be one. However, we show that the estimated I/A slopes from the unconstrained regressions are much less than one, which indicates that the constraint imposed by Hou et al. (2018a) is inappropriate. Running a regression by imposing the incorrect restriction can lead to biased parameter estimates (Greene, 2012), hence biasing the fitted dependent variable and the associated $E_t[\varepsilon_{I/A}]$ derived from it.

The remainder of this paper is organized as follows. Section 2 describes the data and variables. In Sections 3 and 4, we analyze the relation between $E_t[d^1 I/A]$, $E_t[\varepsilon_{I/A}]$, and $E_t[I/A]$ and the cross-section of stock returns and discuss the potential reasons why the two expected investment measures and the expected investment change measure yield contrasting predictions. The final section reviews our conclusions.

2. Data and variable definitions

We obtain stock returns from the Center for Research in Security Prices (CRSP) and accounting data from the merged CRSP-Compustat database from 1961 to 2017. We begin with all firms traded on the NYSE, Amex, and Nasdaq and then exclude securities other than common stocks (firms with CRSP share codes of 10 or 11). Following Hou et al. (2015), we also exclude financial firms (firms with one-digit standard industrial classification codes of six) and firms with negative book equity.[5] We use CRSP delisting returns. If a delisting return is missing and the delisting is performance-related, we impute a return of -30% for NYSE and Amex stocks (Shumway, 1997) and -55% for Nasdaq stocks (Shumway and Warther, 1999) to avoid survivorship bias.

In accordance with Fama and French (1993, 2015), we form all our annual accounting variables at the end of June in year t by using annual accounting information from fiscal year-end $t-1$ from Compustat. In computing the book-to-market ratio, we use the market value of equity (price per share times the number of shares outstanding from CRSP) as of December in year $t-1$. For firm capitalization, we use the market value of the firm's equity from CRSP at the end of June in year t.

We follow Hou et al. (2018a) and use their expected I/A changes ($E_t[d^1I/A]$) as a proxy for the expected I/A growth in the "capital gain" component.[6] At the end of June of each year t, we compute $E_t[d^1I/A]$ with the log of Tobin's q, denoted by $\log(q)$, and operating cash flows, denoted by COP, for the fiscal year ending in calendar year $t-1$; we also calculate the Fama and MacBeth (1973) cross-sectional regression coefficients estimated from the prior ten-year rolling window.[7] We require a minimum of five years. Both the dependent variables and regressors are winsorized cross-sectionally

[5] We use the Fama and French (1993) definition of the book value of equity, and supplement the Compustat information with the Davis et al. (2000) book values of equity from Ken French's website. The book value of equity is defined as shareholders' equity (item SEQ), plus balance sheet deferred taxes and investment tax credits (item TXDITC), minus the book value of preferred stock. We set missing values of balance sheet deferred taxes and investment tax credits equal to zero. To calculate the book value of preferred stock, we set it equal to the redemption value (item PSTKRV), if available, or the liquidation value (item PSTKL) or the par value (item PSTK), in that order. If shareholders' equity is missing, we set it equal to the value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), if available, or total assets (item AT) minus total liabilities (item LT).

[6] The expected I/A changes in Hou et al. (2018b) are constructed at the monthly frequency and the annual accounting information used are lagged at least four months as of month t instead of the standard six months. To facilitate the comparison of our results with those in the prior asset pricing literature, in this paper, we use the construction method as in Hou et al. (2018a). However, our results are robust using the Hou et al.'s (2018b) method.

[7] Tobin's q is defined as the market equity (price per share times the number of shares outstanding from CRSP) plus long-term debt (item DLTT) and short-term debt (item DLC) scaled by the book value of total assets (item AT). We

in each June at the 1st and 99th percentiles, and the regressions are estimated via weighted least squares (WLS) with the market equity as the weight in order to alleviate the impact of microcaps (Fama and French, 2008). We further construct an implied expected investment component ($E_t[\varepsilon_{I/A}]$) by adding current I/A back to the Hou et al.'s (2018a) expected I/A changes ($E_t[d^1 I/A]$) as a proxy for the expected investment in Equation (1).

We also use the expected I/A ($E_t[I/A]$) instead of I/A changes, which directly links to the expected I/A item in Equation (1). $E_t[I/A]$ is constructed closely following Hou et al. (2018a), except that we further add the current I/A as an explanatory variable.

3. Empirical results

In this section, we conduct both parametric and nonparametric tests to assess the predictive power of $E_t[d^1 I/A]$, $E_t[\varepsilon_{I/A}]$, and $E_t[I/A]$ over future stock returns.

3.1. Fama and MacBeth regressions

Table 1 presents average Fama and MacBeth (1973) slope coefficient estimates (multiplied by 100) and their associated Newey-West t-statistics to compare the explanatory power of $E_t[d^1 I/A]$, $E_t[\varepsilon_{I/A}]$, and $E_t[I/A]$. We test the robustness of our results to the inclusion of existing asset pricing characteristics by adding the following controls. The first set of controls is the market beta ($\hat{\beta}_{Mkt}$), the natural logarithm of the book-to-market ratio (log(B/M)), and the natural logarithm of the market value of equity (log(ME)). In the second set of controls, we further add past returns for the prior month ($R_{1,0}$) and for the prior 12-month period excluding month $t-1$ ($R_{12,2}$). The final set of controls includes those controls in the first and second sets as well as the quarterly return on equity (ROEQ).[8] We estimate the regressions monthly using data from July 1967 through December 2017. Standardized regressors are calculated as the difference between the variables and their cross-sectional means, divided by the cross-sectional standard deviations. All the variables are winsorized at the 1st and 99th percentiles.

follow Ball et al. (2016) and measure operating cash flows as total revenue (item REVT) minus cost of goods sold (item COGS), minus selling, general, and administrative expenses (item XSGA), plus research and development expenditures (item XRD), if available, minus change in accounts receivable (item RECT), minus change in inventory (item INVT), minus change in prepaid expenses (item XPP), plus change in deferred revenue (item DRC plus item DRLT), plus change in trade accounts payable (item AP), and plus change in accrued expenses (item XACC), scaled by the book value of total assets (item AT). All changes are annual changes, and the missing changes are set to zero.

[8] ROEQ is defined as income before extraordinary items (item IBQ) scaled by the one-quarter-lagged book equity. Quarterly book equity is defined as shareholders' equity (item SEQQ), plus balance sheet deferred taxes and investment tax credit (item TXDITCQ) if available, minus the book value of preferred stock (item PSTKQ). If shareholders' equity is missing, we set it equal to the value of common equity (item CEQQ) plus the book value of preferred stock, or total assets (item ATQ) minus total liabilities (item LTQ).

Echoing the evidence found by Hou et al. (2018a), the results in Columns (1)–(4) indicate a positive and significant relation between $E_t[d^1I/A]$ and the cross-section of future stock returns. The Newey-West t-statistics from the monthly regressions of realized returns on $E_t[d^1I/A]$ range from 5.732 to 10.707. However, when substituting $E_t[d^1I/A]$ with a direct expected investment measure, we find that $E_t[I/A]$ is indeed a negative and statistically significant return predictor, which is inconsistent with the q theory. The univariate regression results reported in Column (9) indicate a negative and statistically significant relation between $E_t[I/A]$ and the cross-sectional stock returns. The average slope from the monthly regressions of realized returns on $E_t[I/A]$ alone is -0.458 with a Newey-West t-statistic of -8.770. Column (12) in Table 1 controls for all variables simultaneously, including $\hat{\beta}_{MKT}$, size, book-to-market, momentum, short-term reversal, and profitability. In this more general specification, the average slope of $E_t[I/A]$ remains negative, -0.381, and highly significant, with a Newey-West t-statistic of -9.650. Consistently, the average slopes for the implied component of expected investment $E_t[\varepsilon_{I/A}]$ derived from $E_t[d^1I/A]$ also remain negative and highly significant after controlling for stock-specific variables, with Newey-West t-statistics ranging from -7.485 to -9.179, which further confirms the findings of $E_t[I/A]$ that expected investment is negatively linked to future stock returns.

3.2. Portfolio-level analysis

Given that the Fama and MacBeth (1973) regressions are sensitive to outliers, we also perform portfolio tests, which provide a potentially more robust method to evaluate predictive ability without imposing parametric assumptions between the variables. At the end of June of each year, we form decile portfolios by sorting individual stocks based on $E_t[d^1I/A]$, $E_t[\varepsilon_{I/A}]$, and $E_t[I/A]$, partitioned at the NYSE market capitalization breakpoint. We rebalance the portfolios annually at the end of June and the cross-sectional return predictability results are reported from July 1967 to December 2017.

As can be seen, the excess return on the high-minus-low decile portfolio formed on $E_t[d^1I/A]$ is approximately 104 basis points per month (t-statistic = 7.316). In contrast, consistent with the Fama and MacBeth cross-sectional regressions presented in Table 1, the highest excess return is for the $E_t[I/A]$ strategy: approximately -131 basis points per month, with a Newey-West t-statistic of -7.977 and a Sharpe ratio of -1.192. The excess return of $E_t[\varepsilon_{I/A}]$ is approximately -81.7 basis points (t-statistic = -5.210; Sharpe ratio = -0.841), which is smaller than that for $E_t[I/A]$. Later in Section 4, we further investigate the reason for this weaker predictability in $E_t[\varepsilon_{I/A}]$.

We find similar results from evaluating the strategies with five popular asset pricing models, including the capital asset pricing model (CAPM), the Fama and French (1993) three-factor model (FF3), the Carhart (1997) four-factor model (C4), the Hou et al. (2015) q-factor model (Q4), the Fama and French (2015) five-factor model (FF5).[9] For example, the high-minus-low decile portfolio

[9]The Fama–French and momentum factors are from the Kenneth French's Data Library. The Hou et al.'s factors are from Professor Lu Zhang.

formed on $E_t[d^1I/A]$ produces a FF5 alpha of 90.3 basis points (t-statistic = 6.400) and a Q4 alpha of 75.7 basis points per month (t-statistic = 4.897). Similar to the Fama and MacBeth (1973) cross-sectional regressions presented in Table 2, $E_t[I/A]$ continues to exhibit negative return predictability. The high-minus-low strategy produces a FF5 alpha of -89 basis points per month, with a Newey-West t-statistic of -6.806, and a Q4 alpha of -92.7 basis points per month, with a Newey-West t-statistic of -6.018. The equivalent strategy formed on $E_t[\varepsilon_{I/A}]$ produces a FF5 alpha of -69.3 basis points (t-statistic = -4.630) and a Q4 alpha of -75 basis points (t-statistic = -4.538).

4. Further analysis

Thus far, we have shown that the expected I/A growth ($E_t[d^1I/A]$) positively correlates with expected returns, whereas the expected I/A ($E_t[I/A]$) and the implied component of the expected I/A derived from $E_t[d^1I/A]$ ($E_t[\varepsilon_{I/A}]$) negatively predict future returns. In this section, we investigate the potential reasons for this contradiction.

4.1. Mismatch of investment characteristics

We first determine how these variables link to average investment characteristics in extreme deciles and present the results in Table 3. Panel A of Table 3 shows that stocks in the lowest (highest) $E_t[d^1I/A]$ decile are those with high (low) current investment, which explains their positive relation with future returns. More importantly, we can observe that this variable fails to align with true future firm investment. Stocks in the lowest (highest) deciles formed on $E_t[d^1I/A]$ are stocks that invest aggressively (conservatively) in the future.

For the expected I/A measure ($E_t[I/A]$) in Panel C, we find that it aligns well with true future firm investment, which suggests that $E_t[I/A]$ is an appropriate proxy for a firm's future investment. Specifically, stocks in the lowest (highest) $E_t[I/A]$ decile are those with lowest (highest) future investment. The valued-weighted average future investment rises monotonically from 2.32% to 26.4%, when moving from the lowest to the highest $E_t[I/A]$ decile. In addition, stocks with high (low) $E_t[I/A]$ tend to have low $E_t[d^1I/A]$, although the difference between $E_t[d^1I/A]$ in the lowest and highest $E_t[I/A]$ deciles is small and insignificant, with a p-value of 0.891. Similarly, in Panel B, although not rising monotonically as in $E_t[I/A]$, stocks in the highest (lowest) $E_t[\varepsilon_{I/A}]$ decile are those with high (low) future investment.

Taken together, the results demonstrate that the failure of $E_t[d^1I/A]$ to capture the negative expected investment-return relation documented in Tables 1 and 2 is due largely to the fact that it is a poor proxy for future firm investment.

4.2. Estimation of the slope of I/A

Although the mismatch between $E_t[d^1I/A]$ and future firm investment plays an essential role in its prediction of the incorrect sign for the expected investment-return relation, we would like to know the cause of this mismatch. As noted, $E_t[d^1 I/A]$ is constructed as the difference between an implied component linked to expected investment ($E_t[\varepsilon_{I/A}]$) and current investment (I/A). By construction, Hou et al. (2018a) simply assume that current investment is proportional to expected investment with an autoregression coefficient of one. That is, estimating expected investment in this case is equivalent to conducting a constrained regression of $E_t[I/A]$ on I/A, $\log(q)$, and COP, with the constraint that the slope of I/A must be one. To investigate whether this constraint is appropriate, in this subsection, we run the unconstrained regression and then examine whether the coefficient estimates of I/A are around one.

Figure 1 depicts the coefficient estimation of I/A over time, showing that the estimated I/A slopes $\hat{\beta}_{I/A}$ consistently vary approximately 0.2 for the entire sample period and display a slight downward trend, particularly during recent decades. We also test the null hypothesis that the slope coefficient of I/A is one and find that the null can be easily rejected at the 1% significance level (unreported for brevity). Since the results presented in Figure 1 indicate that the true $\beta_{I/A}$ is much lower than one, running a regression by imposing the incorrect restriction that $\beta_{I/A} = 1$ can lead to biased parameter estimates (Greene, 2012). The corresponding coefficient estimates, in turn, bias the fitted dependent variable and the associated $E_t[\varepsilon_{I/A}]$ derived from it, resulting in the positive $E_t[d^1 I/A]$-return relation and the weaker predictability found using $E_t[\varepsilon_{I/A}]$.

In sum, the results indicate that the $\beta_{I/A}$ estimates are indeed much lower than one, and therefore, conducting the regression with the constraint that $\beta_{I/A} = 1$ can lead to misleading parameter estimates and statistical inferences.

5. Concluding remarks

This paper investigates the relation between expected investment and future stock returns and finds that this expected investment-return relation is negative, which is inconsistent with the multiperiod q theory and the findings of Hou et al. (2018a,b) that the associated relation should be positive. Further analysis reveals that there are two potential reasons for this result. First, the expected investment change measure of Hou et al. (2018a) is a poor proxy for firms' true future investment. In fact, stocks in the lowest (highest) decile are those that invest aggressively (conservatively) in the future. Second, the expected investment change estimate is biased because the regression conducted to yield the expected investment changes and the implied component of expected investment is equivalent to the constrained regression, with the constraint that the slope coefficient of I/A is one, which is found to be inconsistent with the empirical evidence. Imposing the incorrect constraint on the regression can lead to biased parameter estimates and, therefore, misleading inferences.

Overall, our findings indicate a negative relation between expect investment and future stock returns, which is inconsistent with the multiperiod q theory. However, we argue that the literature based on mispricing may offer explanations for this negative relation. For example, the negative relation can be attributable to corporate managers' empire-building tendency and investor misreaction (Titman et al., 2004). That is, corporate managers tend to overinvest when they are subject to agency problems. If investors do not fully understand the agency problem of overinvestment when they value stocks, they may overvalue a firm with high expected investments by overestimating its potential future cash flows. Shleifer and Vishny (1997) argue that because of trading frictions, arbitrage can be costly, and mispricing might not be traded away quickly and completely. The low future returns that follow high expected investment therefore reflect the market correction of the initial misreaction. Another potential explanation is based on investors' extrapolation bias (Lakonishok et al., 1994; Cooper et al., 2008). Investors may excessively extrapolate from firms' current and expected growth when they value stocks. Such excessive extrapolation results in the overvaluation of firms with high expected investments and these firms' low subsequent returns. Although it is beyond the scope of this paper, a comprehensive examination of the sources of the negative expected investment-return relation could be an interesting direction for future research.

Acknowledgements

We would like to thank the editor and an anonymous referee for valuable comments. Qi Lin gratefully acknowledges financial support from the National Natural Science Foundation of China (Grant Numbers: 71703142 and 71471161) and the Key Program of the National Natural Science Foundation of China (Grant Number: 71631005).

References

Ball, R., Gerakos, J., Linnainmaa, J.T., Nikolaev, V., 2016. Accruals, cash flows, and operating profitability in the cross section of stock returns. Journal of Financial Economics 121, 28–45.

Carhart, M.M., 1997. On persistence in mutual fund performance. Journal of finance 52, 57–82.

Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross-section of stock returns. Journal of Finance 63, 1609–51.

Davis, J.L., Fama, E.F., French, K.R., 2000. Characteristics, covariances, and average returns: 1929 to 1997. Journal of Finance 55, 389–406.

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

Fama, E.F., French, K.R., 2008. Dissecting anomalies. Journal of Finance 63, 1653–78.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E.F., MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–36.

Greene, W.H., 2012. Econometric analysis, 7th edition. Stern School of Business, New York University .

Hou, K., Mo, H., Xue, C., Zhang, L., 2018a. Motivating factors. Fisher College of Business Working Paper No. 2018-03-03.

Hou, K., Mo, H., Xue, C., Zhang, L., 2018b. Q5. NBER Working Paper No. w24709.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. Journal of Finance 49, 1541–78.

Liu, L.X., Whited, T.M., Zhang, L., 2009. Investment-based expected stock returns. Journal of Political Economy 117, 1105–39.

Newey, W.K., West, K., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–8.

Shleifer, A., Vishny, R.W., 1997. The limits of arbitrage. Journal of Finance 52, 35–55.

Shumway, T., 1997. The delisting bias in CRSP data. Journal of Finance 52, 327–40.

Shumway, T., Warther, V.A., 1999. The delisting bias in CRSP's Nasdaq data and its implications for the size effect. Journal of Finance 54, 2361–79.

Titman, S., Wei, K.J., Xie, F., 2004. Capital investments and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.

Table 1: Fama and MacBeth (1973) Cross-Sectional Regressions

The table reports average Fama and MacBeth (1973) regression slope coefficients (multiplied by 100) on three measures (EInv), which are the expected investment changes of Hou et al. (2018a) ($E_t[d^1I/A]$), the implied expected investment component derived from $E_t[d^1I/A]$ ($E_t[\varepsilon_{I/A}]$), and the expected investment ($E_t[I/A]$), and their associated t-statistics from cross-sectional regressions that predict monthly returns. The regressions include the following control variables: the market beta estimated using monthly returns over the prior 60 months and requiring a minimum of 36 valid observations, $\hat{\beta}_{Mkt}$; the log-market value of equity, log(ME); the log-book-to-market, log(B/M); prior one-year return skipping a month, $R_{12,2}$; prior one-month return, $R_{1,0}$; and quarterly return on equity, ROEQ, which is defined as income before extraordinary items (Compustat quarterly item IBQ) scaled by the one-quarter-lagged book equity. Standardized regressors are calculated as the difference between the variables and their cross-sectional means, divided by the cross-sectional standard deviations. All the variables are winsorized at the 1st and 99th percentiles. The regressions are estimated monthly using data from July 1967 through December 2017. For those including ROEQ as a control variable, the regressions are estimated monthly using data from January 1972 through December 2017. The Newey and West (1987) heteroscedasticity- and autocorrelation-robust t-statistics (with a lag of 4) are given in brackets. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	$E_t[d^1I/A]$				$E_t[\varepsilon_{I/A}]$				$E_t[I/A]$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
EInv	0.360***	0.384***	0.385***	0.303***	-0.329***	-0.213***	-0.229***	-0.266***	-0.458***	-0.362***	-0.387***	-0.381***
	[5.732]	[10.617]	[10.707]	[9.232]	[-7.693]	[-7.485]	[-8.240]	[-9.179]	[-8.770]	[-9.917]	[-10.312]	[-9.650]
$\hat{\beta}_{Mkt}$		0.000	-0.005	0.012		-0.022	-0.023	0.009		0.007	0.005	0.033
		[0.006]	[-0.066]	[0.145]		[-0.270]	[-0.289]	[0.111]		[0.096]	[0.068]	[0.411]
log(ME)		-0.355***	-0.322***	-0.369***		-0.220***	-0.191**	-0.277***		-0.212**	-0.178**	-0.256***
		[-4.435]	[-4.141]	[-4.887]		[-2.615]	[-2.357]	[-3.590]		[-2.450]	[-2.124]	[-3.241]
log(B/M)		0.110***	0.135***	0.180***		0.118***	0.141***	0.166***		0.002	0.019	0.054
		[3.013]	[3.541]	[4.319]		[3.186]	[3.654]	[4.071]		[0.050]	[0.562]	[1.471]
$R_{12,2}$			0.244***	0.168**			0.258***	0.159**			0.247***	0.151**
			[3.745]	[2.437]			[3.972]	[2.312]			[3.765]	[2.178]
$R_{1,0}$			-0.789***	-0.780***			-0.780***	-0.785***			-0.791***	-0.794***
			[-11.540]	[-10.671]			[-11.431]	[-10.728]			[-11.569]	[-10.834]
ROEQ				0.341***				0.432***				0.403***
				[9.078]				[10.884]				[10.158]
Constant	1.191***	1.267***	1.268***	1.301***	1.236***	1.265***	1.265***	1.303***	1.236***	1.265***	1.265***	1.303***
	[4.173]	[4.601]	[4.601]	[4.589]	[4.410]	[4.599]	[4.599]	[4.596]	[4.410]	[4.599]	[4.599]	[4.596]
Adj. R^2	0.610%	3.125%	4.447%	4.638%	0.336%	3.008%	4.346%	4.611%	0.612%	3.122%	4.457%	4.718%

Table 2: Portfolios of Stocks Sorted by $E_t[d^1I/A]$, $E_t[\varepsilon_{I/A}]$, and $E_t[I/A]$

For each month, equal-weighted decile portfolios are formed by sorting individual stocks based on three measures, which are the expected investment changes of Hou et al. (2018a) ($E_t[d^1I/A]$), the implied expected investment component derived from $E_t[d^1I/A]$ ($E_t[\varepsilon_{I/A}]$), and the expected investment ($E_t[I/A]$). The table reports the average excess returns (RET–RF) and risk-adjusted returns (alphas) for the decile portfolios and the zero-investment High–Low decile portfolios formed on three investment measures, respectively. The factor models include the Capital Asset Pricing Model (CAPM), the Fama and French (1993) three-factor model (FF3), the four-factor model augmenting the Fama and French (1993) three-factor model with a Carhart (1997) conventional momentum factor (C4), the Fama and French (2015) five-factor model (FF5), and the Hou et al. (2015) q-factor model (Q4). Newey and West (1987) heteroscedasticity- and autocorrelation-robust t-statistics (with a lag of 4) are given in brackets. The last row reports the annualized Sharpe ratios (SR) and information ratios (IR), respectively. The sample period is July 1967–December 2017.

	$E_t[d^1I/A]$						$E_t[\varepsilon_{I/A}]$						$E_t[I/A]$					
Decile	RET–RF	CAPM	FF3	C4	FF5	Q4	RET–RF	CAPM	FF3	C4	FF5	Q4	RET–RF	CAPM	FF3	C4	FF5	Q4
1 (Low)	0.233	-0.466	-0.555	-0.263	-0.250	0.039	1.051	0.394	0.162	0.408	0.373	0.683	1.431	0.857	0.535	0.718	0.593	0.844
	[0.636]	[-2.217]	[-3.957]	[-1.574]	[-1.421]	[0.167]	[2.803]	[1.677]	[1.016]	[2.259]	[2.170]	[3.200]	[4.395]	[4.160]	[4.297]	[5.136]	[4.493]	[5.858]
2	0.622	-0.023	-0.210	-0.001	-0.123	0.091	1.069	0.497	0.222	0.399	0.234	0.426	1.216	0.646	0.349	0.519	0.360	0.577
	[2.074]	[-0.156]	[-2.425]	[-0.009]	[-1.121]	[0.668]	[3.677]	[3.013]	[2.395]	[3.700]	[2.159]	[3.174]	[4.023]	[3.603]	[3.284]	[4.328]	[3.059]	[4.190]
3	0.818	0.194	-0.007	0.171	0.001	0.183	1.016	0.452	0.212	0.377	0.217	0.411	1.133	0.553	0.264	0.427	0.270	0.457
	[2.817]	[1.313]	[-0.088]	[1.847]	[0.009]	[1.522]	[3.807]	[3.198]	[2.615]	[3.862]	[2.059]	[2.980]	[4.021]	[3.581]	[2.995]	[4.208]	[2.601]	[3.578]
4	0.851	0.250	0.026	0.176	0.004	0.173	0.940	0.385	0.178	0.308	0.158	0.300	1.025	0.460	0.217	0.357	0.203	0.381
	[3.110]	[1.903]	[0.363]	[2.634]	[0.046]	[1.745]	[3.771]	[2.978]	[2.635]	[4.091]	[2.010]	[2.771]	[3.758]	[3.070]	[2.830]	[3.921]	[2.310]	[3.490]
5	0.960	0.371	0.156	0.297	0.124	0.279	0.953	0.399	0.206	0.301	0.169	0.272	0.961	0.387	0.170	0.294	0.166	0.304
	[3.606]	[2.816]	[2.328]	[3.886]	[1.771]	[3.140]	[3.867]	[3.256]	[3.253]	[4.436]	[2.362]	[2.940]	[3.656]	[2.888]	[2.525]	[3.915]	[2.215]	[3.395]
6	0.977	0.397	0.199	0.320	0.176	0.308	0.996	0.440	0.264	0.356	0.233	0.355	0.875	0.290	0.109	0.242	0.097	0.262
	[3.719]	[3.081]	[2.976]	[4.528]	[2.178]	[2.923]	[4.042]	[3.634]	[4.264]	[5.734]	[3.664]	[4.466]	[3.343]	[2.181]	[1.566]	[3.037]	[1.251]	[2.570]
7	1.001	0.427	0.251	0.380	0.211	0.355	0.796	0.226	0.078	0.179	0.040	0.143	0.785	0.190	0.030	0.142	0.030	0.165
	[3.931]	[3.358]	[3.527]	[5.059]	[2.551]	[3.417]	[3.293]	[2.058]	[1.341]	[2.905]	[0.593]	[1.525]	[2.946]	[1.459]	[0.458]	[1.958]	[0.443]	[1.992]
8	1.122	0.551	0.380	0.499	0.344	0.480	0.917	0.317	0.198	0.307	0.197	0.332	0.738	0.117	0.010	0.174	0.050	0.215
	[4.458]	[4.414]	[6.006]	[7.281]	[5.262]	[6.016]	[3.571]	[2.684]	[3.250]	[4.600]	[2.941]	[3.990]	[2.714]	[0.915]	[0.136]	[2.116]	[0.541]	[1.690]
9	1.170	0.594	0.442	0.555	0.415	0.559	0.723	0.079	0.009	0.161	0.077	0.227	0.568	-0.081	-0.128	0.027	-0.058	0.094
	[4.585]	[4.789]	[7.084]	[8.333]	[6.150]	[6.283]	[2.633]	[0.643]	[0.139]	[2.222]	[0.964]	[2.100]	[2.016]	[-0.651]	[-1.761]	[0.338]	[-0.661]	[0.825]
10 (High)	1.271	0.669	0.581	0.688	0.653	0.795	0.233	-0.486	-0.517	-0.253	-0.320	-0.067	0.120	-0.611	-0.536	-0.277	-0.297	-0.083
	[4.606]	[4.606]	[7.511]	[7.664]	[7.409]	[6.661]	[0.729]	[-3.172]	[-5.201]	[-2.519]	[-2.380]	[-0.396]	[0.370]	[-3.861]	[-5.705]	[-2.411]	[-2.218]	[-0.440]
High–Low	1.038	1.134	1.136	0.951	0.903	0.757	-0.817	-0.880	-0.679	-0.661	-0.693	-0.750	-1.311	-1.468	-1.071	-0.995	-0.890	-0.927
	[7.316]	[8.767]	[9.258]	[7.401]	[6.400]	[4.897]	[-5.210]	[-5.613]	[-4.431]	[-4.353]	[-4.630]	[-4.538]	[-7.977]	[-8.588]	[-8.824]	[-7.165]	[-6.806]	[-6.018]
SR(IR)	1.189	1.348	1.421	1.254	1.204	1.024	-0.841	-0.916	-0.789	-0.768	-0.882	-0.961	-1.192	-1.424	-1.412	-1.323	-1.237	-1.171

Table 3: Average investment characteristics

The table reports the value-weighted average current and future investment characteristics (I/A_t and I/A_{t+1}) along with the expected investment changes of Hou et al. (2018a) ($E_t[d^1I/A]$), the implied expected investment component derived from $E_t[d^1I/A]$ ($E_t[\varepsilon_{I/A}]$), and the expected investment ($E_t[I/A]$) in each decile used to form the High–Low hedge returns. We also conduct the mean-comparison t-test for the null hypothesis that the average investment characteristics in the lowest and highest deciles (deciles 1 and 10) have the same mean but unknown variance, and report the associated p-values in the last row of each panel. The sample is annual from 1966 to 2016 and includes all NYSE/Amex/NASDAQ common stocks (excluding financial stocks and stocks with negative book equity).

	Panel A: $E_t[d^1I/A]$				Panel B: $E_t[\varepsilon_{I/A}]$				Panel C: $E_t[I/A]$				
Decile	$E_t[d^1I/A]$	I/A_t	I/A_{t+1}	$E_t[\varepsilon_{I/A}]$	$E_t[\varepsilon_{I/A}]$	I/A_t	I/A_{t+1}	$E_t[d^1I/A]$	$E_t[I/A]$	I/A_t	I/A_{t+1}	$E_t[d^1I/A]$	$E_t[\varepsilon_{I/A}]$
1	-0.137	0.553	0.277	0.302	-0.212	-0.150	0.066	-0.065	-0.006	-0.057	0.023	-0.015	-0.069
2	-0.081	0.400	0.155	0.216	-0.075	-0.032	0.057	-0.043	0.031	-0.010	0.051	-0.022	-0.031
3	-0.064	0.178	0.162	0.110	-0.033	0.003	0.068	-0.036	0.051	0.023	0.053	-0.025	-0.003
4	-0.051	0.163	0.124	0.106	-0.001	0.030	0.074	-0.030	0.067	0.038	0.068	-0.026	0.011
5	-0.040	0.140	0.108	0.092	0.028	0.050	0.088	-0.022	0.083	0.063	0.080	-0.021	0.042
6	-0.030	0.115	0.103	0.085	0.058	0.077	0.112	-0.019	0.099	0.078	0.101	-0.027	0.051
7	-0.019	0.108	0.107	0.088	0.094	0.103	0.137	-0.009	0.117	0.094	0.123	-0.017	0.077
8	-0.006	0.112	0.114	0.105	0.140	0.138	0.142	0.002	0.140	0.121	0.124	-0.015	0.106
9	0.014	0.121	0.124	0.133	0.222	0.216	0.184	0.007	0.173	0.165	0.149	-0.007	0.157
10	0.061	0.140	0.183	0.201	0.632	0.749	0.292	-0.012	0.269	0.503	0.264	-0.014	0.434
Spread$(10-1)$	0.198	-0.413	-0.094	-0.101	0.844	0.899	0.226	0.053	0.276	0.559	0.241	0.001	0.503
p(Spread)	0.000	0.000	0.097	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.891	0.000

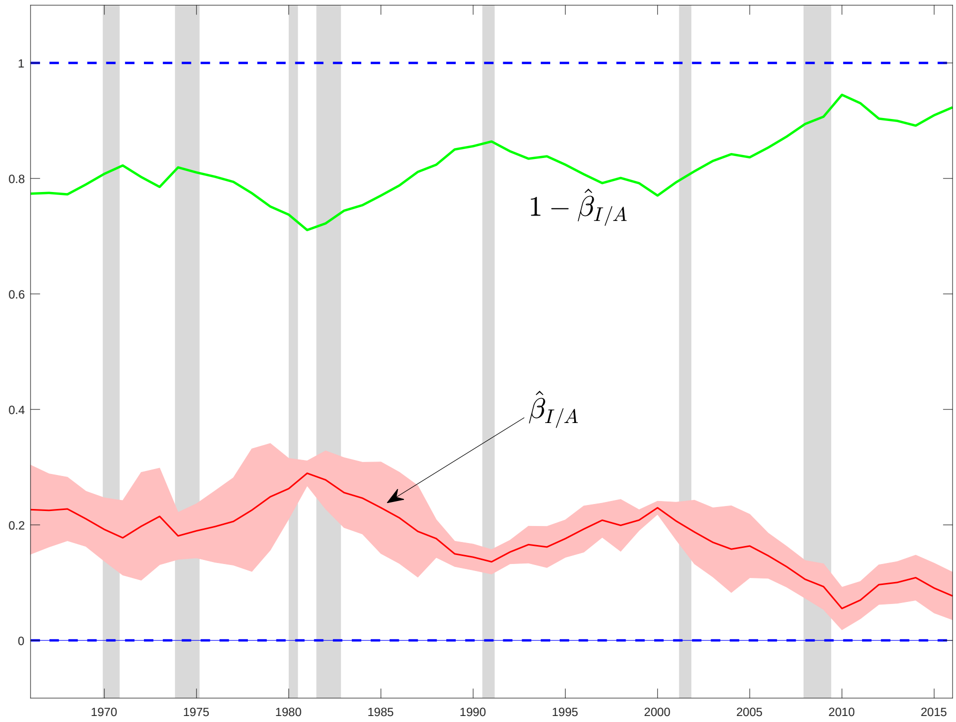

Figure 1: The $\beta_{I/A}$ estimates and the difference between 1 and $\hat{\beta}_{I/A}$ over time.

The red lines is the $\beta_{I/A}$ estimates ($\hat{\beta}_{I/A}$) while the associated shaded areas represent the 95% confidence intervals. The green line is the difference between 1 and $\hat{\beta}_{I/A}$. The vertical bars depict NBER-defined recessions. The sample is annual from 1966 to 2016 and includes all NYSE/Amex/Nasdaq common stocks (excluding financial stocks and stocks with negative book equity).

Highlights

- The expected investment-return relation is negative and inconsistent with the multiperiod q theory.
- The expected investment change measure is not an appropriate proxy for future investment.
- The mismatch of investment characteristics and the incorrect constraint imposed on the regression are two potential reasons.