What Drives Marginal Q and Investment Fluctuations?

Time-Series and Cross-Sectional Evidence

Ilan Cooper Paulo Maio Chunyu Yang[1]

This version: March 2022[2]

[1]Cooper, ilan.cooper@bi.no, Department of Finance, Norwegian Business School (BI); Maio, paulo.maio@hanken.fi, Department of Finance and Economics, Hanken School of Economics; Yang, chunyu.yang@bi.no, Department of Finance, Norwegian Business School (BI).

[2]We thank Nick Roussanov and seminar participants at Ben Gurion University (Department of Management) and the 2018 European Economic Association Meeting (Cologne). Previous versions circulated with the titles "What drives Q and investment fluctuations?", "What drives marginal Q fluctuations?", and "What drives marginal Q fluctuations? time-series and cross-sectional evidence". Any remaining errors are our own.

Abstract

We explore whether marginal Q and investment fluctuate due to revisions in expected marginal profits or discount rates, and by how much of each. We infer marginal Q from the marginal cost of investment, derive a present-value relation, and conduct a VAR-based variance decomposition for marginal Q. We find that discount rates (expected investment returns) drive the bulk of fluctuations in average Q and investment in the time series, but play no role in driving the cross-section of portfolios' average Q and investment. That is, marginal profits are the sole determinant of the cross-section of marginal Q and investment.

Keywords: Tobin's q; Investment; Present-value model; Investment return; Variance decomposition; VAR implied predictability; Aggregation bias; Marginal profit of capital; cross-section of q

JEL classification: E22; E27; G10; G12; G17; G31

1 Introduction

Marginal Q is the present value of all future marginal profits entailed by installing an extra unit of capital. Thus, similar to stock prices, variations in marginal Q are driven by shocks to expected cash flows (expected marginal profits in the case of marginal Q) as well as by discount rate shocks (shocks to expected investment returns in the case of marginal Q). That aggregate investment varies substantially over the business cycle implies that the marginal value of capital is also highly volatile. In the time series, aggregate investment (and consequently also aggregate marginal Q) negatively predicts stock returns (Cochrane, 1991). Simultaneously aggregate investment and marginal Q should also be related to future marginal profits. In the cross-section, firms with low investment (implying low marginal Q) have on average higher subsequent stock returns than firms with high investment (Anderson and Garcia-Feijoo, 2006, Xing, 2008, and Cooper, Gulen, and Schill, 2008). Cross-sectional differences in investment (and marginal Q) should, however, also forecast cross-sectional differences in subsequent marginal profits. The relative importance of these two contributing factors to the time-series and cross-sectional variations of marginal Q and investment remains largely open empirical questions.

Exploring these questions is important for several reasons. First, investment is an immensely important macroeconomic variable as it facilitates economic growth. Second, investment varies substantially over the business cycle, and more so than output or consumption. Third, efficient allocation of capital is very important for the economy. Studying the cross-sectional determinants of investment and marginal Q contributes to our understanding of resource allocation in the economy. Fourth, the exploration informs us of the drivers of managers' valuations of their firms, as opposed to studying the sources of fluctuations of investors' valuations of these firms (which is obtained by examining the sources of fluctuations in equity valuation ratios, such as the dividend-to-price ratio, or the book-to-market ratio).

Extensive literature studies the sources of fluctuations in (scaled) stock prices (for example, Cochrane, 2008, 2011). In contrast, the determinants of variations in marginal Q are relatively unexplored. Our contribution is twofold. Our first contribution is to explore and quantify the roles of the sources of fluctuations in the time-series of realized marginal Q at both the aggregate level and at the portfolio level. In our setting, marginal Q is a linear function of investment. Thus, by exploring the sources of fluctuations in marginal Q we also explore the sources of variation in investment. The observation that time-series variation in marginal Q and investment is driven by discount rate fluctuations is not new. For example, Cochrane (1991) shows that the aggregate investment-to-capital ratio negatively predicts future stock market returns at quarterly and annual horizons. This finding suggests that investment responds to discount rate variations at least to some extent, implying that the marginal value of capital also varies due to changes in discount rates. The contribution of our analysis is that by using a structural model of investment, a present-value relation, and a VAR-based variance decomposition, we are able to measure exactly how much the time-series variation in expected marginal profits and in expected investment returns contribute each to the time-series variation of marginal Q and investment at both the aggregate level as well as at the portfolio level.

Our second contribution concerns the cross-sectional variation of portfolios' average marginal Q and average investment. We derive a decomposition for the cross-section of average marginal Q (and given the linear relation between marginal Q and investment, also for the cross-section of average investment). This relation implies that portfolios with high average marginal Q (and average investment) should have higher average marginal profits and/or lower average investment returns. Relatedly, there is a vast literature examining the investment-cash flow sensitivity (starting with Fazzari, Hubbard, and Petersen, 1988). Marginal profits are somewhat related to cash flows and hence the relation of that literature to our paper. That literature does not explore the relative roles of average marginal profits and average investment returns in driving the cross-sectional variation in average investment and average marginal Q. To the best of our knowledge our paper is the first to quantify the roles of cash flows (average marginal profits) and discount rates (average investment returns) in driving the cross-section of portfolios' average investment and average marginal Q.

In order to study the roles of expected marginal profit shocks and discount rate shocks in driving marginal Q, we employ a classical model of optimal investment (as in Liu, Whited, and Zhang, 2009) and derive a present-value relation for the model-implied marginal Q. In the model, the firm's optimal investment behavior implies that marginal Q, that is, the marginal value of capital, is equal to the marginal cost of investment. Assuming a (standard) functional form for the adjustment cost function implies that the marginal cost of investment is a linear function of the investment-to-capital ratio. Belo, Xue, and Zhang (2013) use the same model to infer marginal Q from investment data. Thus, our approach is the supply approach to valuation, as in Belo *et al.* (2013) who use the supply side, that is, the firm's optimization conditions, to identify marginal Q. Whereas stock prices are determined by investors in the stock market, investment is undertaken by the firm's manager. Thus, our approach of using investment data to infer marginal Q from the marginal adjustment cost of investment implies that we are studying the marginal value that firms' managers, not necessarily investors, attribute to an extra unit of capital.

In the estimation of the model's parameters, namely the share of capital in profit and the adjustment cost parameter, we follow Gonçalves, Xue, and Zhang (2020) and correct for aggregation bias when conducting the GMM estimation (see also Belo, Gala, Salomao, and Vitorino, 2022). That is, we estimate the model's parameters by using firm-level data to match two moment conditions. First, we match portfolio-level stock returns to a weighted average of firm-level levered investment returns (the investment return moment). Second, we match weighted average Tobin's marginal Q in the data to a weighted average model-implied Tobin's Q (the valuation moment). Following the parameter estimation, we construct time series of marginal Q, marginal profits, and investment returns at the aggregate level.

Armed with the estimated time-series for investment returns and its components, we derive a dynamic present-value relation for the log marginal Q (q), in which q is positively correlated with both future multi-period log marginal profits and the future q at some terminal date, and negatively correlated with future multi-period log investment returns. This present-value relation is analogous to the present-value relation associated with the log stock price derived in Campbell, Lo, and MacKinlay (1997) and is similar to the present-value relation derived in Lettau and Ludvigson (2002). Such relationship gives rise to a variance decomposition for q at each forecasting horizon, which contains the fractions of the variance of current q attributed to the predictability of future investment returns, marginal profits, and q.

We use two methods to estimate empirically the variance decomposition for q: first-order restricted VAR (as in Cochrane, 2008), and first-order unrestricted VAR (as in Larrain and Yogo, 2008 and Maio and Xu, 2020). For aggregate marginal Q (and thus, aggregate investment) the two methods produce qualitatively similar variance decompositions at both intermediate and long horizons. Specifically, the bulk of variation in q turns out to be investment return predictability, with predictability of the marginal profitability of capital assuming a secondary role (although statistically significant when the decomposition is based on the restricted VAR). On the other hand, predictability of future q only plays a relevant role at very short horizons.[1] To have an idea of these predictability patterns, the shares of long-run (infinite horizon) investment return (marginal profits) predictability (at the aggregate level) are 0.61 (0.37) and 0.80 (0.18) under the restricted and unrestricted VAR approaches, respectively. In fact, in the case of the variance decomposition based on the unrestricted VAR, we cannot reject the null that *all* the variation in q stems from long-run investment return predictability.

These findings are robust to a host of robustness checks. Specifically, the findings are robust to using median stock returns and investment returns instead of value-weighted stock returns and investment returns, as well as to conducting the GMM estimation of the structural investment model based on decile portfolios sorted by marginal Q. Our findings are also robust to using a bootstrap simulation (based on the restricted VAR), which represents an alternative statistical inference for the implied horizon-specific predictive slope estimates (that complements the standard asymptotic inference). Further, we obtain qualitatively similar results by estimating the variance

[1] This finding is consistent with Eberly, Rebelo, and Vincent (2012) that lagged investment is a strong predictor of current investment.

decomposition for q based on long-horizon regressions (direct approach) rather than relying on the first-order VAR (indirect approach). The findings are also robust to varying the values of the capital depreciation parameter. As an additional robustness check, we re-estimate the technology parameters using the methodology in Liu, Whited, and Zhang (2009). This methodology does not account for aggregation bias. Reassuringly, the long-run predictability mix is qualitatively similar to that in our benchmark setting, even though the structural parameter estimates differ to some degree: the discount rate channel explains most (71%) of the variation in q at the aggregate level.

Given that the share of capital in production and the adjustment cost parameters are estimated with errors, we conduct comparative statics by experimenting with several possible values of these parameters. We test several combinations of these parameters such that the fraction of adjustment costs in output varies in the range of 0% to 20% (the range surveyed in Bloom, 2009). Importantly, we use the same values of investment and sales as in the data, but varying the parameters yields different series of marginal profits, q, and investment returns. We find the following. First, the share of investment return (marginal profits) predictability declines (increases) with the share of capital parameter. Intuitively, if the share of capital is zero, then marginal profits are constant at zero and are not predictable at all. As the share of capital rises, the relative predictability of marginal profits versus investment returns rises. Second, the shares of return (marginal profits) predictability tend to increase (decrease) monotonically with the adjustment costs parameter. Intuitively, if adjustment costs are zero, marginal Q equals to one at each point in time, implying that investment returns are less time-varying. Overall, for a wide range of plausible parameter values the bulk of the variation in q is driven by investment return predictability.

We choose to explore the determinants of the time-series fluctuations not only for aggregate marginal Q (and aggregate investment), but also for portfolio-level marginal Q (and investment). Our purpose in doing so is to examine whether the results at the aggregate level carry over to a more disaggregated level.[2] Hence, we extend the time-series variance decomposition analysis for equity

[2]Relatedly, at the aggregate portfolio level, the equity discount rate channel dominates the cash-flow channel (Campbell, 1991, Campbell and Ammer, 1993, and Cochrane, 2008, 2011). However, at the individual firm level,

portfolios sorted on the book-to-market ratio, asset growth, and operating profitability. These equity characteristics are associated with some of the major anomalies in asset pricing. Each group of decile portfolios (e.g., asset growth deciles) has a different GMM estimation of the structural investment model and thus a different set of estimates for the two key investment parameters. Such analysis enables us to put in perspective the evidence obtained for the aggregate portfolio. Overall, the results suggest that what drives the variation in the q of different equity portfolio groups is the investment return channel, with the cash-flow channel being marginal in most cases. Hence, this result is consistent with the evidence obtained for the market portfolio. However, there are some important differences on the long-run predictability mix within the cross-section of firms. In particular, predictability of future marginal profits plays a critical role in explaining the variation of q for low-profitability firms.

In the last part of the paper, we estimate a variance decomposition for average q in the cross-section of the 30 equity portfolios, which represents an important part of our analysis.[3] Under the cross-sectional decomposition, firms/portfolios that have a higher average q tend to register a higher average marginal profitability of capital and/or a lower average investment return. Hence, similarly to the dynamic present-value relation, there is a positive (negative) correlation between q and marginal profitability (investment return). In order to obtain the shares in the cross-sectional variance decomposition, we estimate cross-sectional regressions of average return or average marginal profits onto average q. The results show that the cash-flow channel is the source of cross-sectional variation in average q, that is, firms with higher average q are firms with higher average marginal profitability of capital, instead of being firms with lower average investment return. In fact, average investment return is positively correlated with average q, which is inconsistent with the prediction

Vuolteenaho (2002) finds that individual stock returns are mainly driven by cash-flow news. Maio and Santa-Clara (2015) find that in contrast to the aggregate stock market portfolio, for portfolios of small and value stocks dividend yields are mainly related to future dividend changes and less so to future returns.

[3] We choose to study the cross-section of marginal Q of portfolios sorted on characteristics known to generate large spreads in stock returns. In the model we employ, (levered) investment returns are equal to stock returns. Hence, our use of these portfolios provides a potentially large spread in investment returns increasing the power of the tests relating average investment returns to average marginal Q.

of the cross-sectional decomposition. This implies that from an economic viewpoint all the cross-sectional variation in average q stems from the cash-flow channel. This result is in clear contrast with the findings for the time-series analysis, in which the discount-rate channel explains most of the variation of q (associated with either the market or equity portfolios) over time.

Related to our cross-sectional tests, Cochrane (2011) finds mixed evidence on the driving forces of the cross-sectional variation of the dividend-to-price ratios of portfolios sorted on market capitalization or the book-to-market ratio.[4] In related work, Cohen, Polk, and Vuolteenaho (2003) decompose the cross-sectional variance of firms' book-to-market ratios. They find that in contrast to aggregate time-series results, variation in expected stock returns causes only a relatively small fraction of the cross-sectional variance of book-to-market ratios. The bulk of the dispersion in book-to-market ratios is explained by expected profitability spreads, as well as the persistence of valuation ratios. This result resembles our cross-sectional result. Cohen *et al.* (2003) note that "If the cross section of valuation ratios is largely driven by rational cash-flow expectations, the conclusion that aggregate valuation ratios are exclusively driven by irrational investor sentiment is perhaps premature." Similarly, our result that the cross-section of marginal Q and investment is driven solely by average marginal profit spreads implies that managers' investment decisions are closely related to economic fundamentals.

The paper most closely related to ours is Abel and Blanchard (1986), who also analyze the empirical implications of a present-value relation for marginal Q (in levels) at the aggregate level. The main goal of Abel and Blanchard (1986) is to compute a marginal Q measure (i.e. the present value of future marginal profits) by using a VAR to extract expected marginal profits and discount rates and understand how its two components (namely expected future marginal profits and discount rates) drive its variability over time. Abel and Blanchard (1986) then relate the marginal Q series to investment in a separate step. They conclude that variations in marginal Q are due more to discount rate variations than to variations in expected marginal profits. They also find that the

[4] As shown in his Table AIV (first two columns), the average stock return channel is dominant for the book-to-market portfolios, while the average dividend growth channel is dominant for the size portfolios.

marginal profits component has a larger and more significant effect on investment than the cost of capital component.

Among several other differences in the empirical designs of the two studies, there are several major differences. First, Abel and Blanchard extract from a VAR estimation both discount rates and expected marginal profits from which they compute a present-value series (i.e. a time-series of the marginal values of capital). They then compute the standard deviations of the two components of Q and find that discount rates vary more. This leads them to conclude that discount rates contribute more to the variation of marginal Q than expected marginal profits. However, they do not conduct a proper variance decomposition in the sense that they do not calculate the weights associated with the variance and covariance terms, as well as the corresponding standard errors.[5] Thus, they are not able to formally quantify how important shocks to discount rates and expected profits are in terms of driving the dynamics of marginal Q. That is, their results cannot inform us of the statistical and economic significance of the driving forces of Q. In contrast, our variance decomposition informs us how much of the variation in q is due to predictability of investment returns and how much is due to predictability of marginal profits, while also quantifying the statistical significance of these effects. Hence, our paper focuses on the predictive information in the log marginal Q for future investment returns and marginal profits, something that is not addressed in Abel and Blanchard (1986). Second, Abel and Blanchard (1986) study the aggregate time-series level of marginal Q and investment. However, in this study we also explore the determinants of the time-series and cross-sectional fluctuations in portfolio-level marginal Q and investment, which represents an important analysis in the paper.

Lettau and Ludvigson (2002) use a dynamic present-value relation for q to motivate their empirical design, in which traditional predictors of the equity premium (such as the dividend yield, term spread, or default spread) are used to forecast future aggregate investment growth (controlling

[5] See, for example, Campbell (1991), Campbell and Ammer (1993), Maio (2014), and Guo, Kontonikas, and Maio (2020), for properly defined variance decompositions (based on the variances and covariances of the components) associated with stock and bond returns.

for investment-based variables). There are two key differences between the two papers. First, Lettau and Ludvigson (2002) do not compute a variance decomposition for q and hence they do not test the relative importance of discount rate shocks and expected marginal profits shocks in driving q. Second, the investment-based predictors that they employ in forecasting investment growth (such as the average Q or profit growth) are not directly obtained from a structural model of investment.[6] Yashiv (2016) estimates a variance decomposition of the aggregate log marginal Q-to-productivity ratio. The present-value and variance decomposition associated with that variable are very distinct from those associated with q. This implies that our time-series evidence for the aggregate portfolio is not directly comparable to Yashiv's results.[7]

It is important to note that the discussion in the literature regarding the investment-cash flow sensitivity is irrelevant for our purpose. That is, the present value relation we derive holds regardless of whether or not firms' investment is sensitive to contemporaneous cash flows. According to our present-value relation, the model-implied marginal Q (which is a linear function of the investment to capital ratio) must always forecast future values of the marginal profit of capital, future investment returns, or both.

The rest of the paper is organized as follows. In Section 2, we present a model of a firm's optimal investment decisions. Section 3 describes the data and the econometric methodology for estimating the production and adjustment costs parameters and the components of investment returns. Section 4 provides the results for the time-series variance decomposition for the aggregate portfolio, while Section 5 contains a sensitivity analysis. In Sections 6 and 7, we conduct respectively time-series and cross-sectional variance decompositions for equity portfolios. The paper concludes in Section 8.

[6] Chen, Da, and Larrain (2016) find that discount rates play no role in driving shocks to the time-series of aggregate investment growth. Importantly, they study a present-value relation that pertains to total investment, which includes investment in liquid assets such as cash and cash equivalent assets. Most likely, investment in cash does not incur adjustment costs (see, for example, Gonçalves *et al.*, 2020) and therefore the cash component should not forecast future returns or cash flows. This renders their findings difficult to interpret.

[7] Moreover, Yashiv (2016) relies on a restricted VAR, which is likely to yield a severe misspecification on the estimated long-run variance decomposition (see the discussion in Cooper, Maio, and Yang, 2021).

2 The Background Model

In this section, we provide the details of the underlying structural investment model.

2.1 Theoretical model

We employ the model in Liu, Whited, and Zhang (2009) in order to derive our variables of interest. The firm is assumed to have linearly homogenous production function and adjustment cost function. The factors of production are capital, as well as costlessly adjustable inputs, such as labor. The firm is a price taker, and in each period chooses optimally the costlessly adjustable inputs to maximize operating profits, defined as revenues minus the cost of the costlessly adjustable inputs. Taking operating profits as given, the firm chooses optimal investment and debt to maximize the value of equity.

Let $\Pi\left(K_{i,t}, X_{i,t}\right)$ denote the maximized operating profits of firm i at time t, where K is the stock of capital and X is a vector of aggregate and idiosyncratic shocks. The firm is assumed to have a Cobb-Douglas production function with constant returns to scale. The marginal operating profit of capital is given by

$$\frac{\partial \Pi\left(K_{i,t}, X_{i,t}\right)}{\partial K_{i,t}} = \alpha \frac{Y_{i,t}}{K_{i,t}}, \tag{1}$$

where $\alpha > 0$ is the share of capital and Y is sales.

We assume standard quadratic functional form for the adjustment cost function,

$$\Phi\left(I_{i,t}, K_{i,t}\right) = \frac{a}{2}\left(\frac{I_{i,t}}{K_{i,t}}\right)^2 K_{i,t}, \tag{2}$$

where $a > 0$ is the adjustment cost parameter. Taxable profits equal operating profits minus capital depreciation minus interest expenses.

The firm's investment return is given by

$$R_{i,t+1} = \frac{(1-\tau_{t+1})\left[\alpha \frac{Y_{i,t+1}}{K_{i,t+1}} + \frac{a}{2}\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)^2\right] + \tau_{t+1}\delta + (1-\delta)\left[1 + (1-\tau_{t+1})\, a \frac{I_{i,t+1}}{K_{i,t+1}}\right]}{\left[1 + (1-\tau_t)\, a \frac{I_{i,t}}{K_{i,t}}\right]}. \quad (3)$$

The marginal value of an additional unit of capital appears in the numerator, whereas the marginal cost of investment is in the denominator. $(1-\tau_{t+1})\,\alpha Y_{i,t+1}/K_{i,t+1}$ is the after-tax marginal operating profit of capital, $(1-\tau_{t+1})\,(a/2)\,(I_{i,t+1}/K_{i,t+1})^2$ is the after-tax marginal reduction in adjustment costs in period $t+1$ that stems from the existence of an extra unit of capital installed in period t, $\tau_{t+1}\delta$ is the marginal depreciation tax shield, and $(1-\delta)\,[1 + (1-\tau_{t+1})\,a\,(I_{i,t+1}/K_{i,t+1})]$ is the marginal value at $t+1$ of the undepreciated part of the unit of capital installed in period t (which under optimal investment at $t+1$ is equal to the marginal cost of investment at $t+1$).

We define marginal profit of capital, M, as follows:

$$M_{i,t+1} \equiv (1-\tau_{t+1})\left[\alpha \frac{Y_{i,t+1}}{K_{i,t+1}} + \frac{a}{2}\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)^2\right] + \tau_{t+1}\delta. \quad (4)$$

Thus, the marginal profit of capital is the sum of the after-tax marginal operating profit of capital and reduction in adjustment costs due to the existence of the extra unit of capital, plus the depreciation shield. Optimal investment entails equating the marginal value of capital (Q) to the marginal cost of investment. Hence, optimal investment behavior implies that

$$Q_{i,t+1} = 1 + (1-\tau_{t+1})\, a \frac{I_{i,t+1}}{K_{i,t+1}}. \quad (5)$$

Therefore, the investment return for firm i can be rewritten as

$$R_{i,t+1} = \frac{M_{i,t+1} + (1-\delta) Q_{i,t+1}}{Q_{i,t}}, \quad (6)$$

and the levered investment return $R^{Iw}_{i,t+1}$ depends on the investment return $R_{i,t+1}$, the after-tax corporate bond return $R^{Ba}_{i,t+1}$, and the market leverage $w_{i,t}$:

$$R^{Iw}_{i,t+1} = \frac{R_{i,t+1} - w_{i,t} R^{Ba}_{i,t+1}}{1 - w_{i,t}}. \tag{7}$$

2.2 A model of aggregation and aggregate investment return

Given the firm-level parameters: the capital share (α) and the adjustment cost parameter (a), using Equation (4) to (6), for each firm, we can construct a time series of the firm's investment returns and the associated two components: the firm's marginal profitability of capital (M) and the firm's marginal value of capital (Q). At the aggregate market level, we need to construct the aggregate investment return and a similar representation as Equation (6) but with two aggregate components, namely aggregate marginal profitability of capital and aggregate marginal value of capital. This is needed in order to derive the present-value relation connecting the aggregate marginal value of capital to future aggregate marginal profitability of capital and future aggregate investment returns. In this section, we present a model of aggregation of firms and present the model-implied weights. Our derivation follows Cooper *et al.* (2021).

Let N be the number of firms in the market. Each firm optimizes by equating the marginal adjustment costs of investment to the marginal value of capital. Each firm makes an investment $I_{i,t}$ at time t, and exiting time t with a level of capital stock $K_{i,t+1}$. Given the constant returns to scale assumption, applying the result of Hayashi (1982) implies that the marginal value of capital is equal to the average value of capital, and hence the firm's value at the end of time t is given by $K_{i,t+1}Q_{i,t}$ where $Q_{i,t}$ is the marginal value of capital at the end of time t.[8] The aggregate market value is therefore $\sum_{i=1}^{N} K_{i,t+1}Q_{i,t}$. We measure the aggregate marginal value of capital at the end of time t, denoted by Q_t, by assuming that Q_t can price the total capital stock value at the end of

[8]This notation is consistent with the notation in Belo, Xue, and Zhang (2013).

time t if multiplied by the total capital stock at the end of time t, as follows:

$$\left(\sum_{i=1}^{N} K_{i,t+1}\right) Q_t = \sum_{i=1}^{N} K_{i,t+1} Q_{i,t}. \tag{8}$$

Equivalently, Q_t can be expressed as

$$Q_t = \sum_{i=1}^{N} \left(\frac{K_{i,t+1}}{\sum\limits_{j=1}^{N} K_{j,t+1}} \right) Q_{i,t}. \tag{9}$$

Thus, Q_t is a weighted average of individual firms' Q values where the weight of firm i is proportional to firm i's capital stock at the end of time t. Notice that an extra unit of capital in the economy invested according to the existing capital allocation in the economy, that is, invested proportionally to the fraction of capital of each firm from the total capital in the economy will indeed have a value of Q_t.[9]

For the same N firms at time $t+1$, we can measure the aggregate marginal Q at time $t+1$, Q_{t+1}, by assuming that Q_{t+1} can price the total firm value at $t+1$:

$$Q_{t+1} = \sum_{i=1}^{N} \left(\frac{K_{i,t+2}}{\sum_{j=1}^{N} K_{j,t+2}} \right) Q_{i,t+1}.$$

We assume that an extra unit of capital for the aggregate economy at time t is invested according to the firms' proportion of their capital stock at the end of time t. Therefore, the aggregate marginal profit of that extra unit of capital (M_{t+1}) is a capital stock weighted average of firms' marginal profits of capital. That is,

$$M_{t+1} = \sum_{i=1}^{N} \left(\frac{K_{i,t+1}}{\sum\limits_{j=1}^{N} K_{j,t+1}} \right) M_{i,t+1}. \tag{10}$$

[9] The allocation of an extra unit of capital in the aggregate economy according to firms' proportions of capital stocks keeps unchanged the distribution of capital in the economy.

Finally, we define the aggregate investment return as the ratio of the aggregate marginal benefit of investment at time $t+1$ to the aggregate marginal cost of investment at time t:

$$R_{t+1} \equiv \frac{M_{t+1} + (1-\delta)\, Q_{t+1}}{Q_t}. \tag{11}$$

For an investor who holds the economy's stock of capital, an extra unit of capital at time t costs Q_t. This extra unit of capital generates profit M_{t+1} at time $t+1$ and depreciate to $1-\delta$ unit exiting time $t+1$ with a continuation value of $(1-\delta)\, Q_{t+1}$.

3 Estimating the Investment Return and its Components

In this section, we provide structural estimates of firm-level parameters and aggregate measures of the investment return and the respective components, which are based on the model presented in the last section.

3.1 Methodology

We follow Gonçalves, Xue, and Zhang (2020) and estimate the firm-level parameters, namely the capital share (α) and the adjustment cost parameter (a), using one-step GMM to fit the investment Euler equation moment for each testing portfolio jointly with an additional moment, namely the valuation moment as in Belo, Xue, and Zhang (2013). We include the valuation moment in the estimation because in the benchmark setting we consider the aggregate market portfolio as the testing portfolio. Hence, using only the investment Euler equation leads to an unidentified estimation with one moment but two parameters (the capital share and the adjustment cost parameter).[10] With two moments and two parameters, the estimation is exactly identified and the two moments fit perfectly.

Specifically, for a given set of testing portfolios (indexed by j), the first set of moment conditions

[10] Similarly, Belo, Xue, and Zhang (2013) base their tests on both the investment Euler equation and the valuation equation.

corresponds to testing whether the average stock return equals the average levered investment return for each testing portfolio j,

$$e_j^r \equiv \mathrm{E}_T[R^S_{j,t+1} - R^{Iw}_{j,t+1}(\alpha, a)] = 0, \tag{12}$$

where $\mathrm{E}_T(\cdot)$ denotes the sample moment, $R^S_{j,t+1}$ is the portfolio stock return, and $R^{Iw}_{j,t+1}$ is the portfolio levered investment return that depends on parameters α and a.

The second set of moment conditions tests whether the average Tobin's Q in the data equals the average Q predicted by the model,

$$e_j^q \equiv \mathrm{E}_T\left[\widetilde{Q}_{j,t} - \left(1 + (1-\tau_t)a\frac{I_{j,t}}{K_{j,t}}\right)\frac{K_{j,t+1}}{A_{j,t}}\right] = 0, \tag{13}$$

where $A_{j,t}$ is the book value of assets and $\widetilde{Q}_{j,t}$ is the Tobin's Q in the data defined as the ratio of the sum of market equity and total debt to the total assets, $\widetilde{Q}_{j,t} \equiv (P_{j,t} + B_{j,t+1})/A_{j,t}$, where $P_{j,t}$ is the market value of the firm's equity and $B_{j,t+1}$ is the firm's book value of debt.

Firm-level accounting variables and, thus, firm-level investment returns are subject to the issue of outliers. The outliers in firm-level investment returns can contaminate the aggregate portfolio-level investment returns and lead to noisy parameter estimates from the GMM estimation. To alleviate the impact of outliers, we follow Gonçalves, Xue, and Zhang (2020) and construct firm-level investment returns using winsorized firm-level accounting variables, then compute value-weighted portfolio levered investment returns to match with value-weighted portfolio stock returns.[11] As a robustness check (see Section 5), instead of winsorization, we also follow an alternative approach that is employed in Belo, Gala, Salomao, and Vitorino (2022), where portfolio median is used to aggregate firm-level investment returns to portfolio level since the median is known to be robust to outliers.

[11] We winsorize firm-level accounting variables at the 1-99% level.

3.2 Data

We largely follow Gonçalves, Xue, and Zhang (2020) and Belo, Xue, and Zhang (2013) in measuring accounting variables and in aligning their timing with the timing of stock returns. Our sample consists of all common stocks on NYSE, Amex, and Nasdaq from 1963 to 2018. The firm-level data are from the merged CRSP and COMPUSTAT industrial database. We include all firms with fiscal year ending in the second half of the calendar year. We exclude firms with primary standard industrial classifications between 4900 and 4999 (utilities) and between 6000 and 6999 (financials). We also delete firm-year observations for which total assets, capital stock, or sales are either zero or negative.

Capital stock ($K_{i,t}$) is net property, plant, and equipment (Compustat annual item PPENT). Investment ($I_{i,t}$) is capital expenditures (Compustat annual item CAPX) minus sales of property, plant, and equipment (Compustat annual item SPPE, zero if missing). Total debt ($B_{i,t+1}$) is long-term debt (Compustat annual item DLTT, zero if missing) plus short-term debt (Compustat annual item DLC, zero if missing). $A_{i,t}$ is total assets (Compustat annual item AT). Market equity ($P_{i,t}$) is the stock price per share (CRSP item prc) times the number of shares outstanding (CRSP item shrout). Market leverage ($w_{i,t}$) is the ratio of total debt to the sum of total debt and market equity. We follow Cochrane (1991) and assume a depreciation rate (δ) equal to 0.1. Output ($Y_{i,t}$) is sales (Compustat annual item SALE). Market leverage, $w_{i,t}$, is the ratio of total debt to the sum of total debt and the market value of equity. We measure the tax rate (τ_t) as the statutory corporate income tax (from the Commerce Clearing House, annual publications). The after-tax corporate bond returns ($R^{Ba}_{i,t+1}$) are computed from $R^{B}_{i,t+1}$ using the average of tax rates in year t and $t+1$. For the pre-tax corporate bond returns ($R^{B}_{i,t+1}$) we use the ratio of total interest and related expenses (Compustat annual item XINT) scaled by the total debt ($B_{i,t+1}$).[12]

At the end of June of year t, we construct the aggregate "market" portfolio. That is, a portfolio whose value is the value of the aggregate capital stocks and whose return is the aggregate investment

[12] As shown in Gonçalves, Xue, and Zhang (2020), doing so increases the sample coverage by 12.7% as compared to the prior studies that use credit rating imputation such as Liu, Whited, and Zhang (2009).

returns. Alternatively, we sort all stocks on Tobin's Q at the end of June of year t into deciles based on the NYSE breakpoints. For each testing portfolio, we compute annual value-weighted stock returns from July of year t to June of year $t+1$. We construct annual levered investment returns to match with annual stock returns and annual valuation ratios to match with annual Tobin's Q. To construct the matching annual levered investment returns, we use capital at the end of fiscal year $t-1$ ($K_{i,t}$), the tax rate, investment, and capital at the end of year t (τ_t, $I_{i,t}$, and $K_{i,t+1}$), as well as other variables at the end of year $t+1$ (τ_{t+1}, $Y_{i,t+1}$, and $I_{i,t+1}$). To match with $\widetilde{Q}_{i,t}$ for portfolios formed at the end of June of year t, we take $I_{i,t}$ from the fiscal year ending in calendar year t and $K_{i,t}$ from the fiscal year ending in year $t-1$.[13]

3.3 Structural parameter estimates

In the basecase GMM estimation, the testing portfolio is the aggregate stock market portfolio and the portfolio returns are measured as the value-weighted returns. The estimation is exactly identified and the two moments fit perfectly. The estimate of capital share (α) is 0.08, which is similar to the results in Gonçalves, Xue, and Zhang (2020).[14] The 8% estimate seems rather small relative to common values assigned to this parameter in the macroeconomic literature. However, this estimate refers to the firm-level, not the aggregate level. When we conduct the GMM estimation using the Liu, Whited, and Zhang (2009) methodology (i.e. estimating the parameters at the aggregate portfolio-level) our estimated share of capital rises to 0.23 which is closer to the value assigned to this parameter in the macroeconomic literature, and our main results turn out similar.[15] Relatedly, Basu and Fernald (1997) find that the estimates of returns to scale generally rise with the level of aggregation.

[13] Compustat records both stock and flow variables at the end of year t. In the model, however, stock variables dated t are measured at the beginning of year t, and flow variables dated t are over the course of year t. To capture this timing difference, we follow Liu, Whited, and Zhang (2009) and take, for example, for the year 2003 the beginning-of-year capital ($K_{i,2003}$) from the 2002 balance sheet and any flow variable over the year, such as $I_{i,2003}$, from the 2003 income or cash flow statement.

[14] Gonçalves, Xue, and Zhang (2020) report estimates of capital share varying from 5.04% to 7.53% across different testing portfolios in Table 5 Panel B.

[15] Barkai (2020) finds that the share of capital since 1985 has been mostly well below 0.25.

The estimate of adjustment cost parameter (a) is 15.18, which is higher than the corresponding estimates in Gonçalves, Xue, and Zhang (2020) based on testing portfolios different from our setup, but similar to the estimates in Liu, Whited, and Zhang (2009).[16] The reason for the difference in the adjustment cost estimate is that we are fitting both the investment return moment and the valuation moment, whereas Gonçalves *et al.* (2020) fit only the investment return moment. Untabulated results when using the same portfolios and sample as in their study and fitting both the return and valuation moments produce very similar capital share and adjustment cost estimates to those estimates that we report. The estimated magnitude of the adjustment costs is 11.09% of sales, which is in line with those reported in prior studies.[17]

Given the parameters α and a that are estimated using firm-level data, we compute the aggregate investment return and its components, by plugging these parameter estimates in Equations (9) to (11), together with firm-level accounting variables.

4 Variance Decomposition for q: Time-Series Analysis with Market Portfolio

In this section, we evaluate the forecasting performance of q for both future investment returns and marginal profits by deriving and estimating a variance decomposition for the log Q. The objective is to assess what are the sources of predictability that drive the variation in aggregate log Q over time.

[16] Gonçalves, Xue, and Zhang (2020) report estimates of adjustment cost parameter varying from 0.72 to 5.66 in Table 5 (Panel B) and from 1.63 to 8.11 in Table 3 (Panel B). Liu, Whited, and Zhang (2009) report estimates varying from 11.5 to 28.9 when matching two moments: expected returns and variances. Belo *et al.* (2022) (e.g., their Table 3) report adjustment costs in the range of 21 to 29 in the one-capital specification.

[17] For example, Cooper and Priestley (2016) find that implied adjustment costs represent 12.21% of sales across a host of manufacturing industries. Bloom (2009) surveys the estimates of convex adjustment costs to be between zero and 20% of revenue.

4.1 A present-value relation

We start by deriving a dynamic present-value relation for the log Q, which represents the basis for the empirical analysis conducted in the rest of the section.

Our methodology relies on the definition of the realized gross investment return (R) presented in Equation (11). This definition is analogous to the usual definition of the gross stock return with Q playing the role of the stock price and M being the analog of dividends. By conducting a log-linear transformation of the investment return in Equation (11), and proceeding along the lines of Campbell and Shiller (1988) and Campbell, Lo, and MacKinlay (1997), we derive the following approximate difference equation in log Q,

$$q_t \approx const. + \rho q_{t+1} - r_{t+1} + (1-\rho) m_{t+1}, \tag{14}$$

where $q_t \equiv \ln(Q_t)$ is the log Q at time t; $r_{t+1} \equiv \ln(R_{t+1})$ represents the log investment return at time $t+1$; and $m_{t+1} \equiv \ln(M_{t+1})$ denotes the log marginal profit at time $t+1$. In this setting, variables denoted with lower-case letters represent the logs of the corresponding variables in upper-case letters.

ρ plays an important role in the analysis, representing a (log-linearization) discount coefficient that depends on the mean of the log marginal profits-to-Q ratio ($mq_t \equiv m_t - q_t$),

$$\rho \equiv \frac{e^{\ln(1-\delta)-\overline{mq}}}{1+e^{\ln(1-\delta)-\overline{mq}}},$$

where $\overline{mq}$ represents the average of mq_t.

By iterating the equation above forward, we obtain the following present-value dynamic relation for q at each forecasting horizon H:

$$q_t \approx const. - \sum_{h=1}^{H} \rho^{h-1} r_{t+h} + \sum_{h=1}^{H} \rho^{h-1}(1-\rho) m_{t+h} + \rho^H q_{t+H}. \tag{15}$$

Under this present-value relation, current q (q_t) is positively correlated with both future multi-period log profits (m_{t+h}) and the future q at terminal date $t + H$ (q_{t+H}), while being negatively correlated with future multi-period log investment returns (r_{t+h}). This present-value relation is analogous to the present-value relation associated with the stock price derived in Campbell, Lo, and MacKinlay (1997), where q replaces the log stock price (p), m is analogous to the log dividend (d), and r replaces the log stock return.

At an infinite horizon, by assuming the following transversality (or no-bubbles) condition,

$$\lim_{H \to \infty} \rho^H q_{t+H} = 0,$$

we obtain the following long-run present-value relation:

$$q_t \approx const. - \sum_{h=1}^{\infty} \rho^{h-1} r_{t+h} + \sum_{h=1}^{\infty} \rho^{h-1}(1-\rho) m_{t+h}. \tag{16}$$

Hence, at very long horizons, only predictability of future investment returns and predictability of marginal profits drives the variation in the current q.[18] Which of these two components matters most in terms of driving the dynamics of q remains an empirical question, which will be addressed in the following sections.

Table 1 (Panel A) presents the descriptive statistics for the variables in the present-value relation for q. Both the investment return and q have a volatility of 10%. On the other hand, both q (with a first-order autocorrelation of 0.61) and m (0.54) are considerably more persistent than r (0.12). Figure 1 plots the time series of r, m, and q. All three variables appear to be mean-reverting to a large degree, and hence, stationary. Both r and m seem to be procyclical variables, as they tend to decline around most recession periods and rise during economic booms. On the other hand, q appears to be less correlated with the business cycle.

Panel B of Table 1 contains the correlations among the three variables. The correlation between

[18] Lettau and Ludvigson (2002) derive a related dynamic accounting decomposition for the log Q. However, their present-value relation is based on a second-order Taylor expansion.

m and q is very high, at 0.95. Thus, times of high marginal profitability of capital correspond to times when both the marginal value of capital and investment are high. We can also see that the investment return is positively correlated with both m and q, albeit to a lower degree, as indicated by the correlations around 0.60. The reason is that shocks to both marginal profits and Q are also shocks to contemporaneous returns.

Regarding the critical parameter ρ, which is a function of the average mq ratio, we obtain an estimate of 0.85 for our sample period (1964–2018). This estimate is somewhat smaller than corresponding estimates for the analogue parameter values in present-value relations associated with stock returns, which are typically above 0.90 (see Campbell and Vuolteenaho, 2004; Cochrane, 2008, 2011; Maio, 2013; Maio and Santa-Clara, 2015, among others).[19]

4.2 Restricted VAR

Following Cochrane (2008), we specify the following first-order restricted VAR,

$$r_{t+1} = \pi_r + \lambda_r q_t + \varepsilon^r_{t+1}, \tag{17}$$

$$m_{t+1} = \pi_m + \lambda_m q_t + \varepsilon^m_{t+1}, \tag{18}$$

$$q_{t+1} = \pi_q + \phi q_t + \varepsilon^q_{t+1}, \tag{19}$$

where the εs represent forecasting errors. This VAR system is estimated by multiple-equation OLS (see Hayashi, 2000), with Newey and West (1987) t-statistics (computed with one lag).[20]

By combining the VAR above with the present-value relation in Equation (15), we obtain an approximate identity involving the predictability coefficients associated with q_t, at each forecasting

[19] In those studies, the value of ρ is either calibrated or estimated as a function of the average log dividend-to-price ratio.

[20] Using zero lags (White, 1980) leads to similar statistical inference.

horizon H:

$$
\begin{aligned}
1 &\approx b_m^H - b_r^H + b_q^H, \\
b_r^H &\equiv \lambda_r \frac{1-\rho^H \phi^H}{1-\rho\phi}, \\
b_m^H &\equiv (1-\rho)\lambda_m \frac{1-\rho^H \phi^H}{1-\rho\phi}, \\
b_q^H &\equiv \rho^H \phi^H.
\end{aligned}
\tag{20}
$$

This equation can be interpreted as a variance decomposition for q. The predictive coefficients, $-b_r^H$, b_m^H, and b_q^H, represent the fraction of the variance of current q attributed to the predictability of future investment returns, marginal profits, and q, respectively. Hence, these slopes measure the weight (of the predictability) of each of these variables ($\sum_{h=1}^{H} \rho^{h-1} r_{t+h}$, $\sum_{h=1}^{H} \rho^{h-1}(1-\rho) m_{t+h}$, and $\rho^H q_{t+H}$) in terms of driving the variation in the current q. This relation also imposes a quantitative constraint on the predictability associated with q in the sense that the slopes need to add (approximately) to one. Hence, if at some forecasting horizon H, q_t forecasts neither future investment returns nor future marginal profits, then it must forecast its own future value at time $t+H$. Otherwise q would not vary over time, something that is counterfactual, as discussed above.

In this variance decomposition, the predictive slopes at each forecasting horizon H are obtained from the one-period VAR slopes. Cochrane (2008, 2011) specifies a similar variance decomposition for the dividend yield. The expressions above imply that the relative shares of predictability (e.g., b_r^H / b_m^H) are invariant with the forecasting horizon. Further, the multi-horizon slopes represent mechanical transformations of the one-year VAR slopes, which means that the short-run VAR dynamics dictate all the implied long-run dynamics. The first-order VAR addresses the concern of the lack of statistical power at long horizons associated with long-horizon regressions.

We can also compute the variance decomposition for an infinite horizon ($H \to \infty$):

$$\begin{aligned} 1 &\approx b_m^{lr} - b_r^{lr}, \\ b_r^{lr} &\equiv \frac{\lambda_r}{1-\rho\phi}, \\ b_m^{lr} &\equiv \lambda_m \frac{1-\rho}{1-\rho\phi}. \end{aligned} \tag{21}$$

In the long-run decomposition, all the variation in current q is associated with either return or marginal profits predictability. The VAR approach enables one to estimate this long-run decomposition, something that is not feasible under the direct method. The t-statistics associated with both the multi-horizon and long-run predictive coefficients are computed from the t-statistics corresponding to the VAR slopes by using the delta method. The full details on the derivation of the variance decomposition are available in the online appendix.

Following Cochrane (2008), we compute t-statistics for two joint null hypotheses of long-run predictability: the first null assumes that there is only marginal profits predictability,

$$H_0 : b_r^{lr} = 0, b_m^{lr} = 1,$$

while the second null hypothesis assumes that there is only return predictability:

$$H_0 : b_r^{lr} = -1, b_m^{lr} = 0.$$

The results for the baseline variance decomposition are shown in Table 2 (Panel A) and Figure 2 (Panels A and C). The return channel plays a dominant role at nearly all forecasting horizons. Apart from $H = 1$, predictability of future investment returns is the major source of variation in q. Indeed, the return slope estimates are strongly significant (1% level) at all horizons beyond one year, with magnitudes around or above 0.60 at most horizons. Marginal profits predictability assumes a secondary role, with weights close to 0.40 at intermediate and long horizons. Despite the smaller magnitudes, it turns out that the m coefficient estimates are also strongly significant (1% or 5% level) at all horizons. This arises from the fact that the corresponding one-year VAR slope estimate is largely significant (t-ratio of 4.58). In fact, the VAR equation for m has a larger fit than the corresponding return equation, as indicated by the explanatory ratio of 0.28 (versus 0.08 for r). However, there is a downward effect caused by the term $1-\rho$, which is substantially smaller than one (0.15). Therefore, the larger magnitude of the m VAR coefficient (1.24 versus -0.29) is not enough to generate larger multi-horizon predictive slopes than the coefficients associated with future investment returns (in magnitude). The shares associated with predictability of future (single-date) q are economically and statistically significant at very short horizons. However, such estimates decay to zero, and become insignificant, at a relatively rapid pace ($H > 3$).

We observe that the sum of the variance decomposition is very close to one (above 0.98) at all forecasting horizons. This shows that the present-value relation for q is quite accurate. In other words, ignoring the higher-order terms, which are absent from the first-order Taylor approximation underlying the present-value relation for q, does not have a significant impact on the variance decomposition.

At very long horizons, the results in Table 2 (Panel A) indicate that the return and marginal profits coefficients are -0.61 and 0.37, respectively. We clearly reject the null that $b_r^{lr} = 0, b_m^{lr} = 1$ (t-ratios around 4 in magnitude). Yet, we also reject the null that $b_r^{lr} = -1, b_m^{lr} = 0$, with t-ratios around 2.50. Hence, in statistical terms, both the return and marginal profits channels matter in terms of driving q. However, the size of the return channel is nearly twice as large as the size of the marginal profits channel. These results for the long-run predictability mix associated with q are not particularly surprising. In fact, since $1-\rho$ is a number substantially below one (around 0.15), in order to obtain (at least) similar long-run shares for m and r, we would need the magnitude of the marginal profits VAR slope estimate to be almost seven times larger than that of the VAR return slope estimate. The VAR estimation results confirm a higher magnitude of λ_m relative to λ_r (1.24 versus 0.29), but such discrepancy is clearly not enough to generate similar long-run weights. In order words, despite the larger short-run predictability of q for m than for r, such difference does not compensate for the pre-specified tilt (embedded in the decomposition) towards a larger weight for return predictability.

4.3 Simulation

Next, we conduct a bootstrap simulation of the restricted VAR model estimated above. The objective is to account for the relatively poor small-sample properties of long-horizon predictability and the question of whether the asymptotic inference is valid when assessing the statistical significance of the implied multi-horizon slopes (see Valkanov, 2003; Torous, Valkanov, and Yan, 2004; Boudoukh, Richardson, and Whitelaw, 2008, among others for a discussion on this issue). In related work, Cochrane (2008) and Maio and Santa-Clara (2015) conduct VAR-based Monte-Carlo simulations to assess the predictive ability of the dividend yield for future stock returns and dividend growth. One key advantage of a bootstrap simulation relative to a monte-carlo simulation is that we can skip the normally-distributed assumption for the variables in the system.

To assess predictability of future returns, we impose a null hypothesis where q does not forecast the future investment return. Under this null, all the variation in q comes from predicting future marginal profits. Thus, we simulate the first-order VAR by imposing the restrictions, both in the predictive slopes and residuals, associated with this null hypothesis,

$$\begin{pmatrix} r_{t+1} \\ m_{t+1} \\ q_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{1-\rho\phi}{1-\rho} \\ \phi \end{pmatrix} q_t + \begin{pmatrix} \rho\varepsilon^q_{t+1} + (1-\rho)\varepsilon^m_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^q_{t+1} \end{pmatrix}, \tag{22}$$

where all the variables in the VAR are demeaned.

To assess predictability of future marginal profits, we simulate an alternative VAR specification under the null hypothesis that q does not forecast future m. This means that all the variation in q

emanates from predicting future investment returns:

$$\begin{pmatrix} r_{t+1} \\ m_{t+1} \\ q_{t+1} \end{pmatrix} = \begin{pmatrix} \rho\phi - 1 \\ 0 \\ \phi \end{pmatrix} q_t + \begin{pmatrix} \rho\varepsilon^q_{t+1} + (1-\rho)\varepsilon^m_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^q_{t+1} \end{pmatrix}. \tag{23}$$

We conduct a bootstrap experiment associated with each of the VARs specified above. We draw the VAR residuals (10,000 times) with replacement from the original VAR estimates. The realization of q for the base period is chosen randomly from the original time-series of q_t. We compute the pseudo p-values associated with the implied VAR return slopes at each horizon, which represent the fractions of simulated estimates of the return coefficients (from the simulations associated with the first VAR above) that are lower than the corresponding estimates found in the data. Similarly, the pseudo p-values associated with the marginal profits coefficients represent the fractions of pseudo estimates of the profitability slopes (obtained from the simulations under the second VAR presented above) that are higher than the corresponding sample estimates. The full details of the bootstrap simulation are available in the online appendix.

We note that our bootstrap simulation does not account for the fact that the variables in the predictive system are nested variables, that is, they are estimated with error (rather than observed). In principle, we could simulate the structural model (described in Section 2) inside the bootstrap experiment in order to account for the estimation error in those variables. However, such procedure is problematic in our case, since we have to employ non-linear GMM estimation of the structural model, as described in Section 3. Specifically, it is likely that for many of the pseudo samples, the numerical optimization underlying the GMM estimation does not converge properly, which would lead to a problematic or infeasible bootstrap simulation. Perhaps, more important, the bootstrap simulation conducted in this subsection produces p-values for the return slopes that are very small. Hence, it is unlikely that incorporating such additional source of statistical uncertainty would turn the most relevant predictive coefficients (in the variance decomposition) insignificant.

The results associated with the bootstrap simulation are presented in Figure 3. It turns out that the p-values associated with the return slopes are below 1% at all forecasting horizons beyond the two-year horizon. In the case of the marginal profit coefficient estimates, the corresponding p-values are below 1% at all horizons beyond the four-year horizon. At very short horizons, both r and m coefficient estimates are significant at the 10% or 5% level. Therefore, these results are consistent with the asymptotic inference presented in the previous subsection, that is, both the return and marginal profits slope estimates are strongly statistically significant at nearly all forecasting horizons.

4.4 Unrestricted VAR

In this subsection, we estimate an alternative variance decomposition for q, based on a less restrictive first-order VAR.

Specifically, we consider an unrestricted VAR(1):

$$r_{t+1} = \pi_r + \gamma_r r_t + \theta_r m_t + \lambda_r q_t + \varepsilon^r_{t+1}, \tag{24}$$

$$m_{t+1} = \pi_m + \gamma_m r_t + \theta_m m_t + \lambda_m q_t + \varepsilon^m_{t+1}, \tag{25}$$

$$q_{t+1} = \pi_q + \gamma_q r_t + \theta_q m_t + \phi q_t + \varepsilon^q_{t+1}. \tag{26}$$

This specification accounts for relevant predictability of lagged returns and marginal profits on all three variables in the system, something that the benchmark VAR misses. Indeed, Maio and Xu (2020) show that the restricted VAR(1), in a similar context, can be severely misspecified. This originates an implausible long-run variance decomposition for aggregate stock price ratios, such as the earnings yield or dividend yield.

The VAR above can be presented in matrix form,

$$\begin{pmatrix} r_{t+1} \\ m_{t+1} \\ q_{t+1} \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_m \\ \pi_q \end{pmatrix} + \begin{pmatrix} \gamma_r & \theta_r & \lambda_r \\ \gamma_m & \theta_m & \lambda_m \\ \gamma_q & \theta_q & \phi \end{pmatrix} \begin{pmatrix} r_t \\ m_t \\ q_t \end{pmatrix} + \begin{pmatrix} \varepsilon^r_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^q_{t+1} \end{pmatrix}. \tag{27}$$

Equivalently, the VAR can be defined as

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{28}$$

where the last equation defines the variables of interest.

The benchmark restricted VAR(1) is nested in this general specification, with

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & \lambda_r \\ 0 & 0 & \lambda_m \\ 0 & 0 & \phi \end{pmatrix}.$$

Consider the indicator vectors, $\mathbf{e}_r \equiv (1, 0, 0)'$, $\mathbf{e}_m \equiv (0, 1, 0)'$, and $\mathbf{e}_q \equiv (0, 0, 1)'$, which represent the position of each state variable in the VAR. As in the benchmark case, the VAR is estimated by applying multiple-equation OLS, with Newey-West t-ratios. The t-ratios of the implied horizon-specific coefficients are produced by employing the delta method. The covariance matrix of the state variables is given by $\mathbf{\Sigma} \equiv \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}_t')$. Given these definitions, and following Larrain and Yogo (2008) and Maio and Xu (2020), we derive the following variance decomposition for q at each horizon H,

$$\begin{aligned} 1 &\approx b_m^H - b_r^H + b_q^H, \qquad (29)\\ b_r^H &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I}-\rho^H\mathbf{A}^H)(\mathbf{I}-\rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\mathbf{\Sigma}\mathbf{e}_q},\\ b_m^H &\equiv \frac{(1-\rho)\mathbf{e}_m'\mathbf{A}(\mathbf{I}-\rho^H\mathbf{A}^H)(\mathbf{I}-\rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\mathbf{\Sigma}\mathbf{e}_q},\\ b_q^H &\equiv \frac{\rho^H\mathbf{e}_q'\mathbf{A}^H\mathbf{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\mathbf{\Sigma}\mathbf{e}_q}, \end{aligned}$$

where $\mathbf{I}$ represents a conformable identity matrix.

Further details on the derivation of this variance decomposition are available in the online appendix. The expressions above show that the relative shares of predictability (e.g., b_r^H/b_m^H) change with the forecasting horizon, in contrast to the restricted VAR case. In other words, the unrestricted VAR enables for a decoupling between the short-run and implied long-run forecasting dynamics.

At an infinite horizon, it turns out that $\lim_{H\to\infty}\rho^H\mathbf{A}^H$ approaches to a matrix of zeros. Thus, the corresponding long-run VAR-based variance decomposition for q is given by

$$\begin{aligned} 1 &\approx b_m^{lr} - b_r^{lr}, \qquad (30)\\ b_r^{lr} &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I}-\rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\mathbf{\Sigma}\mathbf{e}_q},\\ b_m^{lr} &\equiv \frac{(1-\rho)\mathbf{e}_m'\mathbf{A}(\mathbf{I}-\rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\mathbf{\Sigma}\mathbf{e}_q}. \end{aligned}$$

As in the restricted VAR case, the t-ratios for the implied infinite-horizon slope estimates are obtained by using the delta method.

The estimation of the unrestricted VAR above yields the following results,

$$\begin{pmatrix} r_{t+1} \\ m_{t+1} \\ q_{t+1} \end{pmatrix} = \widehat{\boldsymbol{\pi}} + \begin{pmatrix} 0.41(\underline{2.43}) & -0.06(-0.29) & -0.39(-0.91) \\ 0.59(\mathit{1.65}) & 0.13(0.31) & 0.63(0.65) \\ 0.39(\mathbf{2.98}) & -0.08(-0.56) & 0.59(\mathit{1.74}) \end{pmatrix} \begin{pmatrix} r_t \\ m_t \\ q_t \end{pmatrix} + \begin{pmatrix} \widehat{\varepsilon}^r_{t+1} \\ \widehat{\varepsilon}^m_{t+1} \\ \widehat{\varepsilon}^q_{t+1} \end{pmatrix},$$

with R^2 estimates of 0.18, 0.33, and 0.45, respectively. The numbers in parentheses represent the t-ratios, with bold, underlined, and italic numbers denoting significance at the 1%, 5%, and 10%, respectively.

These results show that r_t helps to forecast a rise in all three variables in the system, as the respective coefficient estimates are positive and statistically significant in all cases (marginally so in the equation for marginal profits). On the other hand, the slope estimates associated with m_t are largely insignificant in all cases. Moreover, the estimate of λ_m is cut to about half the magnitude of the corresponding estimate in the restricted VAR (0.63 versus 1.24) and becomes largely insignificant. In comparison, the estimate of λ_r increases in magnitude (-0.39 versus -0.29), but also with no statistical significance. The estimate of ϕ is similar to that obtained in the baseline VAR (0.59 versus 0.61), albeit the statistical significance becomes relatively weak (10% level). These results suggest that the short-run dynamics associated with the restricted and unrestricted VARs can differ by a good deal. In particular, there is a kind of a "substitution effect" in predictive power among some of the variables: Restricting to zero the slopes associated with lagged r, magnifies the forecasting role of q for future marginal profits.

The horizon-specific variance decompositions based on the unrestricted VAR(1) are presented in Figure 2 (Panels B and D). The results point to an even more dominant role of return predictability in comparison to the benchmark VAR. Specifically, the long-run (infinite horizon) return and marginal profit slope estimates are -0.80 and 0.18, respectively. Further, while the return coefficient estimates are strongly significant at all forecasting horizons, there is significance for the m slopes estimates only at very short horizons ($H < 3$). Untabulated results indicate that we cannot reject the null (at the 10% level, t-ratio of 1.20) that all the variation in q stems from long-run return predictability. As in the restricted VAR, predictability of future q is the major driving force at $H = 1$, but this effect dies off at a faster pace.

All in all, the punch line of this section is that predictability of future investment returns (the discount-rate channel) is the major driving force of variation in q. Predictability of future marginal profits (the cash-flow channel) plays a secondary role at most. This pattern is even more evident under the more robust unrestricted VAR.

5 Sensitivity Analysis

In this section, we provide a sensitivity analysis to the empirical results discussed in the previous section. To save space and keep the focus, most results are based on the restricted VAR method.

5.1 Alternative investment series

We conduct the variance decomposition for q by using alternative time series of the investment variables.

First, the data is generated from GMM estimation of the structural investment model based on ten Tobin's Q-sorted portfolios, as in Belo, Xue, and Zhang (2013). The resulting GMM estimates are similar to the base case. The estimate of capital share (α) is 0.07. The estimate of the adjustment cost parameter (a) is 20.31 and the corresponding ratio of adjustment-cost-to-sales is equal to 14.84%. The model is not rejected by the χ^2-test, with a p-value around 0.44. The results tabulated in Table 2 (Panel B) show that the predictability mix is very similar to that estimated with the benchmark data. Specifically, the long-run return and profits slopes are -0.62 and 0.36, respectively. The statistical significance is also very close to that obtained in the benchmark case.

Second, the investment data are associated with the median firm, rather than the value-weighted average. This is in line with Belo, Gala, Salomao, and Vitorino (2022) who use the portfolio median to aggregate firm-level investment returns to portfolio level since the median is robust to outliers. In particular, for a given portfolio of firms, they compute the portfolio median of firm-level investment returns to match with the portfolio median of stock returns. Using artificial data simulated from known firm-level parameters, they show that matching the portfolio median in the GMM estimation can recover the true firm-level parameters without bias. The estimate of α is 0.08, while the estimate of a is 20.89 (with the corresponding ratio of adjustment-cost-to-sales being equal to 15.27%). As shown in Table 2 (Panel C), the long-run predictability mix is identical to that estimated under the benchmark case.

5.2 Higher-order VAR

Next, we estimate a variance decomposition for q, based on a restricted second-order VAR. The rationale is that the second lag of q might provide useful information for predicting the three variables in the system.

The restricted VAR(2) specification is given by

$$r_{t+1} = \pi_r + \lambda_{r1} q_t + \lambda_{r2} q_{t-1} + \varepsilon^r_{t+1}, \tag{31}$$

$$m_{t+1} = \pi_m + \lambda_{m1} q_t + \lambda_{m2} q_{t-1} + \varepsilon^m_{t+1}, \tag{32}$$

$$q_{t+1} = \pi_q + \phi_1 q_t + \phi_2 q_{t-1} + \varepsilon^q_{t+1}. \tag{33}$$

The VAR(2) is estimated by multiple-equation OLS, with Newey–West t-statistics (computed with two lags). We can write the VAR above as a VAR(1) in the companion form:

$$\begin{pmatrix} r_{t+1} \\ m_{t+1} \\ q_{t+1} \\ r_t \\ m_t \\ q_t \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_m \\ \pi_q \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \lambda_{r1} & 0 & 0 & \lambda_{r2} \\ 0 & 0 & \lambda_{m1} & 0 & 0 & \lambda_{m2} \\ 0 & 0 & \phi_1 & 0 & 0 & \phi_2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r_t \\ m_t \\ q_t \\ r_{t-1} \\ m_{t-1} \\ q_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon^r_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^q_{t+1} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{34}$$

or equivalently,

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}. \tag{35}$$

The indicator vectors are defined as follows,

$$\mathbf{e}_r \equiv (1,0,0,0,0,0)', \tag{36}$$

$$\mathbf{e}_m \equiv (0,1,0,0,0,0)', \tag{37}$$

$$\mathbf{e}_q \equiv (0,0,1,0,0,0)', \tag{38}$$

while the covariance matrix of $\mathbf{z}_t$ corresponds to

$$\boldsymbol{\Sigma} = \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}_t') = \begin{pmatrix} \mathrm{var}(r_t) & \mathrm{Cov}(r_t,m_t) & \mathrm{Cov}(r_t,q_t) & 0 & 0 & 0 \\ \mathrm{Cov}(r_t,m_t) & \mathrm{var}(m_t) & \mathrm{Cov}(m_t,q_t) & 0 & 0 & 0 \\ \mathrm{Cov}(r_t,q_t) & \mathrm{Cov}(m_t,q_t) & \mathrm{var}(q_t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{39}$$

In order to obtain the predictive slopes at each forecasting horizon H, we use the definitions above into the same formulas presented in the previous section for the case of the unrestricted VAR(1). Similar to the unrestricted VAR case, the t-ratios for the long-horizon coefficients are obtained by applying the delta method.

The estimation of the VAR(2) above yields the following results,

$$\begin{pmatrix} r_{t+1} \\ m_{t+1} \\ q_{t+1} \end{pmatrix} = \widehat{\boldsymbol{\pi}} + \begin{pmatrix} -0.02(-0.16) & -0.42(-\mathbf{2.75}) \\ 1.67(\mathbf{6.36}) & -0.67(\underline{-2.01}) \\ 0.86(\mathbf{7.82}) & -0.39(-\mathbf{3.31}) \end{pmatrix} \begin{pmatrix} q_t \\ q_{t-1} \end{pmatrix} + \begin{pmatrix} \widehat{\varepsilon}^{r}_{t+1} \\ \widehat{\varepsilon}^{m}_{t+1} \\ \widehat{\varepsilon}^{q}_{t+1} \end{pmatrix},$$

with R^2 estimates of 0.18, 0.33, and 0.45, respectively. The numbers in parentheses represent the

t-ratios, with bold, underlined, and italic numbers denoting significance at the 1%, 5%, and 10%, respectively.

These results show that q_{t-1} helps to forecast (a decline in) all three variables in the system, as the respective slope estimates are negative and statistically significant in all cases. On the other hand, the estimate of λ_{r1} becomes close to zero and largely insignificant. We also observe that the estimates for both λ_{m1} and ϕ_1 register larger magnitudes than the corresponding estimates under the restricted VAR(1). Notably, there is a large increase in the fit of the forecasting regressions associated with r (from 0.08 to 0.18) and q (from 0.35 to 0.45), in comparison to the baseline VAR.

Untabulated results show that the long-run predictive slopes associated with future r and m are -0.69 and 0.29, respectively, and both of these estimates are statistically significant (at the 5% or 1% level). This indicates a slightly larger share of return predictability (69% versus 61%) in comparison to the restricted VAR(1) case.

5.3 Alternative δ

We employ an alternative value for the depreciation rate ($\delta = 0.1219$). This value represents the capital-weighted average firm-level depreciation rate in the sample, where the depreciation rate has the difference between Compustat item DP and Compustat item AM in the numerator, and Compustat item PPENT in the denominator.

The VAR estimation results are displayed in Table 2 (Panel D). We can see that the results are very similar to those in the benchmark case. Specifically, the long-run r and m coefficient estimates are -0.60 and 0.38, respectively. Regarding the statistical significance of these estimates, we get the same qualitative inference as in the benchmark setting.

5.4 Alternative structural estimation

We conduct the variance decomposition for q by using other time series of the investment variables. In contrast with the rest of the paper, we rely on the structural estimation method employed in Liu, Whited, and Zhang (2009). Using the aggregate market portfolio as the testing portfolio and matching the value-weighted portfolio returns, the estimated capital share (α) is 0.23, which is higher than the basecase estimate. On other hand, the estimated adjustment costs are similar to the basecase, as indicated by both the estimated parameter (a) of 16.10 and the estimated adjustment-costs-to-sales ratio (Φ/Y) of 11.76%.

The VAR estimation results are displayed in Table 2 (Panel E). The predictability mix is qualitatively similar to that estimated under the benchmark case, with a somewhat higher degree of return predictability. Specifically, the long-run return and marginal profits coefficient estimates are -0.71 and 0.30, respectively. This means that return predictability accounts for more than twice the share of marginal profits profitability in terms of explaining the variation in q. Both of these estimates are strongly significant (1% level), which means that both channels are important from a statistical viewpoint. In comparison to the benchmark case, the larger role for the discount-rate channel stems from both a larger magnitude of the VAR return slope estimate (-0.35 versus -0.29) and a lower magnitude of the m coefficient estimate (0.82 versus 1.24).

5.5 Direct approach

We estimate the variance decomposition for q by using the direct approach.

Following Cochrane (2008, 2011) and Maio and Santa-Clara (2015), we estimate weighted long-horizon regressions of future multiperiod log investment returns, future multiperiod log marginal profits, and future q on the current q:

$$\sum_{h=1}^{H} \rho^{h-1} r_{t+h} = a_r^H + b_r^H q_t + \varepsilon_{t+H}^r, \tag{40}$$

$$\sum_{h=1}^{H} \rho^{h-1} (1-\rho) m_{t+h} = a_m^H + b_m^H q_t + \varepsilon_{t+H}^m, \tag{41}$$

$$\rho^H q_{t+H} = a_q^H + b_q^H q_t + \varepsilon_{t+H}^q. \tag{42}$$

The estimation is conducted by equation-by-equation OLS and the t-statistics for the direct predictive slopes are based on Newey and West (1987) standard errors with $H-1$ lags (i.e., the Bartlett Kernel with a bandwidth of H). These standard errors incorporate a correction of the bias induced by using overlapping observations in the regressions presented above.

Similarly to Cochrane (2011), by combining the present-value relation for q in Equation (15) with the predictive regressions presented above, we obtain an approximate identity involving the predictability coefficients associated with q_t, at each forecasting horizon H:

$$1 \approx -b_r^H + b_m^H + b_q^H. \tag{43}$$

If the first-order VAR does not fully capture the dynamics of the data generating process for r, q, and m, it follows that the corresponding variance decomposition will be a poor approximation of the true decomposition for q, as discussed in Cochrane (2008) and Maio and Xu (2020). This problem does not exist under the direct approach, which a priori should yield the most correct estimates for the variance decomposition (see Cochrane, 2008, 2011; Maio and Santa-Clara, 2015). The minus side of the direct approach is that with small or moderate samples, the statistical power of the long-horizon regressions is negatively affected at very long horizons, given the substantial decline in the number of usable observations. For example at $H = 20$, 20 observations are lost by running the corresponding long-horizon regression.

The term structure of direct variance decompositions are presented in Figure 4. At the one-year horizon, the dominant source of variation in current q is its own predictability, with a share of 52%. This result emanates from the existence of some short-run persistence in this variable, as indicated in Table 1. Yet, such effect dies off quickly. Indeed, for forecasting horizons beyond one year, the key driving force of variation in q becomes return predictability, with shares above 65% at most horizons. At very long horizons, there is a slightly lower share of return predictability, with weights below 56%. The negative return slope estimates are strongly statistically significant (at the 1% or 5% level) at all horizons. In comparison, the positive coefficient estimates associated with future marginal profits are significantly smaller in magnitude, with shares around or below 30% at most forecasting horizons. At very long horizons ($H > 18$), we obtain a slightly larger role for m predictability, with weights around 40%. These slope estimates are statistically significant (at the 5% level) at most forecasting horizons, with the few exceptions occurring at some short and intermediate horizons.

Overall, the results from the direct approach are largely consistent with those based on the indirect or VAR approach used in our main empirical analysis.

5.6 Comparative statics

In this subsection, we conduct a comparative statics exercise. Specifically, we estimate a range of long-run variance decompositions associated with q for a set of artificial series of the key investment variables in the system (r, m, and q). The artificial time-series are obtained from calibration of the two key structural parameters of the theoretical model presented in Section 2 (α and a).[21] The goal of this analysis is to assess if the predictability mix associated with q, that we obtained above, holds for a reasonable range of those two underlying parameters.

The simulation results when the variance decomposition is based on the restricted VAR are presented in Table 3. Table 4 displays the simulation results based on the unrestricted VAR. We calibrate five different values for α (0.05, 0.15, 0.30, 0.50, and 0.70) and five values for a (1.37, 6.85, 13.69, 20.53, and 27.38). In the case of a, these values are associated with a calibration of the adjustment cost-to-output ratio (Φ/Y) of 0.01, 0.05, 0.10, 0.15, and 0.20, respectively. Hence, we have a total of 25 (5×5) different artificial data sets, which are used in the computation of the variance decomposition.[22] To save space and keep the focus, we report only the long-run (infinite horizon) variance decompositions for q.

The first key pattern that emerges from Tables 3-4 is that the share of return (marginal profits)

[21] Importantly, we use the same series of investment and sales as in the data. However, different parameter values will yield different series of r, m, and q.

[22] These ratios are consistent with Bloom (2009). Bloom (2009, Table IV) surveys the estimates of convex adjustment costs to be between zero and 20% of revenues.

predictability declines (increases) with α. Such pattern is especially predominant in the case of the unrestricted VAR: at high values of this parameter, that is $\alpha = 0.50, 0.70$, it turns out that the return slopes (that have the correct sign) are below 60% (in magnitude) in most cases, despite the large statistical significance. At the other end of the spectrum, for $\alpha = 0.05$, the estimates of b_r^{lr} are around or above 0.80 (in magnitude) in most cases. In the case of the restricted VAR, the difference in return slope estimates is not as large. On the other hand, we observe exactly the opposite pattern for the m slopes, and this holds for both VAR specifications. For example, we obtain estimates of b_m^{lr} around 0.60 for $\alpha = 0.05, a = 1.37$, while the corresponding estimates are around 2 for $\alpha = 0.70, a = 1.37$, and this holds under both VAR frameworks. We also observe across the board that the weights of return predictability tend to be larger under the unrestricted VAR than under the restricted VAR method. This pattern is especially evident among low values of α and is consistent with the evidence from Section 4.

Higher values of α entail higher volatilities of both marginal profits and investment returns. However, the volatility of marginal profits rises substantially more. For example, for a value of a of 1.37, a rise in the share of capital in production from 0.05 to 0.7, implies that the volatility of investment returns is six times higher, whereas the volatility of marginal profits is about 12 times higher. Consequently, the rise in the covariance of marginal profits and lagged q is substantially higher than the rise in the covariance of investment returns and lagged q.

The second key result from Tables 3-4 is that, for a given value of α, the share of return (marginal profits) predictability tends to increase (decrease) monotonically with a. Indeed, apart from the extreme case of $\alpha = 0.05$, the long-run return slope estimates have the wrong sign (positive) at very low values of a (1.37), albeit most of these estimates not being significant at the 10% level. Consequently, at those pairs of calibrated structural parameters, we obtain shares for long-run marginal profit predictability above 100%, which are strongly significant. This means that, in economic terms, marginal profits predictability explains all the variation in q, for extreme low values of a. However, for $a \geq 13.69$, it follows that the discount rate channel is dominant for most choices of α. Intuitively, in the extreme, if the value of the adjustment cost is zero, marginal Q is always one, implying no investment return predictability.

All in all, the results of this subsection show that the dominant role of return predictability in terms of driving variation in q is robust to a plausible range of the key parameters in the structural investment model. However, these simulation results also show that it is possible to find a relevant (and even dominant) share of marginal profits predictability under less plausible values for those structural parameters.

6 Variance Decomposition for q: Time-Series Analysis with Equity Portfolios

In this section, we extend the analysis in Section 4 for equity portfolios sorted on different equity characteristics. We are interested in studying whether the results for the sources of fluctuations in aggregate q extend to other portfolios and explore possible heterogeneity at the more disaggregated level. We study portfolios sorted on book-to-market (Rosenberg, Reid, and Lanstein, 1985; Fama and French, 1992), asset growth (Titman, Wei, and Xie, 2004; Cooper, Gulen, and Schill, 2008; Fama and French, 2008), and operating profitability (Haugen and Baker, 1996; Fama and French, 2008; Novy-Marx, 2013). These represent some of the main characteristics driving the cross-section of equity returns.

6.1 Data and variables

We conduct the GMM estimation of the structural investment model based on decile portfolios sorted on book-to-market (BM10), asset growth (IA10), and operating profitability (OP10). More concretely, we use the same sample as described in Section 3 to construct testing deciles based on the NYSE breakpoints at the end of June of each year and rebalance at the end of June of next year.

To form the BM10 deciles, at the end of June of each year t, we sort stocks on the book-to-market ratio, which is the book equity for the fiscal year ending in calendar year $t-1$ divided by the market equity (from CRSP) at the end of December of $t-1$.[23] To form the IA10 deciles, at the end of June of each year t, we sort stocks on asset growth, defined as total assets (Compustat annual item AT) for the fiscal year ending in year $t-1$ divided by total assets for the fiscal year ending in $t-2$ (Cooper *et al.*, 2008). Finally, to form the OP10 deciles, at the end of June of each year t, we sort stocks on operating profitability, which is measured with accounting data for the fiscal year ending in year $t-1$ and is revenue (Compustat annual item REVT) minus the cost of goods sold (Compustat annual item COGS), minus selling, general, and administrative expenses (Compustat annual item XSGA), minus interest expense (Compustat annual item XINT), all divided by book equity (Fama and French, 2015).

For each equity characteristic, we estimate the capital share (α) and the adjustment cost parameter (a) by fitting both the investment Euler equation moments and the valuation moments across that group of portfolio deciles. Then, given the parameter estimates, we construct the series of investment returns and the series of its two components (M and Q) for each decile portfolio.

Using the physical capital model with a depreciation rate of 0.1 and the firm-level aggregation as in Gonçalves, Xue, and Zhang (2020), we obtain GMM estimates of α as 0.08 (BM10), 0.07 (IA10), and 0.07 (OP10). The corresponding estimates of a are 15.83 (BM10), 15.74 (IA10), and 15.66 (OP10). Our estimates of α are similar to the estimates in Gonçalves *et al.* (2020) and also similar to our estimate in Section 3. Our estimates of a are higher than in Gonçalves *et al.* (2020) due to the fact that we fit additional moment conditions, namely the valuation moments.[24] As

[23] Following Davis *et al.* (2000), we measure book equity as stockholders' book equity, plus balance sheet deferred taxes and investment tax credit (Compustat annual item TXDITC) if available, minus the book value of preferred stock. Stockholders' equity is the value reported by Compustat (item SEQ), if it is available. If not, we measure stockholders' equity as the book value of common equity (item CEQ) plus the par value of preferred stock (item PSTK), or the book value of assets (item AT) minus total liabilities (item LT). Depending on availability, we use redemption (item PSTKRV), liquidating (item PSTKL), or par value (item PSTK) for the book value of preferred stock.

[24] As a quality assurance, we follow Gonçalves, Xue, and Zhang (2020) to fit only the investment Euler equation moments. We obtain estimates of α and a that are similar to theirs for the book-to-market and asset growth deciles.

robustness checks, we consider several variations of the GMM estimates. Similar to the findings in Gonçalves *et al.* (2020), using the aggregation approach as in Liu, Whited, and Zhang (2009) or the two-capital model as in Gonçalves *et al.* (2020) we obtain higher estimates of α but the estimates of a are not affected much. Using a higher depreciation rate of 0.12 only slightly changes the estimates of α and a.

The descriptive statistics associated with the extreme low (L) and high (H) deciles within each anomaly group (BML, BMH, IAL, IAH, OPL, and OPH) are reported in Table 5. We can see that all three variables in the VAR are more volatile among low-profitability firms (OPL) in comparison to high-profitability firms (OPH). A similar, albeit weaker, pattern holds for high-investment firms (IAH) in comparison to low-investment firms (IAL). In the case of the two extreme book-to-market deciles, the differences in volatility of the three variables are clearly less pronounced. It turns out that BMH, IAH, and OPL show a higher persistence in both m and q in comparison to the BML, IAL, and OPH portfolios, respectively.

Similarly to the case of the market portfolio discussed in Section 4, we find that m and q are strongly correlated, with correlations above 0.84 in all six cases. Nonetheless, these correlations are larger for BML and IAH in comparison to BMH and IAL, respectively. We can also see that the correlations between r and m are above 0.54 in all cases. Across extreme deciles, such correlation is larger among BML and IAH in comparison to BMH and IAL, respectively. Overall, these statistics show some differences for the key variables among the three anomaly groups and among the extreme deciles within each group.

6.2 Results

The results for the variance decompositions associated with each of the equity portfolios are presented in Table 6. To save space and keep the focus, we restrict the analysis to the "extreme" six portfolios described above. We can see that the discount rate channel is dominant for most equity portfolios, as the the long-run return slope estimates (in magnitude) are in most cases around or above 0.74 (0.70) when the estimation is based on the restricted (unrestricted) VAR. This indicates an even larger share of investment return predictability than that obtained for the market portfolio (61%). Indeed, when the estimates are based on the unrestricted VAR, for most portfolios we cannot reject (at the 10% level) the null that all the variation in q stems from return predictability ($b_r^{lr} = -1$). By comparing the portfolios associated with each sort, we find that the share of investment return predictability tends to be higher among low-investment firms in comparison to high-investment firms. A similar pattern holds for growth firms relative to value firms, although in this case, such patterns holds under the unrestricted VAR.

The key exception to the dominance of the discount-rate channel is the lowest OP decile. When the long-run coefficient estimates are based on the restricted VAR, we obtain relatively comparable shares of return and marginal profits predictability (0.57 versus 0.42) and both of these estimates are significant at the 5% level. When the long-run estimates are based on the unrestricted VAR, the results are even more extreme, as the weight associated with marginal profits predictability has almost twice the magnitude of the share of return predictability (0.64 versus 0.34). The long-run m slope estimate is marginally significant (10% level with a t-ratio of 1.89), while the return coefficient is largely insignificant (t-ratio of -0.91). Further, the null that all variation in q stems from marginal profits predictability, $b_m^{lr} = 1$, is not rejected at the 10% level (t-ratio of -1.06). Hence, the low-profitability portfolio appears as an outlier relative to all other five portfolios in the sense that the dynamics of the corresponding q is entirely driven by the cash-flow channel when the estimates are based on the unrestricted VAR.

Table 7 displays the results when the portfolio series of q, r, and m are based on the structural estimation method from Liu, Whited, and Zhang (2009). Across most portfolios, we find a slightly larger share of investment return predictability than in the benchmark case documented above and this pattern is specially evident when the estimates are based on the unrestricted VAR. This implies that the return channel is clearly dominant across most portfolios in terms of driving the time-series dynamics of q. The key exception is again the OPL portfolio. Under both the restricted and unrestricted VAR cases, we find that the shares of long-run return and marginal profits predictability are nearly the same (about 50%). We also find that both slope estimates are strongly significant (1% level), that is, both sources of long-run predictability drive the variation in q for that portfolio. In comparison to the results displayed in Table 6, there is a smaller role for marginal profits predictability under the unrestricted VAR (0.50 versus 0.64) for OPL. Nonetheless, such estimate is still largely significant in economic terms.

Overall, the results of this section suggest that what drives the variation in the q of different equity portfolio groups is the investment return channel, with the cash-flow channel being marginal in most cases. This result is consistent with the evidence obtained for the market portfolio in the previous sections. However, there are some important differences on the long-run predictability mix within the cross-section of firms. In particular, predictability of future marginal profits plays a critical role in explaining the variation of q for low-profitability firms.

7 Variance Decomposition for q: Cross-Sectional Analysis

In this section, we estimate a variance decomposition for average q in the cross-section of equity portfolios.

7.1 Methodology

We start by deriving a cross-sectional accounting decomposition for average q. The difference equation in q presented in Equation (14) can be defined for a given firm/portfolio i,

$$q_{i,t} \approx \rho q_{i,t+1} - r_{i,t+1} + (1-\rho) m_{i,t+1}, \tag{44}$$

where $q_{i,t}$ denotes the log Q for firm/portfolio i at time t, while $r_{i,t+1}$ and $m_{i,t+1}$ represent respectively the investment return and marginal profits for firm/portfolio i at time $t+1$. For simplicity, we assume that ρ is constant across firms/portfolios.

By assuming stationarity in $q_{i,t}$, $\mathrm{E}(q_{i,t}) = \mathrm{E}(q_{i,t+1})$, and rearranging, we have:

$$\mathrm{E}(q_{i,t}) \approx \mathrm{E}(m_{i,t+1}) - \frac{1}{1-\rho}\,\mathrm{E}(r_{i,t+1}). \tag{45}$$

This equation represents the cross-sectional analogue of the present-value relation for aggregate q in Equation (15): firms/portfolios that have a higher average q tend to register a higher average marginal profitability of capital and/or a lower average investment return. Hence, similarly to the dynamic present-value relation from Section 4, there is a positive (negative) correlation between q and marginal profitability (investment return).

To simplify notation, we define $\mathrm{E}(q_{i,t}) \equiv \mu_{q,i}$, $\mathrm{E}(m_{i,t+1}) \equiv \mu_{m,i}$, and $\mathrm{E}(r_{i,t+1}) \equiv \mu_{r,i}$, leading to

$$\mu_{q,i} \approx \mu_{m,i} - \frac{1}{1-\rho}\mu_{r,i}. \tag{46}$$

Multiplying both sides by $\mu_{q,i} - \mathrm{E}^*(\mu_{q,i})$, and taking expectations in the cross-section, we obtain,

$$\mathrm{var}^*(\mu_{q,i}) \approx \mathrm{Cov}^*(\mu_{m,i}, \mu_{q,i}) - \frac{1}{1-\rho}\,\mathrm{Cov}^*(\mu_{r,i}, \mu_{q,i}), \tag{47}$$

where $*$ denotes a cross-sectional moment.

Finally, dividing both sides by $\mathrm{var}^*(\mu_{q,i})$, we obtain the following cross-sectional variance decomposition for average q ($\mu_{q,i}$),

$$1 \approx c_m - \frac{c_r}{1-\rho}. \tag{48}$$

Under this decomposition, the cross-sectional dispersion in average q is attributed to cross-sectional investment return correlation (captured by the term $-c_r(1-\rho)$) and/or cross-sectional marginal profit correlation (given by c_m). Cochrane (2011) employs a similar cross-sectional variance decomposition for the average log dividend-to-price ratio.

The slopes in the previous equation are obtained from the following single cross-sectional regressions (estimated by OLS):

$$\mu_{r,i} = a_r + c_r \mu_{q,i} + \zeta_{r,i}, \quad (49)$$

$$\mu_{m,i} = a_m + c_m \mu_{q,i} + \zeta_{m,i}. \quad (50)$$

Hence, these cross-sectional regressions play the role of the first-order VAR in the case of the time-series variance decompositions estimated in the previous sections. This cross-sectional decomposition shares a property with the time-series decomposition. Since $1 - \rho$ is a number substantially below one, the decomposition is a priori "forced" to produce a larger share for the return channel (captured by $-c_r/(1-\rho)$) than for the cash-flow channel (captured by c_m).

7.2 Results

We start by presenting summary statistics for the variables of interest (μ_r, μ_m, and μ_q) across the BM10, IA10, and OP10 decile groups. The results in Table 8 show that all three variables are more volatile within BM10 in comparison to the corresponding variables within IA10 and especially OP10. We also observe that the cross-sectional means of μ_r, μ_m, and μ_q do not vary much across the three portfolio groups. Regarding the linear association among the variables, we can see that both μ_r and μ_m are strongly correlated with μ_q within BM10 or IA10, as indicated by the large correlation coefficients (around or above 0.90). In the case of the OP10 deciles, the correlation between μ_r and μ_q is substantially more modest (0.41).

The results for the cross-sectional variance decomposition for average q are displayed in Table 9, Panel A. We conduct cross-sectional regressions containing all 30 portfolios (BM10+IA10+OP10). This ensures a higher statistical power of the resulting estimates in comparison to conducting the regressions for each group of deciles. Given the relatively small sample size of our cross-section (30 portfolios), we report both OLS and heteroskedasticity-robust (White, 1980) t-ratios. The estimate of ρ used in the analysis is based on a simple cross-sectional mean (across the 30 portfolios) of $mq_{i,t}$,

$$\rho \equiv \frac{e^{\ln(1-\delta)-\overline{\widehat{mq}}}}{1+e^{\ln(1-\delta)-\overline{\widehat{mq}}}},$$

where $\overline{\widehat{mq}}$ represents the time-series average of $\widehat{mq}_t = (1/30)\sum_{i=1}^{30} mq_{i,t}$.

We can see that the marginal profit slope estimate is clearly above one (2.18) and strongly significant (1% level) based on both t-ratios. The R^2 of 0.89 indicates a large fit in the cross-sectional regression associated with average m. The large magnitude of c_m stems from the estimate of c_r assuming the wrong sign (0.18). Hence, the weight associated with average r is negative, implying that from an economic viewpoint all the cross-sectional variation in average q stems from the cross-sectional correlation between μ_q and average m (rather than average r). The sum of the variance decomposition, $c_m - c_r/(1-\rho)$, is relatively close to one at 0.92. Therefore, these results clearly indicate that firms with higher average q are firms with higher average marginal profitability of capital instead of lower average investment return.

Panel B of Table 9 displays the results when the portfolio structural investment series are estimated under the method of Liu, Whited, and Zhang (2009). The magnitudes of the slope estimates are smaller than in the benchmark case: μ_m and μ_r are 1.71 and 0.12, respectively, with strong significance in both cases. The corresponding shares in the decomposition are 1.71 and -0.71, respectively, which implies that the sum of the decomposition is about one. Hence, the approximation of the variance decomposition is better than in the benchmark case.

In Panel C of Table 9, all the investment series are based on a different calibration of the depreciation rate parameter ($\delta = 0.1219$). The results are quite similar to those in the benchmark case, with shares of marginal profits and return predictability of 2.07 and -1.13, respectively. This implies that $c_m - c_r/(1-\rho)$ is marginally closer to one than in the baseline case (0.94 versus 0.92). In Panel D, the cross-sectional regression includes the Tobin's Q deciles instead of the BM deciles. The shares associated with μ_m and μ_r are 2.10 and -1.27, respectively. This implies that $c_m - c_r/(1-\rho)$ is more distant from one (0.86) than in the baseline regression, which means that the variance decomposition is less accurate in this setup. Nonetheless, in economic terms, all the variation in average q stems from the cash-flow channel.

These findings are in sharp contrast with the evidence for the time-series decompositions for q analyzed in the previous sections. As discussed above, both decompositions have an ex ante tilt towards the investment return channel. What makes the cross-sectional decomposition distinct is that the cross-sectional return slope has the wrong sign, that is, average q is positively (rather than negatively) correlated with average r. This implies that the positive correlation between average q and average m drives all the cross-sectional dispersion in the first variable in economic terms.

Figure 5 displays the pairs of observed μ_m and μ_r and the fitted values from the corresponding regressions in the benchmark setup (Panel A of Table 9). We can see that the fit of the cross-sectional regression associated with μ_m is clearly higher than that for μ_r, as the pairs of observed and fitted values are much closer to the diagonal line. This represents another way to illustrate the higher R^2 in the regression associated with average m. The plot of the regression for μ_r illustrates well that such variable is positively correlated with average q, exactly the opposite pattern postulated by the cross-sectional variance decomposition. The main outlier in both regressions is the first OP decile, which generates the largest positive errors among all portfolios. This finding is consistent with the evidence presented in Section 6, which shows that the time-series variance decomposition for that portfolio deviates substantially from that of other portfolios.

Overall, the message from this section is clear: the cash-flow channel is the source of cross-sectional variation in average q, that is, firms with higher average q are firms with higher average marginal profitability of capital, instead of being firms with lower average investment return. This result is in clear contrast with the findings for the time-series analysis documented in Sections 4 and 6, in which the discount-rate channel explains most of the variation of q (associated with either the market or equity portfolios) over time. Considering that in our specification marginal Q is a linear function of investment, our results in this section are consistent with the empirical investment literature that finds a positive investment-to-cash-flow sensitivity. However, the novel fact we document is the lack of explanatory power for average investment return in driving the cross-sectional variation in average marginal Q and investment. That is, in the cross-section discount rates play no role.

8 Conclusion

This paper explores the sources of time-series fluctuations in aggregate and portfolio-level realized marginal Q (and investment). In addition, the paper studies the determinants of the cross-section of portfolios' average marginal Q (and average investment). The extant literature studies intensively the sources of variation in (scaled) stock prices but the drivers of marginal Q (and investment) are relatively unexplored. The approach we undertake is the supply approach to valuation. That is, we infer marginal Q from the marginal adjustment cost of investment, as in Belo, Xue, and Zhang (2013).

We employ a parsimonious model of optimal investment behavior with standard production and adjustment cost technologies. The model's optimal investment condition, namely that the marginal value of capital equals the marginal cost of investment, implies that marginal Q is a linear function of the investment-to-capital ratio. Subsequently, we estimate the share of capital and adjustment cost parameters by employing GMM estimation for the aggregate of firms on the Compustat database. Following Gonçalves, Xue, and Zhang (2020), we correct for aggregation bias. We then derive a present-value relation to show that variations in the log of marginal Q (q) (and hence in investment) must reflect shocks to expected future marginal profits of capital, shocks to expected investment returns, or the future value of q at some terminal date, or any combination of these three variables.

We conduct variance decomposition for q using alternative methodologies, namely a first-order restricted VAR and an unrestricted VAR. We find that the bulk of variations in q is due to investment return predictability, whereas predictability of the marginal profits of capital assumes a secondary role (albeit statistically significant in some cases). We conduct several robustness checks, namely using portfolios sorted by Tobin's Q as well as using portfolio medians in the GMM estimation of the structural model; conducting simulation exercises in the variance decomposition; employing weighted long-horizon regressions to estimate the decomposition; estimating the variance decomposition from a second-order VAR; varying the value of the depreciation rate; and conducting the GMM estimation without correcting for aggregation bias. Our main qualitative results are robust in all those checks.

In addition to studying the sources of the time-series variation of q of the aggregate of firms we also explore the determinants of the time-series fluctuations in portfolio-level q. We find that like in the aggregate level, the time-series variation of portfolios' q are largely due to revisions to discount rates as measured by expected investment returns.

Finally, we estimate a variance decomposition for average q in the cross-section of 30 portfolios sorted on book-to-market, asset growth, and profitability. We find that the cash-flow channel (namely marginal profits) is the dominant channel in describing the cross-section of marginal Q and investment. That is, firms with higher average q are firms with higher average marginal profitability of capital, instead of being firms with lower average investment return. This result is markedly different from the time-series results.

A possible extension of our work is to compute time-series and cross-sectional variance decompositions respectively for the realized and average q of individual firms rather than for portfolios of firms. This is left for future research.

Table 1: Descriptive statistics

This table reports descriptive statistics for the log investment return (r), log marginal profits (m), and log Q (q) associated with the market portfolio. The sample is 1964–2018. AR(1) designates the first-order autocorrelation. The correlations between the variables are presented in Panel B.

Panel A					
	Mean	S.D.	Min.	Max.	AR(1)
r	0.06	0.10	−0.19	0.31	0.12
m	−0.85	0.24	−1.30	0.03	0.54
q	1.03	0.10	0.85	1.29	0.61
Panel B (Correl.)					
	r	m	q		
r	1.00	0.62	0.57		
m		1.00	0.95		
q			1.00		

Table 2: VAR estimates

This table reports the VAR(1) estimation results when the predictor is log Q (q). The variables in the VAR are the log investment return (r), log marginal profits (m), and q. The results in Panel A correspond to the baseline set of investment variables. Results in Panels B to E correspond to alternative sets of the investment variables. λ, ϕ denote the VAR slopes associated with lagged q, while t denotes the respective Newey and West (1987) t-statistics (calculated with one lag). R^2 is the coefficient of determination for each equation in the VAR. b^{lr} denote the long-run coefficients (infinite horizon). $t(b_r^{lr} = 0)$ and $t(b_r^{lr} = -1)$ denote the t-statistics associated with the null hypotheses ($b_r^{lr} = 0, b_m^{lr} = 1$) and ($b_r^{lr} = -1, b_m^{lr} = 0$), respectively. The original sample is 1964–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	λ, ϕ	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = -1)$
			Panel A			
r	−0.29	<u>−2.32</u>	0.08	−0.61	**−3.96**	<u>2.52</u>
m	1.24	**4.58**	0.28	0.37	**−4.24**	<u>2.54</u>
q	0.61	**5.91**	0.35			
			Panel B (Q Deciles)			
r	−0.30	<u>−2.38</u>	0.09	−0.62	**−4.16**	<u>2.54</u>
m	1.21	**4.63**	0.28	0.36	**−4.48**	<u>2.55</u>
q	0.61	**5.93**	0.35			
			Panel C (Median firm)			
r	−0.29	<u>−2.31</u>	0.08	−0.61	**−3.99**	<u>2.51</u>
m	1.27	**4.59**	0.28	0.37	**−4.30**	<u>2.52</u>
q	0.61	**5.91**	0.35			
			Panel D (Alternative δ)			
r	−0.30	<u>−2.33</u>	0.08	−0.60	**−3.90**	<u>2.56</u>
m	1.17	**4.53**	0.28	0.38	**−4.14**	<u>2.57</u>
q	0.61	**5.91**	0.35			
			Panel E (Alternative series)			
r	−0.35	**−3.30**	0.13	−0.71	**−6.52**	**2.71**
m	0.82	**4.38**	0.23	0.30	**−6.49**	**2.73**
q	0.61	**6.46**	0.36			

Table 3: Simulation with restricted VAR

This table reports the simulation results for the long-run variance decomposition associated with log Q (q). The simulated series for the log investment return (r), log marginal profits (m), and q are based on different pairs of the calibrated structural parameters α and a from the theoretical model. The implied long-run predictive statistics are based on a restricted VAR(1). b^{lr} denote the long-run coefficients (infinite horizon), while t represent the corresponding t-statistics. The original sample is 1964–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	b_r^{lr}	t	b_m^{lr}	t
$\alpha = 0.05, a = 1.37$	−0.39	(−1.37)	0.60	($\underline{2.15}$)
$\alpha = 0.05, a = 6.85$	−0.58	(**−3.38**)	0.40	($\underline{2.43}$)
$\alpha = 0.05, a = 13.69$	−0.61	(**−3.95**)	0.37	($\underline{2.50}$)
$\alpha = 0.05, a = 20.53$	−0.63	(**−4.20**)	0.36	($\underline{2.53}$)
$\alpha = 0.05, a = 27.38$	−0.63	(**−4.34**)	0.35	($\underline{2.54}$)
$\alpha = 0.15, a = 1.37$	0.08	(0.15)	1.09	($\underline{2.13}$)
$\alpha = 0.15, a = 6.85$	−0.52	(**−2.66**)	0.48	($\underline{2.49}$)
$\alpha = 0.15, a = 13.69$	−0.59	(**−3.66**)	0.40	(**2.58**)
$\alpha = 0.15, a = 20.53$	−0.61	(**−4.06**)	0.38	(**2.60**)
$\alpha = 0.15, a = 27.38$	−0.62	(**−4.27**)	0.37	(**2.61**)
$\alpha = 0.30, a = 1.37$	0.50	(0.71)	1.53	($\underline{2.15}$)
$\alpha = 0.30, a = 6.85$	−0.44	($\underline{-1.99}$)	0.56	($\underline{2.48}$)
$\alpha = 0.30, a = 13.69$	−0.56	(**−3.27**)	0.44	(**2.61**)
$\alpha = 0.30, a = 20.53$	−0.59	(**−3.84**)	0.40	(**2.65**)
$\alpha = 0.30, a = 27.38$	−0.61	(**−4.13**)	0.39	(**2.67**)
$\alpha = 0.50, a = 1.37$	0.82	(0.98)	1.86	($\underline{2.18}$)
$\alpha = 0.50, a = 6.85$	−0.38	(−1.50)	0.63	($\underline{2.45}$)
$\alpha = 0.50, a = 13.69$	−0.53	(**−2.88**)	0.48	(**2.60**)
$\alpha = 0.50, a = 20.53$	−0.57	(**−3.56**)	0.43	(**2.67**)
$\alpha = 0.50, a = 27.38$	−0.59	(**−3.94**)	0.41	(**2.70**)
$\alpha = 0.70, a = 1.37$	1.02	(1.11)	2.07	($\underline{2.21}$)
$\alpha = 0.70, a = 6.85$	−0.34	(−1.23)	0.68	($\underline{2.43}$)
$\alpha = 0.70, a = 13.69$	−0.50	(**−2.60**)	0.50	(**2.58**)
$\alpha = 0.70, a = 20.53$	−0.56	(**−3.32**)	0.45	(**2.66**)
$\alpha = 0.70, a = 27.38$	−0.58	(**−3.75**)	0.42	(**2.70**)

Table 4: Simulation with unrestricted VAR

This table reports the simulation results for the long-run variance decomposition associated with log Q (q). The simulated series for the log investment return (r), log marginal profits (m), and q are based on different pairs of the calibrated structural parameters α and a from the theoretical model. The implied long-run predictive statistics are based on an unrestricted VAR(1). b^{lr} denote the long-run coefficients (infinite horizon), while t represent the corresponding t-statistics. The original sample is 1964–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	b_r^{lr}	t	b_m^{lr}	t
$\alpha = 0.05, a = 1.37$	−0.37	(−0.83)	0.62	(1.42)
$\alpha = 0.05, a = 6.85$	−0.78	$(-\mathbf{4.14})$	0.19	(1.11)
$\alpha = 0.05, a = 13.69$	−0.82	$(-\mathbf{5.19})$	0.16	(1.10)
$\alpha = 0.05, a = 20.53$	−0.83	$(-\mathbf{5.60})$	0.15	(1.10)
$\alpha = 0.05, a = 27.38$	−0.83	$(-\mathbf{5.83})$	0.14	(1.10)
$\alpha = 0.15, a = 1.37$	0.33	(0.79)	1.35	$(\mathbf{3.24})$
$\alpha = 0.15, a = 6.85$	−0.50	$(-\mathit{1.78})$	0.49	$(\mathit{1.77})$
$\alpha = 0.15, a = 13.69$	−0.70	$(-\mathbf{3.19})$	0.29	(1.40)
$\alpha = 0.15, a = 20.53$	−0.76	$(-\mathbf{4.15})$	0.23	(1.30)
$\alpha = 0.15, a = 27.38$	−0.79	$(-\mathbf{4.74})$	0.20	(1.26)
$\alpha = 0.30, a = 1.37$	0.59	(1.41)	1.64	$(\mathbf{3.88})$
$\alpha = 0.30, a = 6.85$	−0.39	$(-\mathit{1.79})$	0.62	$(\mathbf{2.90})$
$\alpha = 0.30, a = 13.69$	−0.55	$(\underline{-2.42})$	0.45	$(\underline{1.99})$
$\alpha = 0.30, a = 20.53$	−0.64	$(-\mathbf{2.94})$	0.35	$(\mathit{1.65})$
$\alpha = 0.30, a = 27.38$	−0.70	$(-\mathbf{3.46})$	0.30	(1.51)
$\alpha = 0.50, a = 1.37$	0.80	$(\mathit{1.84})$	1.86	$(\mathbf{4.25})$
$\alpha = 0.50, a = 6.85$	−0.36	$(\underline{-2.05})$	0.66	$(\mathbf{3.74})$
$\alpha = 0.50, a = 13.69$	−0.50	$(-\mathbf{2.76})$	0.50	$(\mathbf{2.79})$
$\alpha = 0.50, a = 20.53$	−0.56	$(-\mathbf{2.86})$	0.44	$(\underline{2.23})$
$\alpha = 0.50, a = 27.38$	−0.61	$(-\mathbf{3.01})$	0.39	$(\mathit{1.91})$
$\alpha = 0.70, a = 1.37$	0.94	$(\underline{2.14})$	2.01	$(\mathbf{4.49})$
$\alpha = 0.70, a = 6.85$	−0.34	$(\underline{-2.11})$	0.68	$(\mathbf{4.16})$
$\alpha = 0.70, a = 13.69$	−0.49	$(-\mathbf{3.19})$	0.52	$(\mathbf{3.36})$
$\alpha = 0.70, a = 20.53$	−0.54	$(-\mathbf{3.24})$	0.46	$(\mathbf{2.75})$
$\alpha = 0.70, a = 27.38$	−0.58	$(-\mathbf{3.18})$	0.42	$(\underline{2.35})$

Table 5: Descriptive statistics for equity portfolios

This table reports descriptive statistics for the log investment return (r), log marginal profits (m), and log Q (q) associated with alternative equity portfolios. BML, IAL, and OPL represents respectively the lowest decile among the book-to-market, asset growth, and profitability portfolios. BMH, IAH, and OPH represents respectively the highest decile among the book-to-market, asset growth, and profitability portfolios. The sample is 1964–2018. AR(1) designates the first-order autocorrelation. $Corr(x, q)$ designates the correlation between either r or m and q. $Corr(x, m)$ designates the correlation between r and m.

	Mean	S.D.	Min.	Max.	AR(1)	$Corr(x, q)$	$Corr(x, m)$
				Panel A (BML)			
r	0.12	0.19	−0.23	0.71	−0.13	0.62	0.73
m	−0.17	0.35	−0.93	1.05	0.38	0.93	
q	1.32	0.15	0.96	1.74	0.49		
				Panel B (BMH)			
r	0.04	0.17	−0.53	0.52	−0.17	0.54	0.54
m	−1.16	0.35	−1.89	0.31	0.58	0.87	
q	0.86	0.17	0.54	1.49	0.55		
				Panel C (IAL)			
r	0.06	0.21	−0.37	0.63	−0.44	0.65	0.56
m	−0.86	0.35	−1.58	0.24	0.28	0.84	
q	0.98	0.17	0.52	1.46	0.35		
				Panel D (IAH)			
r	0.12	0.25	−0.71	0.93	−0.28	0.62	0.67
m	−0.30	0.45	−1.08	1.31	0.38	0.96	
q	1.28	0.21	0.85	1.94	0.47		
				Panel E (OPL)			
r	0.12	0.27	−0.64	0.90	−0.36	0.53	0.63
m	−0.56	0.58	−1.78	1.39	0.51	0.94	
q	1.04	0.27	0.60	1.94	0.64		
				Panel F (OPH)			
r	0.08	0.17	−0.22	0.57	−0.12	0.58	0.69
m	−0.49	0.33	−1.06	0.19	0.40	0.93	
q	1.21	0.15	0.86	1.58	0.53		

Table 6: VAR estimates for equity portfolios

This table reports the VAR-based variance decompositions associated with alternative equity portfolios. The variables in the VAR are the log investment return (r), log marginal profits (m), and log Q (q). Results in Panels A to F refer to the restricted VAR(1) case, in which the sole predictor is q. In Panels G to L, each variance decompositions is based on an unrestricted VAR(1). BML, IAL, and OPL represents respectively the lowest decile among the book-to-market, asset growth, and profitability portfolios. BMH, IAH, and OPH represents respectively the highest decile among the book-to-market, asset growth, and profitability portfolios. b^{lr} denote the long-run coefficients (infinite horizon). $t(b_r^{lr} = 0)$ and $t(b_r^{lr} = -1)$ denote the t-statistics associated with the null hypotheses ($b_r^{lr} = 0, b_m^{lr} = 1$) and ($b_r^{lr} = -1, b_m^{lr} = 0$), respectively. The original sample is 1964–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = -1)$		b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = -1)$
	Panel A: BML				**Panel G: BML**		
r	−0.75	**−6.41**	<u>2.15</u>	r	−0.85	**−6.25**	1.10
m	0.26	**−6.44**	<u>2.22</u>	m	0.16	**−6.48**	1.19
	Panel B: BMH				**Panel H: BMH**		
r	−0.74	**−5.33**	*1.85*	r	−0.80	**−6.22**	1.55
m	0.24	**−6.11**	*1.92*	m	0.19	**−6.94**	1.60
	Panel C: IAL				**Panel I: IAL**		
r	−0.82	**−14.02**	**3.01**	r	−0.75	**−8.34**	**2.80**
m	0.17	**−15.48**	**3.06**	m	0.23	**−9.01**	**2.72**
	Panel D: IAH				**Panel J: IAH**		
r	−0.74	**−5.94**	<u>2.10</u>	r	−0.70	**−4.07**	*1.75*
m	0.26	**−6.03**	<u>2.08</u>	m	0.27	**−4.54**	*1.71*
	Panel E: OPL				**Panel K: OPL**		
r	−0.57	**−2.59**	*1.94*	r	−0.34	−0.91	*1.79*
m	0.42	**−2.77**	<u>2.01</u>	m	0.64	−1.06	*1.89*
	Panel F: OPH				**Panel L: OPH**		
r	−0.76	**−5.17**	*1.68*	r	−0.86	**−6.90**	1.10
m	0.25	**−5.27**	*1.74*	m	0.14	**−6.98**	1.16

Table 7: VAR estimates for equity portfolios: Alternative series

This table reports the VAR-based variance decompositions associated with alternative equity portfolios. The variables in the VAR are the log investment return (r), log marginal profits (m), and log Q (q). The series of r, m, and q are based on the Liu–Whited–Zhang structural estimation method. Results in Panels A to F refer to the restricted VAR(1) case, in which the sole predictor is q. In Panels G to L, each variance decompositions is based on an unrestricted VAR(1). BML, IAL, and OPL represents respectively the lowest decile among the book-to-market, asset growth, and profitability portfolios. BMH, IAH, and OPH represents respectively the highest decile among the book-to-market, asset growth, and profitability portfolios. b^{lr} denote the long-run coefficients (infinite horizon). $t(b_r^{lr} = 0)$ and $t(b_r^{lr} = -1)$ denote the t-statistics associated with the null hypotheses $(b_r^{lr} = 0, b_m^{lr} = 1)$ and $(b_r^{lr} = -1, b_m^{lr} = 0)$, respectively. The original sample is 1964–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = -1)$		b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = -1)$
	Panel A: BML				**Panel G: BML**		
r	−0.78	**−7.03**	*1.96*	r	−0.93	**−8.97**	0.63
m	0.22	**−7.00**	$\underline{1.96}$	m	0.07	**−8.90**	0.65
	Panel B: BMH				**Panel H: BMH**		
r	−0.80	**−8.34**	$\underline{2.03}$	r	−0.92	**−8.14**	0.67
m	0.20	**−8.32**	$\underline{2.05}$	m	0.08	**−7.94**	0.66
	Panel C: IAL				**Panel I: IAL**		
r	−0.89	**−18.62**	$\underline{2.31}$	r	−0.81	**−9.58**	$\underline{2.27}$
m	0.12	**−18.62**	$\underline{2.42}$	m	0.20	**−9.37**	$\underline{2.36}$
	Panel D: IAH				**Panel J: IAH**		
r	−0.83	**−10.06**	$\underline{2.04}$	r	−0.95	**−10.18**	0.50
m	0.17	**−9.91**	$\underline{2.02}$	m	0.05	**−9.95**	0.47
	Panel E: OPL				**Panel K: OPL**		
r	−0.53	**−3.79**	**3.32**	r	−0.51	**−2.72**	**2.65**
m	0.47	**−3.79**	**3.32**	m	0.50	**−2.70**	**2.65**
	Panel F: OPH				**Panel L: OPH**		
r	−0.73	**−5.79**	$\underline{2.14}$	r	−0.92	**−11.21**	0.98
m	0.27	**−5.89**	$\underline{2.14}$	m	0.08	**−11.18**	0.94

Table 8: Descriptive statistics for time-series means

This table reports descriptive statistics for the average log investment return (μ_r), average log marginal profits (μ_m), and average log Q (μ_q) across equity portfolios. The portfolios represent deciles sorted on book-to-market ratio (BM10), asset growth (IA10), and operating profitability (OP10). The sample is 1964–2018. $Corr(x, \mu_q)$ designates the correlation between either μ_r or μ_m and μ_q.

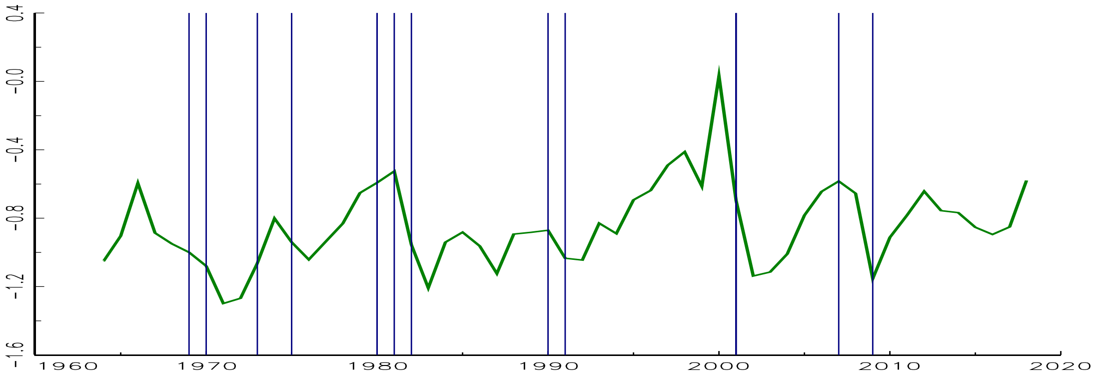
	Mean	S.D.	Min.	Max.	$Corr(x, \mu_q)$
Panel A (BM10)					
μ_r	0.06	0.03	0.03	0.12	0.94
μ_m	−0.80	0.34	−1.16	−0.17	0.99
μ_q	1.07	0.14	0.86	1.32	
Panel B (IA10)					
μ_r	0.05	0.03	0.03	0.12	0.90
μ_m	−0.86	0.28	−1.15	−0.29	0.96
μ_q	1.05	0.12	0.94	1.27	
Panel C (OP10)					
μ_r	0.06	0.02	0.04	0.12	0.41
μ_m	−0.83	0.21	−1.08	−0.50	0.85
μ_q	1.04	0.10	0.92	1.21	

Table 9: Cross-sectional variance decomposition

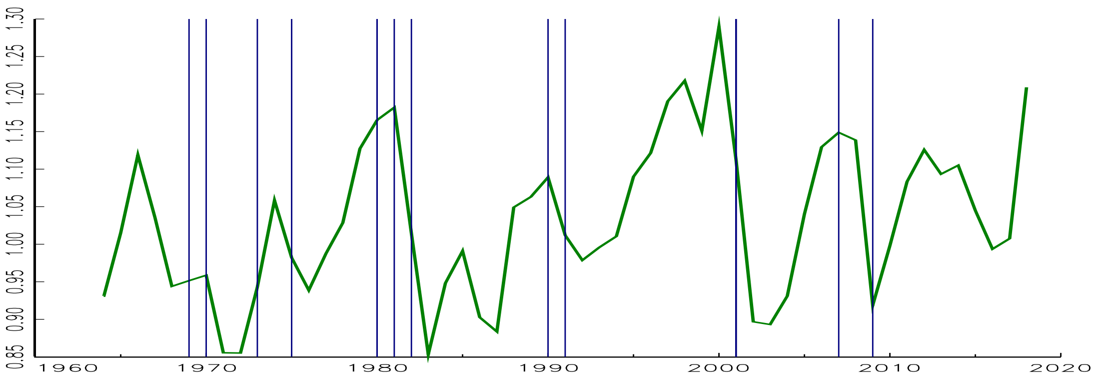
This table reports the results for cross-sectional regressions of either the average investment return (μ_r) or the average marginal profitability of capital (μ_m) onto the average q. c denotes the slope estimate from the corresponding regression. t_{OLS} (t_W) represents the respective OLS (heteroskedasticity-robust) t-statistic. R^2 is the coefficient of determination. "weight" represents the share associated with either return or marginal profits predictability in the cross-section. ρ is the log-linearization coefficient. The cross-section consists of decile portfolios sorted on book-to-market ratio, asset growth, and operating profitability for a total of 30 portfolios. The original sample is 1964–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively. Panel B shows the results when the series of the investment return, marginal profitability, and q are based on the Liu–Whited–Zhang structural estimation method. Panel C refers to the results based on series constructed from an alternative calibrated depreciation rate. In Panel D, the cross-section consists of portfolios sorted on asset growth, operating profitability, and Tobin's Q.

	c	t_{OLS}	t_W	R^2	weight	$c_m - c_r/(1-\rho)$	ρ
				Panel A			
μ_r	0.18	7.06	8.96	0.64	-1.26	0.92	0.86
μ_m	2.18	15.13	19.44	0.89	2.18		
				Panel B (Alternative series)			
μ_r	0.12	5.84	5.54	0.55	-0.71	1.00	0.83
μ_m	1.71	15.29	13.51	0.89	1.71		
				Panel C (Alternative δ)			
μ_r	0.18	6.90	9.20	0.63	-1.13	0.94	0.84
μ_m	2.07	15.16	20.59	0.89	2.07		
				Panel D (Alternative portfolios)			
μ_r	0.18	7.33	8.75	0.66	-1.27	0.84	0.86
μ_m	2.10	15.12	19.74	0.89	2.10		

Panel A (r)



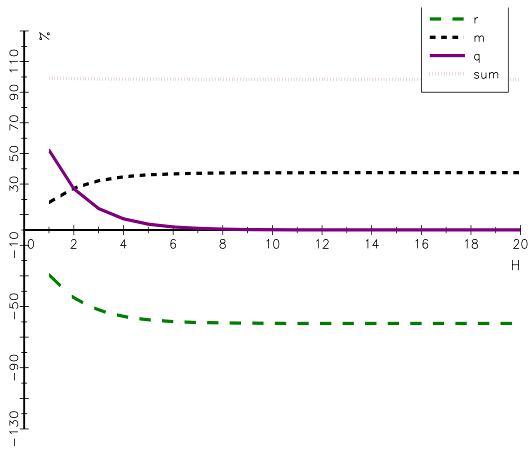
Panel B (m)



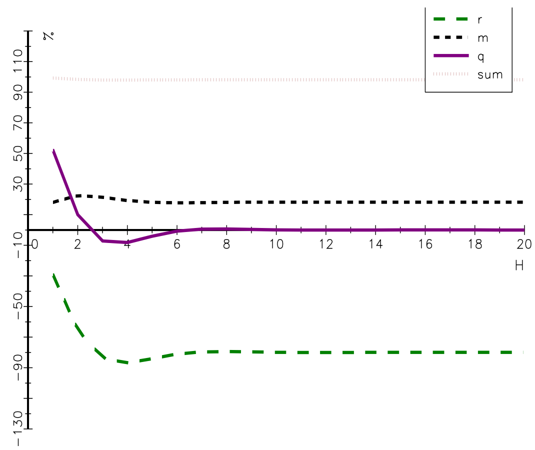
Panel C (q)

Figure 1: Time-Series for r, m, and q

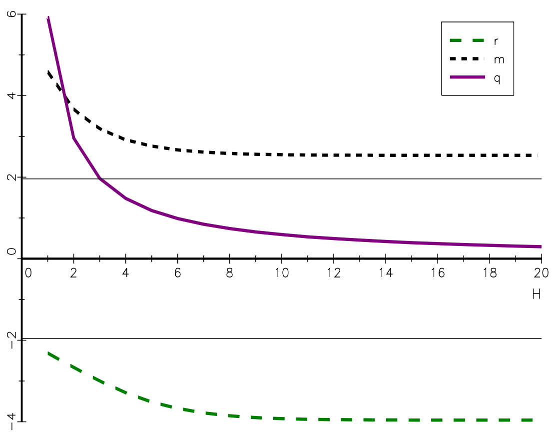
This figure plots the time-series for the log investment return (r), log marginal profit (m), and log Q (q). The bars contain the years with NBER recessions (the 1980 and 2001 recessions are indicated by a single line). The sample is 1964 to 2018.



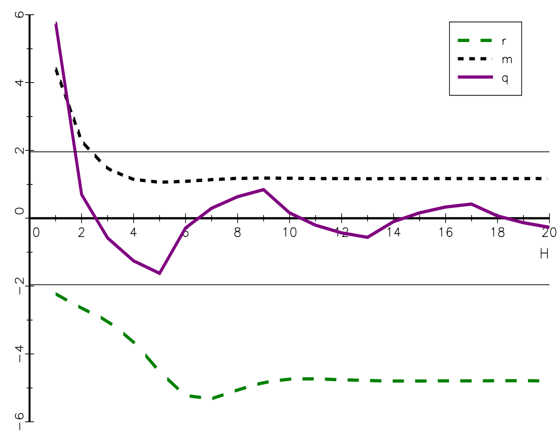
Panel A (Rest. VAR, slopes)



Panel B (Unrest. VAR, slopes)



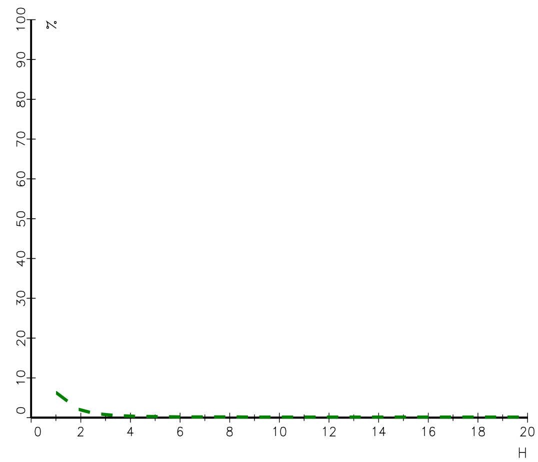
Panel C (Rest. VAR, *t*-stats)



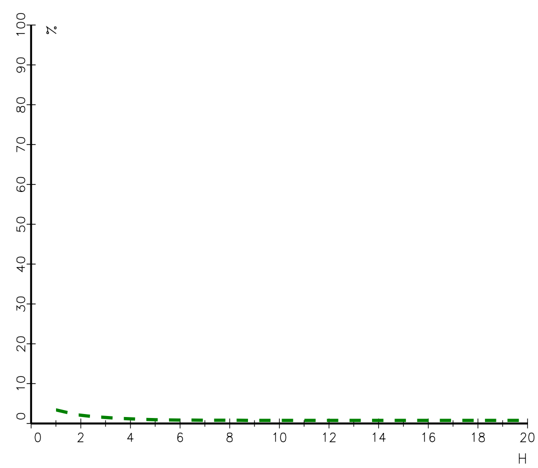
Panel D (Unrest. VAR, *t*-stats)

Figure 2: Variance decomposition

This figure plots the term structure of multiple-horizon predictive coefficients (in %), and respective t-statistics, corresponding to the variance decompositions for log Q (q). The predictive slopes are obtained from either a restricted or an unrestricted first-order VAR. The coefficients are associated with the log investment return (r), log marginal profits (m), and future q. The forecasting variable is q in all three cases. "Sum" denotes the value of the variance decomposition. H represents the number of years ahead. The horizontal lines represent the 5% critical values (-1.96, 1.96). The original sample is 1964 to 2018.



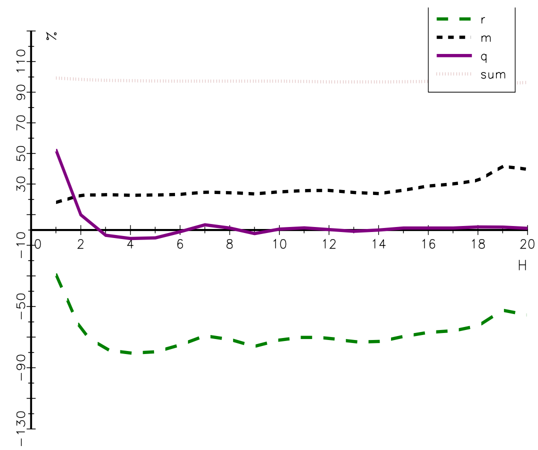
Panel A (r)



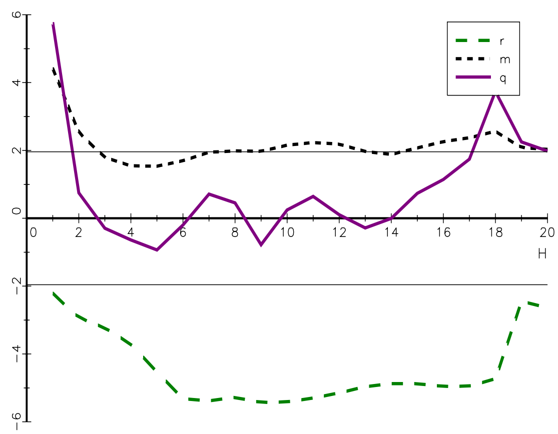
Panel B (m)

Figure 3: Bootstrap simulation

This figure plots the simulated p-values for the restricted VAR-based return (r) and profitability (m) slopes from a Bootstrap simulation with 10,000 replications. The predictive variable is log Q (q). The numbers indicate the fraction of pseudo samples under which the return (profitability) coefficient is lower (higher) than the corresponding estimates from the original sample. H represents the number of years ahead. The original sample is 1964 to 2018.



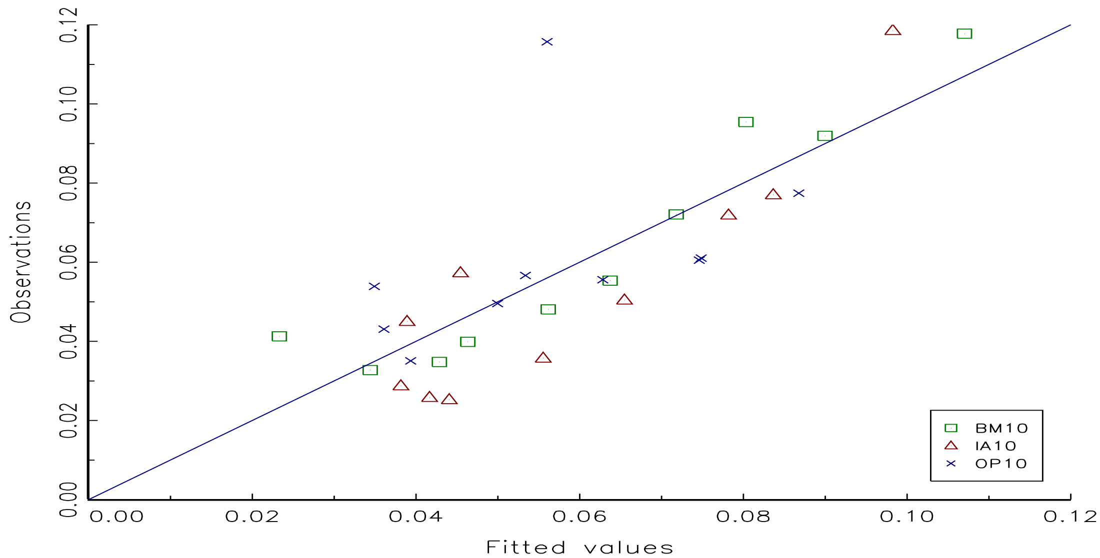
Panel A (slopes)



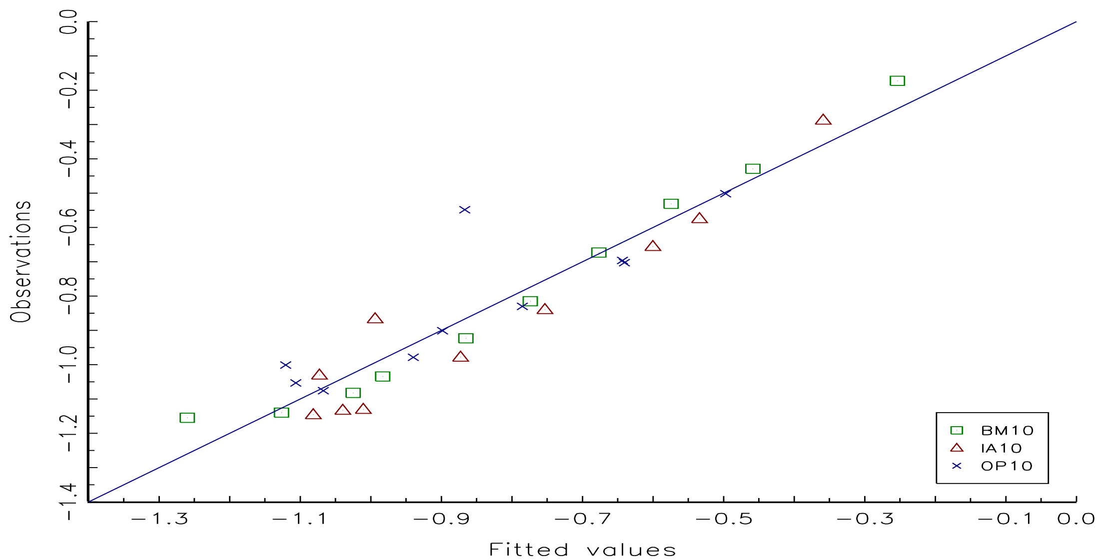
Panel B (t-stats)

Figure 4: Variance decomposition: Direct approach

This figure plots the term structure of multiple-horizon predictive coefficients (in %), and respective t-statistics, corresponding to the variance decompositions for log Q (q). The predictive slopes are obtained from weighted long-horizon regressions. The coefficients are associated with the log investment return (r), log marginal profits (m), and future q. The forecasting variable is q in all three cases. "Sum" denotes the value of the variance decomposition. H represents the number of years ahead. The horizontal lines represent the 5% critical values (-1.96, 1.96). The original sample is 1964 to 2018.



Panel A (μ_r)



Panel B (μ_m)

Figure 5: Cross-sectional regressions

This figure plots the observed versus fitted values in the cross-sectional regressions associated with average investment return (Panel A) and average marginal profits (Panel B) as dependent variables. The sample represents decile portfolios sorted on book-to-market ratio (BM10), asset growth (IA10), and operating profitability (OP10).

References

Abel, A.B. and Blanchard, O.J., 1986. The present value of profits and cyclical movements in investment, *Econometrica*, 54, 249–273.

Anderson, C.W. and Garcia-Feijoo, L., 2006. Empirical evidence on capital investment, growth options, and security returns, *Journal of Finance*, 61, 171–194.

Barkai, S., 2020. Declining labor and capital shares, *Journal of Finance*, 75, 2421–2463.

Basu, S. and Fernald, J.G., 1997. Returns to scale in u.s. production: Estimates and implications, *Journal of Political Economy*, 105, 249–283.

Belo, F., Gala, V.D., Salomao, J., and Vitorino, M.A., 2022. Decomposing firm value, *Journal of Financial Economics*, 143, 619–639.

Belo, F., Xue, C., and Zhang, L., 2013. A supply approach to valuation, *Review of Financial Studies*, 26, 3029–3067.

Bloom, N., 2009. The impact of uncertainty shocks, *Econometrica*, 77, 623–685.

Boudoukh, J., Richardson, M., and Whitelaw, R.F., 2008. The myth of long-horizon predictability, *Review of Financial Studies*, 21, 1577–1605.

Campbell, J.Y., 1991. A variance decomposition for stock returns, *Economic Journal*, 101, 157–179.

Campbell, J.Y. and Ammer, J., 1993. What moves the stock and bond markets? A variance decomposition for long-term asset returns, *Journal of Finance*, 48, 3–37.

Campbell, J.Y., Lo, A.W., and MacKinlay, A.C., 1997. *The econometrics of financial markets*, Princeton University Press Princeton, NJ.

Campbell, J.Y. and Shiller, R.J., 1988. The dividend price ratio and expectations of future dividends and discount factors, *Review of Financial Studies*, 1, 195–228.

Campbell, J.Y. and Vuolteenaho, T., 2004. Bad beta, good beta, *American Economic Review*, 94, 1249–1275.

Chen, L., Da, Z., and Larrain, B., 2016. What moves investment growth?, *Journal of Money, Credit and Banking*, 48, 1613–1653.

Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance*, 46, 209–237.

Cochrane, J.H., 2008. The dog that did not bark: A defense of return predictability, *Review of Financial Studies*, 21, 1533–1575.

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

Cohen, R.B., Polk, C., and Vuolteenaho, T., 2003. The value spread, *Journal of Finance*, 58, 609–641.

Cooper, I., Maio, P., and Yang, C., 2021. The marginal profits-to-q ratio: Reassessing the cash-flow channel, *SSRN Working Paper*.

Cooper, I. and Priestley, R., 2016. The expected returns and valuations of private and public firms, *Journal of Financial Economics*, 120, 41–57.

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

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

Eberly, J., Rebelo, S., and Vincent, N., 2012. What explains the lagged-investment effect?, *Journal of Monetary Economics*, 59, 370–380.

Fama, E.F. and French, K.R., 1992. The cross-section of expected stock returns, *Journal of Finance*, 47, 427–465.

Fama, E.F. and French, K.R., 2008. Dissecting anomalies, *Journal of Finance*, 63, 1653–1678.

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

Fazzari, S.M., Hubbard, R.G., and Petersen, B.C., 1988. Financing constraints and corporate investment, *Brookings Papers on Economic Activity*, 1, 141–195.

Gonçalves, A.S., Xue, C., and Zhang, L., 2020. Aggregation, capital heterogeneity, and the investment capm, *Review of Financial Studies*, 33, 2728–2771.

Guo, H., Kontonikas, A., and Maio, P., 2020. Monetary policy and corporate bond returns, *Review of Asset Pricing Studies*, 10, 441–489.

Haugen, R.A. and Baker, N.L., 1996. Commonality in the determinants of expected stock returns, *Journal of Financial Economics*, 41, 401–439.

Hayashi, F., 1982. Tobin's marginal q and average q: A neoclassical interpretation, *Econometrica*, 50, 213–224.

Hayashi, F., 2000. *Econometrics*, Princeton University Press Princeton, NJ.

Larrain, B. and Yogo, M., 2008. Does firm value move too much to be justified by subsequent changes in cash flow?, *Journal of Financial Economics*, 87, 200–226.

Lettau, M. and Ludvigson, S., 2002. Time-varying risk premia and the cost of capital: An alternative implication of the Q theory of investment, *Journal of Monetary Economics*, 49, 31–66.

Liu, L.X., Whited, T.M., and Zhang, L., 2009. Investment-based expected stock returns, *Journal of Political Economy*, 117, 1105–1139.

Maio, P., 2013. Intertemporal CAPM with conditioning variables, *Management Science*, 59, 122–141.

Maio, P., 2014. Another look at the stock return response to monetary policy actions, *Review of Finance*, 18, 321–371.

Maio, P. and Santa-Clara, P., 2015. Dividend yields, dividend growth, and return predictability in the cross-section of stocks, *Journal of Financial and Quantitative Analysis*, 50, 33–60.

Maio, P. and Xu, D., 2020. Cash-flow or return predictability at long horizons? the case of earnings yield, *Journal of Empirical Finance*, 59, 172–192.

Newey, W.K. and West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, 703–708.

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

Rosenberg, B., Reid, K., and Lanstein, R., 1985. Persuasive evidence of market inefficiency, *Journal of Portfolio Management*, 11, 9–17.

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

Torous, W., Valkanov, R., and Yan, S., 2004. On predicting stock returns with nearly integrated explanatory variables, *Journal of Business*, 77, 937–966.

Valkanov, R., 2003. Long-horizon regressions: Theoretical results and applications, *Journal of Financial Economics*, 68, 201–232.

Vuolteenaho, T., 2002. What drives firm-level stock returns?, *Journal of Finance*, 57, 233–264.

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

Xing, Y., 2008. Interpreting the value effect through the q-theory: An empirical investigation, *Review of Financial Studies*, 21, 1767–1795.

Yashiv, E., 2016. Capital values and job values, *Review of Economic Dynamics*, 19, 190–209.

A Online Appendix: Not for Publication

A.1 Variance Decompositions

In this section, we provide details on the derivations of the VAR-based variance decompositions for q.

A.1.1 Restricted VAR(1)

By multiplying both sides of the present-value relation for q by $q_t - \mathrm{E}(q_t)$, and taking unconditional expectations, we obtain the following variance decomposition for q_t,

$$\mathrm{var}(q_t) \approx (1-\rho)\,\mathrm{Cov}\left(\sum_{h=1}^{H} \rho^{h-1} m_{t+h}, q_t\right) - \mathrm{Cov}\left(\sum_{h=1}^{H} \rho^{h-1} r_{t+h}, q_t\right) + \mathrm{Cov}\left(\rho^H q_{t+H}, q_t\right). \quad \text{(A.1)}$$

By dividing both sides by $\mathrm{var}(q_t)$, we have,

$$1 \approx \beta\left[(1-\rho)\sum_{h=1}^{H} \rho^{h-1} m_{t+h}, q_t\right] - \beta\left(\sum_{h=1}^{H} \rho^{h-1} r_{t+h}, q_t\right) + \beta\left(\rho^H q_{t+H}, q_t\right), \quad \text{(A.2)}$$

where $\beta(y, x)$ denotes the slope from a regression of y on x. This represents the variance decomposition for q based on the direct approach.

By using the property of regression coefficients, $\beta(y+z, x) = \beta(y, x) + \beta(z, x)$, we have:

$$1 \approx (1-\rho)\sum_{h=1}^{H} \rho^{h-1}\beta(m_{t+h}, q_t) - \sum_{h=1}^{H} \rho^{h-1}\beta(r_{t+h}, q_t) + \rho^H \beta(q_{t+H}, q_t). \quad \text{(A.3)}$$

Under the restricted first-order VAR, we have,

$$q_{t+h-1} = \phi^{h-1} q_t + \phi^{h-1}\sum_{l=1}^{h-1} \phi^{-l}(\pi_q + \varepsilon^q_{t+l}), \quad \text{(A.4)}$$

and by combining with the VAR equation for the currency return,

$$r_{t+h} = \pi_r + \lambda_r q_{t+h-1} + \varepsilon^r_{t+h}, \tag{A.5}$$

implies the following equation for r_{t+h}:

$$r_{t+h} = \pi_r + \phi^{h-1}\lambda_r q_t + \phi^{h-1}\lambda_r \sum_{l=1}^{h-1} \phi^{-l}(\pi_q + \varepsilon^q_{t+l}) + \varepsilon^r_{t+h}. \tag{A.6}$$

Since $\mathrm{Cov}(\varepsilon^q_{t+l}, q_t) = 0, l > 0$ and $\mathrm{Cov}(\varepsilon^r_{t+h}, q_t) = 0$, by construction, it follows that

$$\beta(r_{t+h}, q_t) = \phi^{h-1}\lambda_r. \tag{A.7}$$

Similarly, we have,

$$\beta(m_{t+h}, q_t) = \phi^{h-1}\lambda_m. \tag{A.8}$$

On the other hand, given the expanded expression for q_{t+H},

$$q_{t+H} = \phi^H q_t + \phi^H \sum_{l=1}^{H} \phi^{-l}(\pi_q + \varepsilon^q_{t+l}), \tag{A.9}$$

we have

$$\beta(q_{t+H}, q_t) = \phi^H, \tag{A.10}$$

which leads to

$$1 \approx (1-\rho)\sum_{h=1}^{H} \rho^{h-1}\phi^{h-1}\lambda_m - \sum_{h=1}^{H} \rho^{h-1}\phi^{h-1}\lambda_r + \rho^H\phi^H. \tag{A.11}$$

By simplifying the sums above, we obtain the VAR-based variance decomposition associated with

q:

$$
\begin{aligned}
1 &\approx b_m^H - b_r^H + b_q^H, \\
b_m^H &\equiv (1-\rho)\frac{\lambda_m(1-\rho^H\phi^H)}{1-\rho\phi}, \\
b_r^H &\equiv \frac{\lambda_r(1-\rho^H\phi^H)}{1-\rho\phi}, \\
b_q^H &\equiv \rho^H\phi^H.
\end{aligned}
\tag{A.12}
$$

To compute the t-statistics for the predictive coefficients, $\mathbf{b}^H \equiv \left(b_m^H, b_r^H, b_q^H\right)'$, we use the delta method. From the standard errors associated with the VAR slopes, $\mathbf{b} \equiv (\lambda_m, \lambda_r, \phi)'$, we have:

$$
\operatorname{var}\left(\mathbf{b}^H\right) = \frac{\partial \mathbf{b}^H}{\partial \mathbf{b}'}\operatorname{var}\left(\mathbf{b}\right)\frac{\partial \mathbf{b}^H}{\partial \mathbf{b}}. \tag{A.13}
$$

The matrix of derivatives is given by

$$
\frac{\partial \mathbf{b}^H}{\partial \mathbf{b}'} \equiv \begin{bmatrix} \frac{\partial b_m^H}{\partial \lambda_m} & \frac{\partial b_m^H}{\partial \lambda_r} & \frac{\partial b_m^H}{\partial \phi} \\ \frac{\partial b_r^H}{\partial \lambda_m} & \frac{\partial b_r^H}{\partial \lambda_r} & \frac{\partial b_r^H}{\partial \phi} \\ \frac{\partial b_q^H}{\partial \lambda_m} & \frac{\partial b_q^H}{\partial \lambda_r} & \frac{\partial b_q^H}{\partial \phi} \end{bmatrix} = \begin{bmatrix} (1-\rho)\frac{1-\rho^H\phi^H}{1-\rho\phi} & 0 & \frac{-H\lambda_m(1-\rho)\rho^H\phi^{H-1}(1-\rho\phi)+\rho\lambda_m(1-\rho)(1-\rho^H\phi^H)}{(1-\rho\phi)^2} \\ 0 & \frac{1-\rho^H\phi^H}{1-\rho\phi} & \frac{-H\lambda_r\rho^H\phi^{H-1}(1-\rho\phi)+\rho\lambda_r(1-\rho^H\phi^H)}{(1-\rho\phi)^2} \\ 0 & 0 & H\rho^H\phi^{H-1} \end{bmatrix}. \tag{A.14}
$$

A.1.2 Unrestricted VAR(1)

After recursive substitution, the vector of state variables at $t+h$ can be written as,

$$
\mathbf{z}_{t+h} = (\mathbf{I} + \mathbf{A} + ... + \mathbf{A}^{h-1})\boldsymbol{\pi} + \mathbf{A}^h\mathbf{z}_t + \mathbf{A}^{h-1}\boldsymbol{\varepsilon}_{t+1} + ... + \mathbf{A}\boldsymbol{\varepsilon}_{t+h-1} + \boldsymbol{\varepsilon}_{t+h}, \tag{A.15}
$$

or equivalently,

$$
\mathbf{z}_{t+h} = \mathbf{A}^h\mathbf{z}_t + \mathbf{A}^h\sum_{l=1}^{h}\mathbf{A}^{-l}(\boldsymbol{\pi} + \boldsymbol{\varepsilon}_{t+l}). \tag{A.16}
$$

This implies that the regression coefficient of r_{t+h} on q_t is given by

$$\beta(r_{t+h}, q_t) = \frac{\mathrm{Cov}(\mathbf{e}_r'\mathbf{z}_{t+h}, \mathbf{e}_q'\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}_q'\mathbf{z}_t)} = \frac{\mathrm{Cov}(\mathbf{e}_r'\mathbf{A}^h\mathbf{z}_t, \mathbf{e}_q'\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}_q'\mathbf{z}_t)} = \frac{\mathbf{e}_r'\mathbf{A}^h\boldsymbol{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q}, \tag{A.17}$$

where we use the fact that $\mathrm{Cov}(\boldsymbol{\varepsilon}_{t+l}, \mathbf{z}_t) = \mathbf{0}$ for $l > 0$.

By using the result above, it follows that the H-period return slope is given by

$$\begin{aligned}
\sum_{h=1}^{H} \rho^{h-1}\beta(r_{t+h}, q_t) &= \sum_{h=1}^{H} \frac{\mathbf{e}_r'\rho^{h-1}\mathbf{A}^h\boldsymbol{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q} \\
&= \frac{\mathbf{e}_r'}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q}\left(\sum_{h=1}^{H}\rho^{h-1}\mathbf{A}^h\right)\boldsymbol{\Sigma}\mathbf{e}_q \\
&= \frac{\mathbf{e}_r'}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q\rho}\left(\sum_{h=1}^{H}\rho^{h}\mathbf{A}^h\right)\boldsymbol{\Sigma}\mathbf{e}_q \\
&= \frac{\mathbf{e}_r'(\rho\mathbf{A} - \rho^{H+1}\mathbf{A}^{H+1})(\mathbf{I} - \rho\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_q}{\rho\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q} \\
&= \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I} - \rho^H\mathbf{A}^H)(\mathbf{I} - \rho\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q}.
\end{aligned} \tag{A.18}$$

The H-period m slope is defined in an analogous way. The slope associated with future q at $t + H$ is derived as follows:

$$\beta(q_{t+H}, q_t) = \frac{\mathrm{Cov}(\mathbf{e}_q'\mathbf{z}_{t+H}, \mathbf{e}_q'\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}_q'\mathbf{z}_t)} = \frac{\mathrm{Cov}(\mathbf{e}_q'\mathbf{A}^H\mathbf{z}_t, \mathbf{e}_q'\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}_q'\mathbf{z}_t)} = \frac{\mathbf{e}_q'\mathbf{A}^H\boldsymbol{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q}, \tag{A.19}$$

which implies that

$$\rho^H\beta(q_{t+H}, q_t) = \frac{\rho^H\mathbf{e}_q'\mathbf{A}^H\boldsymbol{\Sigma}\mathbf{e}_q}{\mathbf{e}_q'\boldsymbol{\Sigma}\mathbf{e}_q}. \tag{A.20}$$

In the case of the unrestricted VAR(1), the t-ratios associated with the horizon-specific coefficients $\mathbf{b}^H \equiv \left(b_m^H, b_r^H, b_q^H\right)'$, are obtained by using the delta method,

$$\mathrm{var}\left(\mathbf{b}^H\right) = \frac{\partial\mathbf{b}^H}{\partial\mathbf{b}'}\,\mathrm{var}\left(\mathbf{b}\right)\frac{\partial\mathbf{b}^H}{\partial\mathbf{b}}, \tag{A.21}$$

where $\mathbf{b} \equiv (\gamma_m, \theta_m, \lambda_m, \gamma_r, \theta_r, \lambda_r, \gamma_q, \theta_q, \phi)'$. The derivatives are obtained from numerical methods.[25]

A.2 Bootstrap Simulation

The bootstrap simulation associated with the (restricted VAR-based) decomposition for q consists of the following steps.

1. We estimate the first-order restricted VAR,

$$\begin{aligned} r_{t+1} &= \pi_r + \lambda_r q_t + \varepsilon^r_{t+1}, \\ m_{t+1} &= \pi_m + \lambda_m q_t + \varepsilon^m_{t+1}, \\ q_{t+1} &= \pi_q + \phi q_t + \varepsilon^q_{t+1}, \end{aligned}$$

   and save the time-series of residuals (ε^r_{t+1}, ε^m_{t+1}, and ε^q_{t+1}), as well as the estimates of ϕ and ρ.

2. In each replication ($s = 1, ..., 10,000$), we construct pseudo VAR innovations by drawing with replacement from the original VAR residuals:

$$(\varepsilon^{r,s}_{t+1}, \varepsilon^{m,s}_{t+1}, \varepsilon^{q,s}_{t+1})', t = v^s_1, ..., v^s_T,$$

   where the time indices $v^s_1, ..., v^s_T$—which are common for all the three VAR innovations—are created randomly from the original time sequence $1, ..., T$.

3. For each replication, we construct pseudo-samples by imposing the data generating process for r (no-return predictability null),

$$r_{s,t+1} = \rho\varepsilon^{q,s}_{t+1} + (1-\rho)\varepsilon^{m,s}_{t+1},$$

[25] We use the statistical package *Gauss*.

for m (no-profit predictability null),

$$m_{s,t+1} = \varepsilon_{t+1}^{m,s},$$

and for q:

$$q_{s,t+1} = \phi q_{s,t} + \varepsilon_{t+1}^{q,s}.$$

The log Q for the base period (q_1) is picked at random from one of the observations of q_t.

4. In each replication, we use the artificial data to estimate the VAR (1),

$$\begin{aligned} r_{s,t+1} &= \pi_{r,s} + \lambda_{r,s} q_{s,t} + \upsilon_{t+1}^{r,s}, \\ m_{s,t+1} &= \pi_{m,s} + \lambda_{m,s} q_{s,t} + \upsilon_{t+1}^{m,s}, \end{aligned}$$

and estimate the implied long-horizon slopes,

$$\begin{aligned} b_{r,s}^{H} &\equiv \lambda_{r,s} \frac{1-\rho_s^H \phi_s^H}{1-\rho_s \phi_s}, \\ b_{m,s}^{H} &\equiv (1-\rho) \lambda_{m,s} \frac{1-\rho_s^H \phi_s^H}{1-\rho_s \phi_s}, \end{aligned}$$

where ρ_s is the estimate of ρ based on the artificial sample. In result, we have a distribution of the VAR implied slope estimates, $\left\{b_{r,s}^H, b_{m,s}^H\right\}_{s=1}^{10,000}$ for each forecasting horizon H.

5. The p-values associated with the implied VAR slope estimates are calculated as

$$\begin{aligned} p(b_r^H) &= \#\left\{b_{r,s}^H < b_r^H\right\}/10000, \\ p(b_m^H) &= \#\left\{b_{m,s}^H > b_m^H\right\}/10000, \end{aligned}$$

where $\#\left\{b_{m,s}^H > b_m^H\right\}$ denotes the number of simulated slope estimates that are higher than the original slope estimate.