The Marginal Profit-to-Q Ratio: Reassessing the Cash-Flow Channel

Ilan Cooper Paulo Maio Chunyu Yang[1]

This version: April 2021[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]Parts of this study were included in another paper circulating with the title "What drives Q and investment fluctuations?". We thank Ivan Alfaro, Frederico Belo, John Campbell, João Cocco, Winston Dou, Neal Galpin, Howard Kung, Jun Li, Kai Li, Ding Luo, Thien Nguyen, Ali Ozdagli, Stavros Panageas, Riccardo Sabbatucci, Konark Saxena, Amir Yaron, Jin Yu for helpful comments. Any remaining errors are our own.

Abstract

We study a production-based present-value relation that implies that fluctuations in the marginal profit-to-marginal Q ratio (mq) are driven by variations in the expected growth of marginal profits (cash-flow channel), expected investment return changes (discount-rate channel), or both. We find that in contrast to the aggregate dividend-to-price ratio, mq strongly predicts marginal profits growth at both short and long horizons, but not investment returns. mq also predicts (negatively) the growth rates of aggregate earnings, industrial production, and non-farm payrolls. Our findings can guide modeling in which the expected growth rate of marginal profits (at multiple horizons) is time-varying.

Keywords: Tobin's q; Marginal profits-to-q ratio; Investment return; Marginal profit of capital; Variance decomposition; VAR implied predictability; Aggregation bias; Long-horizon regressions; Dividend-to-price ratio; Structural estimation

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

1 Introduction

The neoclassical Q-theory of investment implies that under linearly homogenous technologies, the marginal value of capital, termed marginal Q (Tobin, 1969), is a sufficient statistic to describe investment behavior (Hayashi, 1982). Marginal Q is the present value of all future marginal profits entailed by installing an extra unit of capital. Optimal investment behavior implies that marginal Q is equal to the marginal cost of investment.

In this paper we explore the sources of variation in the logarithm of the ratio of the marginal profit of capital to the marginal value of capital (Tobin's marginal Q), which we intermittently refer to as mq. mq is the production-based analog of the dividend-to-price ratio from the stock market: It plays the same role in the definition of the investment return as the dividend-to-price ratio in relation to the stock return. In particular, mq represents the annual yield driven by the marginal profit of capital (the extra cash-flow due to investment) in relation to the marginal cost of investment (which equals the marginal value of capital under optimal investment behavior). In the same vein, the dividend-to-price ratio represents the annual yield driven by dividends as a fraction of the total cost of investing in the stock (the corresponding price). However, while the stock price is determined by investors in the stock market, it is firms' investment decisions that inform us about their assessment of the marginal value of capital. Identifying the sources of fluctuations in mq is helpful for understanding firms' (scaled) investment behavior and compare it to the behavior of (scaled) stock prices. It is also helpful for understanding the dynamics of profits and marginal profits, an important issue in economic theory, as well as the joint dynamics of investment and profits.

Fluctuations in the marginal profit to marginal Q ratio are potentially caused by both revisions in forecasts of marginal profit growth, of changes in discount rate, or both. Higher discount rates reduce the present value of expected future marginal profits, and hence marginal Q should fall in response. On the other hand, current marginal profits are not affected by discount rate shocks. Consequently, mq could predict positively future investment returns (i.e. discount rates).

The relation between changes in expected marginal profits growth and mq is not as straightforward. Consider a positive and persistent shock to marginal profits. Investment optimally responds positively to the shock, but this in turn could affect negatively the expected future growth rates of marginal profits due to decreasing returns to scale in capital. In this case, marginal Q will forecast negatively the growth rates of future marginal profits. On the other hand, in this case the numerator of mq, namely marginal profits, should also be negatively related to future growth in marginal profits. Thus, both the numerator and denominator of mq rise when the future growth rate of marginal profits are lower. However, it is likely that the denominator rises by less for the following reason. Marginal Q is equal to the marginal cost of investment, which, under (standard) quadratic adjustment costs is a linear function of investment. Due to the convex adjustment costs of investment, investment will respond only gradually to shocks. Overall, it is plausible that mq predicts negatively the growth rate of marginal profits.

As in Liu, Whited, and Zhang (2009) we rely on the q-theory of investment. In the estimation of the model's structural parameters, namely the share of capital in profits 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, 2019). That is, we estimate the parameters using firm-level data to match two moment conditions. First, we match portfolio-level stock returns to a value-weighted average of firm-level investment returns (the investment return moment). Second, we match value-weighted average Tobin's marginal Q in the data to a weighted average model-implied Tobin's Q (the valuation moment).

We derive a dynamic present-value relation for mq, in which mq is negatively correlated with the future cumulative log growth rates in the marginal profit of capital (m), and positively correlated with the future cumulative log investment returns (r). This present-value relation is analogous to the present-value relation associated with the log dividend-to-price price derived in Campbell and Shiller (1988). This relationship gives rise to a variance decomposition for mq at each forecasting horizon, which contains the fractions of the variance of current mq attributed to the predictability of future investment returns, predictability of future marginal profits growth, and the predictability of future mq at a terminal date.

We use two methods to estimate empirically the variance decomposition for mq: A first-order restricted VAR (as in Cochrane, 2008) and an unrestricted VAR (as in Larrain and Yogo, 2008 and Maio and Xu, 2020). The two methods produce qualitatively similar variance decompositions, as all the variation in mq stems from marginal profit growth predictability. Specifically, the long-run marginal profits growth slope estimates are -1.24 and -1.53 under the restricted and unrestricted

VAR approaches, respectively. Hence, in both cases, more than 100% of the variation in mq comes from long-run predictability of marginal profits growth, due to the return slopes having the "wrong" sign. On the other hand, predictability of future mq only plays a relevant role at very short horizons. Therefore, our main finding is that fluctuations in the marginal profit of capital to the marginal value of capital are entirely driven by shocks to expected future marginal profit of capital growth with discount rate shocks assuming no role, as the later source of predictability goes in the wrong direction. This finding stands in stark contrast to the finding in the asset pricing literature showing that the main driving force of the aggregate dividend-to-price ratio is the discount rate channel, that is, the dividend-to-price ratio forecasts stock market returns at long horizons, but does not forecast aggregate dividend growth (see Cochrane, 2008, 2011).

These findings are robust to using median stock and investment returns instead of value-weighted stock 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 decomposition for mq based on long-horizon regressions (direct approach) rather than relying on the first-order VAR (indirect approach). As an additional robustness check, we re-estimate the technology parameters using the methodology employed in Liu, Whited, and Zhang (2009). This methodology does not account for aggregation bias. Reassuringly, the results are very similar for the variance decomposition based on the more robust unrestricted first-order VAR. That is, changes in the expected growth rates of marginal profits account for the bulk of variation in the mq ratio. However, when the variance decomposition is based on the restricted VAR, we obtain an oppositive predictability mix (although with very large standard errors), which suggests a clear misspecification of the restricted first-order VAR when using the alternative definitions of the investment variables.

Since marginal profits represent a dimension of aggregate economic activity, it is relevant to investigate whether mq also helps predicting other dimensions of future business conditions. We show that mq helps to forecast a decline in future economic activity at several forecasting horizons. We employ the growth in aggregate earnings, industrial production, and employees on non-farm payrolls as the alternative measures of broad economic activity. This finding is thus consistent with the predictive ability of mq for the future growth in marginal profits.

The marginal profit of capital (m) and the marginal value of capital (q) are highly correlated (with a correlation coefficient of 0.95). This is consistent with the vast literature that documents the investment cash flow sensitivity. We find that a rise in mq predicts negatively the growth rate of future marginal profits. We also find that m is substantially more volatile than q. We interpret these findings as follows. A positive shock to marginal profit indicates higher marginal profits in the future (because m is somewhat persistent) entailing a rise in q. However, the rise in q is smaller than the rise in m due to the adjustment costs of investment. Hence mq rises. The increase in mq predicts a fall in the expected growth rate of marginal profits for two reasons. First, the profit function exhibits decreasing returns to scale.[1] Second, the productivity shocks that cause a rise in marginal profits are quite quickly mean reverting, implying that a rise in mq is associated with lower expected marginal profit growth.[2]

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). We find the following. First, higher values of the share of capital parameter are associated with lower fractions of marginal profit growth predictability. This occurs because for higher values of the share of capital m becomes more persistent. That is, a shock to m persists longer and hence has a less negative effect on the future growth rate of m. Second, higher levels of the adjustment costs parameter are associated with higher fractions of the marginal profit growth predictability. The reason is that the marginal profit becomes considerably more volatile when this parameter rises.[3] Overall, the comparative analysis indicates that our main result, namely that mq negatively predicts future marginal profit growth, is highly robust to a wide range of plausible parameter values.

In the last part of the paper, we conduct a variance decomposition associated with the ag-

[1] Decreasing returns to scale of the profit function occur for two reasons. First, the marginal product of capital is decreasing. Second, due to any maintenance costs or other fixed costs in production that are proportional to the stock of capital.

[2] Fama and French (2000) find that profitability is mean reverting.

[3] The marginal profit of capital includes the reduction in adjustment costs entailed by having an extra unit of capital. This reduction is larger when the adjustment cost parameter is larger.

gregate dividend yield (dp) corresponding to the firms in our sample. This enables us to put in perspective the main results associated with mq. We find that the discount rate channel drives all the variation in dp from an economic viewpoint: the long-run return slopes are above 100% under both VAR specifications, which implies that the slopes for dividend growth have the wrong sign (positive). These results are consistent with previous evidence on the predictability mix for the dividend-to-price ratio associated with the value-weighted U.S. market portfolio. Therefore, the cash-flow channel seems much more important for the "supply-side" (firms' perspective) of the stock market (i.e., investment return and its components) than the for the usual "demand-side" (investors' perspective) of the market (the stock return and its components).

A natural question to arise from our findings is why our results are so different from the results in the literature concerning the dividend-to-price ratio. Production-based asset pricing models, for example Cochrane (1991), would predict that the sources of variation in the marginal profit-to-marginal Q ratio and the dividend-to-price ratio are the same. One possibility is that dividend smoothing and the changing payout policy by U.S. firms in recent decades (favoring share repurchases over dividends)—see Fama and French (2001), Brav, Graham, Harvey, and Michaely (2005), Leary and Michaely (2011), Chen, Da, and Priestley (2012), among others—might cause the dividend-to-price ratio to be less informative about future equity returns and cash flows than other valuation ratios, like the earnings to price ratio. Indeed, Larrain and Yogo (2008) show that the main driving force of variation in the net payout ratio is long-run cash-flow (rather than return) predictability.[4] This finding is consistent with our findings regarding the marginal profit to marginal Q ratio.

Another possible explanation for the discrepancy between the sources of fluctuations in mq and the dividend-to-price ratio is that managers' assessments of the marginal value of capital differ from those of investors, as in Blanchard, Rhee, and Summers (1993). If managers are less prone to behavioral biases, then corporate investment is more tightly linked with future fundamentals. Corroborating evidence for this conjecture includes the following. Morck, Shleifer, and Vishny

[4] Maio and Xu (2020) find mixed evidence for the aggregate earnings yield. By using the traditional restricted VAR approach, they find a dominant role for the cash-flow channel (earnings growth predictability) in terms of driving the dynamics of the earnings-to-price ratio. This result is in line with the evidence provided in Chen *et al.* (2012). However, by relying on more robust methods—long-horizon regressions or unrestricted VAR—Maio and Xu (2020) obtain a dominant role for the discount-rate channel, which is in line with the abundant existing evidence for the aggregate dividend-to-price ratio.

(1990) find that, although returns can predict investment, this predictive power disappears once they control for fundamentals. Similarly, Blanchard, Rhee, and Summers (1993) find that the stock market does not affect investment, conditional on fundamentals, even though it changes the composition of external finance. Bakke and Whited (2010) find that stock market mispricing does not affect investment, especially that of large firms and firms subject to mispricing. They find that investment does respond to legitimate information in price movements, but only for firms that rely on outside equity financing and whose shares are not mispriced. Thus, it is entirely possible that average investment returns equals average stock returns and that on average the average Q implied by the model equals the average Q in the data (the two moment conditions we estimate), and yet periods of mispricing drive wedges between the respective series. For example, Warusawitharana and Whited (2016) model a first-order autoregressive process for misevaluation, such that the unconditional expectation of the misevaluation is zero. These wedges can lead to different sources of variations in the marginal profit to marginal Q ratio and the dividend-to-price ratio.

In related work, Yashiv (2016) derives a long-run variance decomposition for the log marginal Q-to-marginal profitability of capital, which is approximately equal to (the reciprocal of) our mq variable. By relying on a first-order restricted VAR, he finds a largely dominant role for investment return predictability, a markedly different result from our evidence for mq. However, Yashiv (2016) estimates the structural investment model at the portfolio level, which creates an aggregation bias. Critically, such bias when combined with a first-order restricted VAR (as that employed in Yashiv, 2016), is likely to produce a severe misspecification in the estimated long-run variance decomposition, as suggested by our results. Hence, Yashiv's results appear to suffer from a severe lack of robustness, that is, his long-run predictability mix would not hold by using more robust methods in the estimation of the variance decompositions (unrestricted VAR or long-horizon regressions). Additionally, Yashiv (2016) assumes log utility when estimating his model's parameters. That is, he assumes that consumption growth is the stochastic discount factor. As is well known the consumption-based asset pricing model with log utility cannot successfully describe salient asset pricing facts. In contrast, our estimation does not require any assumption regarding preferences. Another important distinction among the two studies is that Yashiv's model includes labor as a quasi-fixed input in production, whereas our model (as those in Liu, Whited, and Zhang, 2009 and Belo, Xue, and Zhang, 2013, among many others) includes only capital as a costly adjustable input in production.

Our work is also related to the studies of Lettau and Ludvigson (2002) and Cooper *et al.* (2021), which investigate the empirical implications of present-value relations associated with log marginal Q (q). 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 investment growth (controlling for investment-based variables). However, 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 profitability shocks in terms of driving log Q and investment. Moreover, 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. By relying on a similar structural investment model to that used in this study, Cooper *et al.* (2021) estimate a variance decomposition for q and find that the bulk of variation in q is long-run predictability of future investment returns, with predictability of the level of future marginal profits assuming a rather secondary role. This represents exactly the opposite qualitative result that we obtain for the driving forces of mq. In other words, q and mq have clearly distinct dynamics, as indicated by the different driving forces in the respective distinct present-value relations and variance decompositions: In the first case, the discount rate channel is dominant (albeit the cash-flow channel is also important) while in the second case it is all about the cash-flow channel.

Another related study is Chen, Da, and Larrain (2016), who estimate a variance decomposition for unexpected investment growth, which is related to the approach for unexpected stock returns developed in Campbell (1991). They find that unexpected investment growth is largely explained by surprises to current cash flow growth. Apart from many other differences among the two papers, their variance decomposition is about assessing the forecasting role of investment growth for future cash flows and returns, unlike our decomposition, which is centered on the predictive power emanating from mq. Second, their driving forces of investment growth are equity-based variables (e.g., equity cash-flow and discount rate news), as well as the underlying VAR state variables (e.g., stock return, cash-flow growth, and net payout ratio). In contrast, our approach critically relies exclusively on investment-based variables that are estimated from a structural investment model. To the extent that managers' discount rates are different than shareholders' discount rates (as in, for example, Blanchard, Rhee, and Summers, 1993) using expected investment returns rather than stock returns is a more correct approach. Third, importantly, Chen, Da, and Larrain's (2016) present-value relation pertains to total investment which includes investment in liquid assets such as cash and cash equivalent assets, whereas we focus on investment in fixed assets. Most likely, investment in cash does not incur adjustment costs (see, for example, Gonçalves, Xue, and Zhang, 2020) and therefore the cash component should not forecast future returns or cash flows. This implies their findings are difficult to interpret.

Finally, our work also relates to the growing literature that studies stock return predictability from stock price ratios in association with present-value relations. This literature emphasizes the benefits of analyzing jointly the predictability of future stock returns and cash flows. Most of the work is concentrated on the predictability from the dividend yield (Cochrane, 1992, 2008, 2011; Binsbergen and Koijen, 2010; Engsted, Pedersen, and Tanggaard, 2012; Rangvid, Schmeling, and Schrimpf, 2014; Maio and Santa-Clara, 2015; Piatti and Trojani, 2019; Pettenuzo, Sabbatucci, and Timmermann, 2020; Avino, Stancu, and Simen, 2020). On the other hand, several studies compute variance decompositions for other financial ratios such as the earnings yield, book-to-market ratio, or the net payout yield (Cohen, Polk, and Vuolteenaho, 2003; Larrain and Yogo, 2008; Chen, Da, and Priestley, 2012; Maio, 2013a; Maio and Xu, 2020).

The rest of the paper is organized as follows. Section 2 describes the data and the econometric methodology for estimating the structural investment model. We derive a variance decomposition for the log marginal profits-to-Q ratio in Section 3, while Section 4 represents the main empirical analysis conducted in the paper. In Section 5, we provide a sensitivity analysis. Section 6 shows the results for a comparative statics exercise, while Section 7 evaluates the forecasting performance for future economic activity. In Section 8, we estimate a variance decomposition for the dividend-to-price ratio. The paper concludes in Section 9.

2 Model and Econometric Methodology

2.1 Model

This Section describes the model. The purpose of the model is to derive expressions for marginal Q, marginal profits, and investment returns. Subsequently to the econometric estimation of the model's parameters, these variables are used in the derivation of the present value relation. We employ the model in Liu, Whited, and Zhang (2009). The model features a production technology and adjustment costs of investment function. These functions are assumed to be linearly homogenous. Capital is a quasi-fixed factor of production. In addition the firm employs costlessly adjustable inputs. The state variables are the level of capital and an exogenous productivity shock. Given the state variables the firm chooses the levels of the costlessly adjustable inputs in order to maximize operating profits. The firm is a price taker in the goods market. Taking operating profits as given the firm chooses capital investment and debt in order to maximize the value of the firm.

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 $\partial\Pi\left(K_{i,t}, X_{i,t}\right)/\partial K_{i,t} = \alpha Y_{i,t}/K_{i,t}$ where $\alpha > 0$ is the share of capital in profits and Y is sales.

The law of motion of capital is given by $K_{i,t+1} = (1-\delta) K_{i,t} + I_{i,t}$, where $I_{i,t}$ is investment and δ is the capital depreciation rate. Investment entails adjustment costs due to, for example, disruption to production and installation costs. The capital adjustment cost function is assumed to be standard quadratic functional form: $\Phi\left(I_{i,t}, K_{i,t}\right) = a/2\left(I_{i,t}/K_{i,t}\right)^2 K_{i,t}$, where $a > 0$ is the adjustment cost parameter.

We allow firms to finance investment with one-period debt. Taxable profits equal operating profits minus capital depreciation minus interest expenses. The payout of firm i is defined as:

$$D_{i,t} = (1-\tau_t)\left[\Pi\left(K_{i,t}, X_{i,t}\right) - \Phi\left(I_{i,t}, K_{i,t}\right)\right] - I_{i,t} + B_{i,t+1} - R^B_{i,t} B_{i,t} + \tau_t \delta K_{i,t} + \tau_t \left(R^B_{i,t} - 1\right) B_{i,t},$$

where τ_t is the corporate tax rate, $B_{i,t}$ is the amount of debt issued at time $t-1$ and repaid at the beginning of time t, $R^B_{i,t}$ is the gross corporate bond return, $\tau_t \delta K_{i,t}$ is the depreciation tax shield, and $\tau_t\left(R^B_{i,t} - 1\right) B_{i,t}$ is the interest tax shield.

The firm chooses optimal capital investment and debt to maximize the cum-dividend market value of equity:

$$V_{i,t} = \max_{\left\{I_{i,t+s}, K_{i,t+s+1}, B_{i,t+s+1}\right\}_{s=0}^{\infty}} E_t\left[\sum_{s=0}^{\infty} SDF_{t+s} D_{i,t+s}\right],$$

where SDF_{t+1} is the stochastic discount factor from t to $t+1$ and it is correlated with the aggregate component of $X_{i,t+1}$, subject to a transversality condition $\lim_{T\to\infty} E_t\left[SDF_{t+T}B_{i,t+T+1}\right]$.

Proposition 1 in Liu, Whited, and Zhang (2009) states that firms' equity value maximization implies that

$$E_t\left[SDF_{t+1}R_{i,t+1}\right] = 1, \tag{1}$$

in which $R_{i,t+1}$ is the return on investment that is given by

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

The first term in the numerator of the return on investment includes the after-tax marginal operating profit of capital plus the reduction in adjustment costs in period $t+1$ due to the extra unit of capital installed in period t. The second term is the capital depreciation shield, whereas the third term in the numerator is the value of the undepreciated part of the unit of capital installed at time t (where the latter is equal to the marginal cost of investment at time $t+1$). The denominator of the return on investment is the marginal cost of investment, which includes the sum of the purchase price of capital and the marginal adjustment cost of investment. Therefore, the marginal value of an additional unit of capital appears in the numerator, whereas the marginal cost of investment is in the denominator.

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

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

which is the sum of the after-tax marginal product of capital, the reduction in adjustment costs due to the existence of the extra unit of capital, and the depreciation shield.[5]

Optimal investment implies equality between the marginal value of capital and the marginal cost of investment

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

[5] The additional unit of capital reduces tax liabilities as it depreciates. Note that the marginal profit of capital in this definition is different from the marginal operating profit defined earlier.

Given the definitions above we can rewrite the return on investment as

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

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 $w_{i,t}$ and is given by

$$R^{Iw}_{i,t+1} = \frac{R_{i,t+1} - w_{i,t}R^{Ba}_{i,t+1}}{1 - w_{i,t}}. \tag{3}$$

2.2 Aggregate Investment Return

We follow Gonçalves, Xue, and Zhang (2020) and estimate the model's firm-level parameters using firm-level accounting data. That is, we first compute a weighted average of firms' investment returns and test the prediction that on average it is equal to the value weighted stock returns. A question that arises is what weights should be used for the firms' investment returns. In this Section we present a model of aggregation of firms and present the model-implied weights. Our derivation follows Cooper, Maio, and Yang (2021).

Let N be the number of firms in the market. Each firm optimizes by equating the marginal adjustment cost of investment to the marginal value of capital. Firm i makes an investment $I_{i,t}$ at time t, and exiting time t with a level of capital stock $K_{i,t+1}$. The linearly homogenous technologies allow us to apply Hayashi's (1982) result that the marginal value of capital equals the average value of capital. Consequently firm i'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. The aggregate market portfolio's value is thus $\sum_{i=1}^{N} K_{i,t+1}Q_{i,t}$. The aggregate marginal value of capital at time t, denoted by $\bar{Q}_t$, is assumed to price the total capital stock value at the end of time t if multiplied by the total capital stock at the end of time t, as follows:

$$\left(\sum_{i=1}^{N} K_{i,t+1}\right)\bar{Q}_t = \sum_{i=1}^{N} K_{i,t+1}Q_{i,t}. \tag{4}$$

It follows that $\bar{Q}_t$ can be expressed as

$$\bar{Q}_t = \sum_{i=1}^{N} \left(\frac{K_{i,t+1}}{\sum\limits_{j=1}^{N} K_{j,t+1}} \right) Q_{i,t}. \quad (5)$$

Hence, $\bar{Q}_t$ 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 $\bar{Q}_t$. 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.

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

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

The aggregate marginal profit of that extra unit of capital ($\bar{M}_{t+1}$) is a capital stock weighted average of firms' marginal profits of capital. That is,

$$\bar{M}_{t+1} = \sum_{i=1}^{N} \left(\frac{K_{i,t+1}}{\sum\limits_{j=1}^{N} K_{j,t+1}} \right) M_{i,t+1}. \quad (6)$$

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

$$R_{t+1} \equiv \frac{\bar{M}_{t+1} + (1-\delta)\,\bar{Q}_{t+1}}{\bar{Q}_t}. \quad (7)$$

For an investor who holds the economy's stock of capital, an extra unit of capital at time t costs

$\bar{Q}_t$. This extra unit of capital generates profit $\bar{M}_{t+1}$ at time $t+1$ and depreciate to $1-\delta$ unit exiting time $t+1$ with a continuation value of $(1-\delta)\,\bar{Q}_{t+1}$.

2.3 Structural Estimates of Firm-level Parameters

We largely follow Gonçalves, Xue, and Zhang (2020) to estimate the firm-level parameters, the capital share (α) and the adjustment cost parameter (a). These estimates are then used to construct the aggregate investment return and the respective components using Equations (5) to (7).

2.3.1 Moment Conditions

For a given set of testing portfolios (indexed by $j = 1, \cdots, J$), we consider two sets of moment conditions: the investment Euler equation moments and the valuation moments for each testing portfolio j, namely $2J$ moment conditions. The investment Euler equation moment corresponds to testing whether the average stock return equals the average levered investment return. The valuation moment corresponds to testing whether the average Tobin's Q in the data equals the average Q predicted by the model.

$$e_j^r \equiv \mathrm{E}_T[R^S_{j,t+1} - R^{Iw}_{j,t+1}(\alpha, a)] = 0, \tag{8}$$

$$e_j^q \equiv \mathrm{E}_T\left[\widetilde{Q}_{j,t} - \left(1 + (1-\tau_t)a\left(\frac{I_{j,t}}{K_{j,t}}\right)\right)\frac{K_{j,t+1}}{A_{j,t}}\right] = 0, \tag{9}$$

where $E_T(\cdot)$ denotes the sample moment, $R^S_{j,t+1}$ is the portfolio stock return, $R^{Iw}_{j,t+1}$ is the portfolio levered investment return that depends on parameters α and a, $A_{j,t}$ is the total assets, and $\widetilde{Q}_{j,t}$ is the Tobin's Q in the data defined as $\widetilde{Q}_{j,t} \equiv (P_{j,t} + B_{j,t+1})\,/A_{j,t}$.

In the basecase estimation we consider the aggregate market portfolio as the testing portfolio, namely $J = 1$, which leads to two moment conditions. With two moments and two parameters: the capital share and the adjustment cost parameter, the estimation is exactly identified and the two moments fit perfectly. [6]

[6] Our moment conditions are similar as Belo, Xue, and Zhang (2013). Gonçalves, Xue, and Zhang (2020) base their tests on the investment Euler equations only. Imposing only the investment Euler equation for the aggregate market portfolio leads to an unidentified estimation with one moment but two parameters.

2.3.2 Data

We obtain firm-level data from the merged CRSP and COMPUSTAT industrial database. We consider all common stocks on NYSE, Amex, and Nasdaq from 1963 to 2018 excluding firms with primary standard industrial classifications between 4900 and 4999 (utilities) and between 6000 and 6999 (financials). We delete firm-year observations for which total assets, capital stock, or sales are either zero or negative.

We follow Cochrane (1991) and assume a depreciation rate (δ) equal to 0.1.[7] $A_{i,t}$ is total assets (Compustat annual item AT). 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). 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. Tobin's Q ($\widetilde{Q}_{i,t}$) is the ratio of the sum of total debt and market equity to the total assets. Output ($Y_{i,t}$) is sales (Compustat annual item SALE). 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). We measure the tax rate (τ_t) as the statutory corporate income tax (from the Commerce Clearing House, annual publications). We measure the pre-tax corporate bond returns ($R^B_{i,t+1}$) as the ratio of total interest and related expenses (Compustat annual item XINT) scaled by the total debt ($B_{i,t+1}$).[8] 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 $t+1$.

2.3.3 Testing Porfolio and Timing Alignment

We construct testing portfolios at the end of June of year t. In the basecase estimation, the testing portfolio is the aggregate "market" portfolio, namely a portfolio whose value is the value of the aggregate capital stocks and whose return is the aggregate investment returns. Alternatively, we sort all stocks into deciles based on the NYSE breakpoints on Tobin's Q. In the investment Euler equation moment condition, we compute annual value-weighted stock returns from July of year t to

[7] The mean value of the annual depreciation rate is equal to 0.1089 in the our sample of Compustat firms.

[8] Prior studies such as Liu, Whited, and Zhang (2009) use credit rating imputation to measure the pre-tax corporate bond returns. Our approach follows Gonçalves, Xue, and Zhang (2020) that can increase the sample coverage.

June of year $t+1$ for each testing portfolio. To construct the matching annual levered investment returns for each testing portfolio, we compute firm-level levered investment returns using 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 $I_{i,t+1}$). The firm-level levered investment returns are then aggregated to portfolio-level levered investment returns using capital weights. In the valuation moment condition, 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 $t-1$.[9]

Firm-level investment returns, based on firm-level accounting variables, are subject to the issue of outliers. To alleviate the impact of the firm-level outliers, Gonçalves, Xue, and Zhang (2020) winsorize firm-level accounting variables, then construct firm-level investment returns based on winsorized firm-level accounting variables. The firm-level investment returns are then aggregated to the portfolio level as equal- or value-weighted portfolio investment returns. Instead of the winsorization approach, Belo, Gala, Salomao, and Vitorino (2019) use portfolio median return to aggregate firm-level investment returns to the portfolio level since median is known to be robust to outliers. In the basecase estimation, we winsorize firm-level accounting variables at 1-99% level and compute value-weighted portfolio investment returns. As a robustness check, we also apply the median approach without winsorization.

2.3.4 Estimated Parameter Values

In the basecase estimation, we fit the investment Euler equation moment and the valuation moment jointly for the aggregate stock market portfolio using one-step GMM. The estimation is exactly identified and the two moments fit perfectly. Our estimate of the capital share (α) is 0.08, which is similar to the estimates in Gonçalves, Xue, and Zhang (2020).[10] Conventionally, macroeconomists think of the capital share as being approximately 30-40%, based on NIPA data. The 8% number seems rather small on primitive grounds, implying large adjustment costs. However, the

[9] 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 2008 the beginning-of-year capital ($K_{i,2008}$) from the 2007 balance sheet and any flow variable over the year, such as $I_{i,2008}$, from the 2008 income or cash flow statement.

[10] Gonçalves, Xue, and Zhang (2020) report estimates of the capital share varying from 5.04% to 7.53% across different testing portfolios.

estimate of 0.08 pertains to the firm-level, not aggregate level. Indeed, 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 is 0.3. It is plausible that firm-level adjustment costs are considerably larger than aggregate adjustment costs. Indeed as the level of aggregation rises, investment exhibit smoother behavior and is more responsive to shocks (e.g., Doms and Dunne, 1998).

Our estimate of the adjustment cost parameter (a) is 15.18, which is similar to the estimates in Liu, Whited, and Zhang (2009), but higher than the estimates in Gonçalves, Xue, and Zhang (2020) due to the difference in testing portfolios.[11] However, as we show in the comparative statics Section, our results are robust to a wide range of the parameter values, including values of the adjustment cost parameters similar to that in Gonçalves, Xue, and Zhang (2020). This estimate of the adjustment cost parameter leads to an estimated magnitude of the adjustment costs as 11.09% of sales, which is consistent with prior studies.[12]

3 A Present-Value Relation

In this section, we derive a dynamic present-value relation for the log profits-to-Q ratio (mq), which represents the basis for the empirical analysis conducted in the rest of the paper. Such variable is analog to the dividend-to-price ratio embedded in stock returns. Thus, focusing on this ratio facilitates a direct comparison with the extensive asset pricing literature that examines the predictive features of the dividend-to-price ratio.

Our methodology relies on the definition of the realized gross investment return (R) presented in the last section,

$$R_{t+1} = \frac{(1-\delta)Q_{t+1} + M_{t+1}}{Q_t}, \tag{10}$$

where Q represents the aggregate marginal value of capital and M stands for aggregate marginal profits of capital.[13] This definition is analog to the usual definition of the gross stock return with

[11] Liu, Whited, and Zhang (2009) report estimates of the adjustment cost parameter varying from 11.5 to 28.9 across different testing portfolios. Gonçalves, Xue, and Zhang (2020) report such estimates varying from 0.72 to 8.11.

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

[13] To simplify notation, in this section, we ignore the bars on Q and M.

Q playing the role of the stock price and M being the analogue of dividends.

By conducting a log-linear transformation of the investment return in Equation (10), and conducting similar steps as in Campbell and Shiller (1988), we obtain the following difference equation in the log profits-to-Q ratio,

$$mq_t \approx const. + \rho mq_{t+1} + r_{t+1} - \Delta m_{t+1}, \tag{11}$$

where $mq_t \equiv \ln(M_t) - \ln(Q_t) = m_t - q_t$ is the log marginal profits-to-Q ratio at time t; $\Delta m_{t+1} \equiv \ln(M_{t+1}) - \ln(M_t) = m_{t+1} - m_t$ denotes the log growth in marginal profits at time $t+1$; and $r_{t+1} \equiv \ln(R_{t+1})$ represents the log investment return 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 a relevant role in the analysis, which represents a (log-linearization) discount coefficient that depends on the mean of mq (denoted by $\overline{mq}$):

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

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

$$mq_t \approx const. + \sum_{h=1}^{H} \rho^{h-1} r_{t+h} - \sum_{h=1}^{H} \rho^{h-1} \Delta m_{t+h} + \rho^H mq_{t+H}. \tag{12}$$

Under this present-value relation, the current log profits-to-Q ratio is positively correlated with both future multi-period log investment returns (r_{t+h}) and the future profitability ratio at terminal date $t+H$ (mq_{t+H}). On the other hand, mq is negatively correlated with future multi-period log growth in marginal profits (Δm_{t+h}). This dynamic relation is similar to the present-value relation for the log dividend yield developed in Campbell and Shiller (1988): mq plays the role of the log dividend-to-price ratio, the log growth in marginal profits is the analogue of log dividend growth, and the investment return plays the role of the stock return. Similar to the case of the dividend-to-price ratio, in this present-value relation it turns out that both future Δm and r are scaled by one, which implies that both the cash-flow and discount rate effects are on an "equal foot" ex ante.

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

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

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

$$mq_t \approx const. + \sum_{h=1}^{\infty} \rho^{h-1} r_{t+h} - \sum_{h=1}^{\infty} \rho^{h-1} \Delta m_{t+h}. \tag{13}$$

Hence, at very long horizons, only predictability of future investment returns and/or predictability of future marginal profits growth drive the variation in the current mq ratio. Which of these two components matters most in terms of driving the dynamics of mq remains an empirical question, which represents the bulk of our analysis in the following sections.

Table 1 (Panel A) presents the descriptive statistics for the variables in the present-value relation for mq. The log growth in marginal profits is substantially more volatile than the investment return, as indicated by the standard deviations of 0.23 and 0.10, respectively.[14] On the other hand, Δm shows a small negative autocorrelation (-0.08), compared to a small positive serial correlation for r (0.11). The log profits-to-Q ratio is considerably more persistent than the other two variables from the present-value relation, with an autoregressive coefficient close to 0.50. The two components of mq—log marginal profits (m) and log Q (q)—show a slightly higher persistence compared to mq itself, especially in the case of q (autocorrelation of 0.61 versus 0.46). On the other hand, m is substantially more volatile than q, with volatilities of 0.24 and 0.10, respectively. Equation (5) shows that Q is a linear function of the investment-to-capital ratio. The low volatility of q relative to m is consistent with the existence of adjustment costs of investment.

Figure 1 plots the time-series of r, Δm, and mq. All three variables appear to be mean-reverting to a large degree, and hence, stationary. Both r and Δm are clearly procyclical variables, as they tend to decline around most recession periods. This pattern is especially notable during the last two recession periods (2001 and 2007–2009). On the other hand, mq appears as being less correlated with the business cycle. Regarding the critical parameter ρ, which is a function of the average mq ratio, we obtain an estimate of 0.85. This estimate is somewhat smaller than corresponding

[14]The standard deviation of 10% for r is considerably lower than the cross-sectional volatility of the 459 manufacturing industries studied in Cooper and Priestley (2016).

values for the analogue parameter in present-value relations associated with stock returns, which are typically above 0.90 (e.g., Cochrane, 2008, 2011; Maio, 2013b, 2014; Maio and Santa-Clara, 2015).

Panel B of Table 1 shows the correlations among the three variables mentioned above. The correlation between Δm and r is very high, at 0.92. The reason is that shocks to profitability are also shocks to returns, as seen in the definition of the investment return (see Equation (10)). The investment return and mq are also positively correlated, with a correlation of 0.63. This is so also because shocks to marginal profits are also shocks to contemporaneous returns. There is a slightly smaller association between marginal profits growth and mq (correlation of 0.53). Table 1 also displays the correlations with both m and q. It turns out that the main variables of interest (r, Δm, and mq) are somewhat more positively correlated with m than with q. However, there is a very high degree of association between m and q (correlation of 0.95).

4 Variance Decomposition for the Profits-to-Q Ratio

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

4.1 Restricted VAR

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

$$r_{t+1} = \pi_r + \lambda_r mq_t + \varepsilon^r_{t+1}, \quad (14)$$

$$\Delta m_{t+1} = \pi_m + \lambda_m mq_t + \varepsilon^m_{t+1}, \quad (15)$$

$$mq_{t+1} = \pi_{mq} + \phi mq_t + \varepsilon^{mq}_{t+1}, \quad (16)$$

where the εs represent forecast errors. This VAR system is estimated by multiple-equation OLS (see Hayashi, 2000), with heteroskedasticity-robust t-statistics (White, 1980).[15]

[15] Using Newey and West (1987) t-ratios (computed with one lag) leads to similar statistical inference.

By combining the VAR above with the present-value relation in Equation (12), we obtain an approximate identity involving the predictability coefficients associated with mq_t, at each forecasting horizon H:

$$
\begin{aligned}
1 &\approx b_r^H - b_m^H + b_{mq}^H, \\
b_r^H &\equiv \lambda_r \frac{(1-\rho^H \phi^H)}{1-\rho\phi}, \\
b_m^H &\equiv \lambda_m \frac{(1-\rho^H \phi^H)}{1-\rho\phi}, \\
b_{mq}^H &\equiv \rho^H \phi^H.
\end{aligned} \tag{17}
$$

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

In this variance decomposition presented above, 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. The first-order VAR addresses the concern of the lack of statistical power at long horizons associated with long-horizon regressions.[16]

[16] In the next section, we estimate a variance decomposition based on long-horizon regressions.

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

$$\begin{aligned} 1 &\approx b_r^{lr} - b_m^{lr}, \\ b_r^{lr} &\equiv \frac{\lambda_r}{1 - \rho\phi}, \\ b_m^{lr} &\equiv \frac{\lambda_m}{1 - \rho\phi}. \end{aligned} \tag{18}$$

In this decomposition, all the variation in current mq is associated with either return or profits growth long-run predictability. The VAR approach enables one to estimate this long-run decomposition, something that is not feasible under the direct method based on long-horizon regressions. 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 predictability from marginal profits growth,

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

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 VAR-based decomposition associated with mq are displayed in Table 2 (Panel A) and in Figure 2 (Panels A and C). The cash-flow channel plays a clear prominent role at all forecasting horizons. For $H > 1$, more than 100% of the variation in mq is explained by marginal profit growth predictability, with the respective slope estimates being strongly significant (1% level). This means that the return slope estimates assume the wrong sign (negative), being significant (at the 5% level) at most horizons. At an infinite horizon, the estimates of b_r^{lr} and b_m^{lr} are -0.24 and -1.24, respectively. This means that, in economic terms, all the variation in mq is driven by long-run predictability of future marginal profits growth, with long-run return predictability playing no role.

This predictability mix is driven by the VAR(1) mechanics. Specifically, the one-period slope estimate in the equation for Δm has a substantially larger magnitude than the corresponding estimate in the return equation (-0.75 versus -0.14), with the former estimate being significant at the 1% level (while the coefficient estimate associated with future r is only significant at the 10% level). This can also be illustrated by the substantially larger fit in the forecasting regression for Δm (22%) in comparison to the return regression (4%). Such different forecasting performance for future r and Δm is magnified at intermediate and long horizons. We also observe that the sum of the variance decomposition is very close to one at all forecasting horizons. This shows that the present-value relation for mq is quite accurate.

4.2 Bootstrap Simulation

Next, we conduct a bootstrap simulation of the restricted VAR model estimated above. The objective is to account for the 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 predictability of the dividend yield for future stock returns and dividend growth.

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

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho\phi - 1 \\ \phi \end{pmatrix} mq_t + \begin{pmatrix} \varepsilon^{m}_{t+1} - \rho\varepsilon^{mq}_{t+1} \\ \varepsilon^{m}_{t+1} \\ \varepsilon^{mq}_{t+1} \end{pmatrix}, \tag{19}$$

where all the variables in the VAR are demeaned.

To assess predictability of future marginal profits growth, we simulate an alternative VAR specification under the null hypothesis that mq does not forecast future Δm. This means that all the variation in mq comes from predicting future investment returns:

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \begin{pmatrix} 1-\rho\phi \\ 0 \\ \phi \end{pmatrix} mq_t + \begin{pmatrix} \varepsilon^m_{t+1} - \rho\varepsilon^{mq}_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \end{pmatrix}. \tag{20}$$

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 mq for the base period is chosen randomly from the original time-series of mq_t. We compute the pseudo p-values associated with the implied VAR return slopes at each horizon, which represent the fractions of simulated estimates for the return coefficients (from the simulations associated with the first VAR above) that are higher than the estimates found in the data. Similarly, the p-values associated with the marginal profits growth coefficients represent the fractions of pseudo estimates of the profitability coefficients (obtained from the simulations under the second VAR presented above) that are lower than the corresponding sample estimates. The full details of the bootstrap simulation are available in the online appendix.

The results associated with the bootstrap simulation are presented in Figure 3. The simulation confirms the strong statistical significance of the Δm coefficient estimates based on the asymptotic t-ratios, as the corresponding p-values are below 1% at all forecasting horizons. On the other hand, the p-values associated with the return coefficients are substantially above 10% at all horizons. This suggests that the asymptotic p-values for the return slope estimates reported in the previous subsection are likely misleading. In other words, only the predictability coefficient estimates associated with future marginal profits growth present robust statistical significance across inference methods.

4.3 Unrestricted VAR

In this subsection, we estimate an alternative variance decomposition for mq, based on a less restrictive VAR.

Specifically, we consider an unrestricted VAR(1):

$$r_{t+1} = \pi_r + \gamma_r r_t + \theta_r \Delta m_t + \lambda_r mq_t + \varepsilon^r_{t+1}, \tag{21}$$

$$\Delta m_{t+1} = \pi_m + \gamma_m r_t + \theta_m \Delta m_t + \lambda_m mq_t + \varepsilon^m_{t+1}, \tag{22}$$

$$mq_{t+1} = \pi_{mq} + \gamma_{mq} r_t + \theta_{mq} \Delta m_t + \phi mq_t + \varepsilon^{mq}_{t+1}. \tag{23}$$

This specification accounts for relevant predictability of lagged returns and profits growth on all three variables in the VAR, something that the benchmark VAR misses. Indeed, Maio and Xu (2020) show that the restricted VAR(1) can be misspecified, which 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,

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_m \\ \pi_{mq} \end{pmatrix} + \begin{pmatrix} \gamma_r & \theta_r & \lambda_r \\ \gamma_m & \theta_m & \lambda_m \\ \gamma_{mq} & \theta_{mq} & \phi \end{pmatrix} \begin{pmatrix} r_t \\ \Delta m_t \\ mq_t \end{pmatrix} + \begin{pmatrix} \varepsilon^r_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \end{pmatrix}, \tag{24}$$

or equivalently, the VAR can be defined as

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{25}$$

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 $\mathbf{e}_{mq} \equiv (0,0,1)'$, 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 heteroskedasticity-robust t-ratios. The t-ratios of the implied horizon-specific coefficients are produced from the delta method. The covariance matrix of the state variables is given by $\mathbf{\Sigma} \equiv \text{Cov}(\mathbf{z}_t, \mathbf{z}_t')$. Given these definitions, and following Larrain and Yogo (2008) and Maio and Xu (2020), we derive the following variance decomposition for mq at each horizon H,

$$\begin{aligned}
1 &\approx b_r^H - b_m^H + b_{mq}^H, \\
b_r^H &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I} - \rho^H\mathbf{A}^H)(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}, \\
b_m^H &\equiv \frac{\mathbf{e}_m'\mathbf{A}(\mathbf{I} - \rho^H\mathbf{A}^H)(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}, \\
b_{mq}^H &\equiv \frac{\rho^H\mathbf{e}_{mq}'\mathbf{A}^H\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}},
\end{aligned} \tag{26}$$

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. This means that 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 mq is given by

$$\begin{aligned}
1 &\approx b_r^{lr} - b_m^{lr}, \\
b_r^{lr} &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}, \\
b_m^{lr} &\equiv \frac{\mathbf{e}_m'\mathbf{A}(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}.
\end{aligned} \tag{27}$$

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

The estimation results for the unrestricted VAR above are presented in Table 3 (Panel A). We can see that the slopes associated with lagged r and Δm are largely insignificant in all three VAR equations. We also note that the estimates for both λ_r and λ_m increase in magnitude relative to the corresponding estimates in the restricted VAR. On the other hand, the estimate for ϕ is lower than the corresponding estimate in the benchmark VAR, with no significance at the 10% level (although marginally so, with a t-ratio of 1.60). Nonetheless, because the estimates for the γs and θs have relatively small magnitudes, we expect the resulting variance decomposition for mq to be relatively similar to the corresponding decomposition under the benchmark VAR.

The results for the variance decomposition based on the unrestricted VAR(1) are presented in Figure 2 (Panels B and D) and Table 3 (Panel A). The results are indeed qualitatively similar to those in the decomposition corresponding to the benchmark VAR. Specifically, the shares of marginal profits growth predictability are above 100%, and strongly statistically significant (1% level), at all forecasting horizons beyond one year. At $H = 1$, the share of Δm predictability is 0.75, and such estimate is also significant at the 1% level. On the other hand, the return slope estimates assume the wrong sign (negative) at all horizons. The long-run (infinite horizon) return and marginal profit growth coefficient estimates are -0.51 and -1.53, respectively. This suggests an even more extreme amount of long-run profits predictability in comparison to that obtained under the baseline VAR. Since the return slope estimates have the wrong sign, we have the same pattern as in the restricted VAR case: From an economic viewpoint, all the variation in mq stems from predictability of future profits growth at nearly all forecasting horizons. The unique exception to this pattern is at $H = 1$, in which predictability of future mq also plays a relevant role, with a coefficient estimate of 0.40 (which is significant at the 1% level). Yet, this last predictability channel converges to zero rather quickly for horizons beyond one year.

All in all, the punch line of the results associated with this alternative variance decomposition is that predictability of future profitability growth (the cash-flow channel) represents the major driving force of variation in mq by a very large margin.

4.4 Inspecting the Mechanism

In this subsection, we conduct a further decomposition of the variance decomposition associated with mq. The objective is to evaluate which of the components of the profits-to-Q ratio, m or q, is driving the predictability patterns discussed above. This analysis appears relevant as those two variables have a different economic meaning: q is the log of the marginal value of capital and m is the log of the marginal profitability of capital, which represents the short-term benefit of an extra unit of capital. Thus, mq represents the ratio of the short-term benefit of capital to the value of capital. Intuitively, because investment is forward looking, it should respond only to persistent shocks to m.

To keep focus, the analysis of this subsection is based on the long-run variance decomposition of mq associated with the restricted VAR. By using the linearity of the covariance operator, we can decompose the one-period return slope for mq as follows,

$$
\begin{aligned}
\lambda_r &\equiv \frac{\operatorname{Cov}(r_{t+1}, mq_t)}{\operatorname{var}(mq_t)} \\
&= \frac{\operatorname{Cov}(r_{t+1}, m_t)}{\operatorname{var}(mq_t)} - \frac{\operatorname{Cov}(r_{t+1}, q_t)}{\operatorname{var}(mq_t)} \\
&= \frac{\operatorname{Cov}(r_{t+1}, m_t)}{\operatorname{var}(m_t)} \frac{\operatorname{var}(m_t)}{\operatorname{var}(mq_t)} - \frac{\operatorname{Cov}(r_{t+1}, q_t)}{\operatorname{var}(q_t)} \frac{\operatorname{var}(q_t)}{\operatorname{var}(mq_t)} \\
&= \mu_r \omega_m - \eta_r \omega_q, \qquad (28)
\end{aligned}
$$

where $\omega_m \equiv \operatorname{var}(m_t)/\operatorname{var}(mq_t)$ and $\omega_q \equiv \operatorname{var}(q_t)/\operatorname{var}(mq_t)$ represent the variance weights associated with m and q, respectively.

Similarly, the one-period marginal profits growth coefficient is decomposed as follows:

$$
\begin{aligned}
\lambda_m &= \frac{\operatorname{Cov}(\Delta m_{t+1}, mq_t)}{\operatorname{var}(mq_t)} \\
&= \frac{\operatorname{Cov}(\Delta m_{t+1}, m_t)}{\operatorname{var}(m_t)} \frac{\operatorname{var}(m_t)}{\operatorname{var}(mq_t)} - \frac{\operatorname{Cov}(\Delta m_{t+1}, q_t)}{\operatorname{var}(q_t)} \frac{\operatorname{var}(q_t)}{\operatorname{var}(mq_t)} \\
&= \mu_m \omega_m - \eta_m \omega_q. \qquad (29)
\end{aligned}
$$

The coefficient estimates associated with m_t are obtained from the following auxiliary regressions,

$$
\begin{aligned}
r_{t+1} &= \tau_r + \mu_r m_t + \nu^r_{t+1}, \qquad (30) \\
\Delta m_{t+1} &= \tau_m + \mu_m m_t + \nu^m_{t+1}, \qquad (31)
\end{aligned}
$$

while the slopes corresponding to q_t are retrieved from:

$$
\begin{aligned}
r_{t+1} &= \varpi_r + \eta_r q_t + \varsigma^r_{t+1}, \qquad (32) \\
\Delta m_{t+1} &= \varpi_m + \eta_m q_t + \varsigma^m_{t+1}. \qquad (33)
\end{aligned}
$$

Equipped with these definitions, it is straightforward to decompose the implied long-run (infinite horizon) return and marginal profits growth coefficients associated with mq:

$$b_r^{lr} = \frac{\lambda_r}{1-\rho\phi} = \frac{\mu_r}{1-\rho\phi}\omega_m - \frac{\eta_r}{1-\rho\phi}\omega_q, \quad (34)$$

$$b_m^{lr} = \frac{\lambda_m}{1-\rho\phi} = \frac{\mu_m}{1-\rho\phi}\omega_m - \frac{\eta_m}{1-\rho\phi}\omega_q. \quad (35)$$

Hence, the original long-run mq slopes are decomposed into a difference of weighted long-run m and q coefficients, in which the weights are driven by the volatility shares over mq.

The auxiliary regressions when the predictor is m yield the following results,

$$r_{t+1} = -0.03(-0.64) - 0.11(-1.97)m_t, R^2 = 0.06,$$

$$\Delta m_{t+1} = -0.38(-2.79) - 0.46(-3.17)m_t, R^2 = 0.22,$$

where the numbers in parentheses denote the t-ratios.

When q is the predictor, we get the following results:

$$r_{t+1} = 0.35(2.62) - 0.29(-2.13)q_t, R^2 = 0.08,$$

$$\Delta m_{t+1} = 1.01(2.83) - 0.97(-2.70)q_t, R^2 = 0.18.$$

These results show that the four slope estimates are significantly negative. Yet, the coefficient estimates associated with lagged q have larger magnitudes than the corresponding estimates associated with lagged m. Given these estimates and the estimated variance weights, $\omega_m = 2.69$ and $\omega_q = 0.49$, the long-run return and marginal profits growth coefficients associated with mq are decomposed as follows:

$$b_r^{lr} = \frac{\mu_r}{1-\rho\phi}\omega_m - \frac{\eta_r}{1-\rho\phi}\omega_q = -0.47 - (-0.23) = -0.24$$

$$b_m^{lr} = \frac{\mu_m}{1-\rho\phi}\omega_m - \frac{\eta_m}{1-\rho\phi}\omega_q = -2.03 - (-0.79) = -1.24.$$

Hence, these results show that the long-run predictability mix associated with mq is driven by m rather than by q. This stems from the fact that m is substantially more volatile than q, despite the fact that the two variables are highly correlated (as shown in Table 1). The resulting higher volatility weights ($\omega_m > \omega_q$) more than outweigh the larger magnitudes of the slopes associated with lagged q discussed above. In more detail, the estimate of η_r (-0.29) is not negative enough in order to dominate the negative estimate of μ_r (-0.11) and produce an estimate of λ_r with the correct sign (positive). On the other hand, the estimate of η_m (-0.97) is not negative enough in order to attenuate the estimate of μ_m (-0.46) so that the resulting estimate of λ_m is less negative (i.e., lower than one in magnitude), which corresponds to a lower share of profits growth predictability in driving mq. Therefore, the relatively high volatility of m has a large impact on the predictability pattern associated with mq.

We interpret the results in this subsection as being consistent with economic theory. The stochastic process of m is largely driven by productivity shocks. It is a persistent but mean-reverting process. Investment responds immediately to changes in m. m is substantially more volatile than q due to adjustment costs of investment which render q relatively smooth. The overall results show that investment predicts the future level of the marginal profitability of capital and consequently investment predicts negatively the growth rate of the marginal profitability of capital.

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, we concentrate the discussion on the long-run variance decomposition for mq.

5.1 Alternative Investment Series

We conduct the variance decomposition for mq 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 predictability mixes are very similar to those estimated with the benchmark data. Specifically, the long-run return and marginal profits growth slope estimates are −0.27 and −1.27, respectively, in the case of the restricted VAR (Table 2, Panel B). When the estimation is based on the unrestricted VAR (Table 3, Panel B), the corresponding estimates are −0.55 and −1.56, respectively, which are also quite close to the estimates obtained with the benchmark series.

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 (2019) 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%). The predictability patterns are very similar to those estimated in the benchmark case. In particular, the long-run return and profits growth coefficient estimates are −0.25 and −1.25, respectively, in the case of the restricted VAR (Table 2, Panel C). The corresponding estimates in the case of the unrestricted VAR are −0.51 and −1.52, respectively (Table 3, Panel C).

5.2 Higher-Order VAR

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

The restricted VAR(2) specification is given by

$$r_{t+1} = \pi_r + \lambda_{r1} mq_t + \lambda_{r2} mq_{t-1} + \varepsilon^r_{t+1}, \quad (36)$$

$$\Delta m_{t+1} = \pi_m + \lambda_{m1} mq_t + \lambda_{m2} mq_{t-1} + \varepsilon^m_{t+1}, \quad (37)$$

$$mq_{t+1} = \pi_{mq} + \phi_1 mq_t + \phi_2 mq_{t-1} + \varepsilon^{mq}_{t+1}. \quad (38)$$

The VAR(2) is estimated by multiple-equation OLS, with Newey and West (1987) t-statistics (computed with one lag).[17]

[17] Using two lags provides similar statistical significance.

We can write the VAR above as a VAR(1) in the companion form:

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \\ r_t \\ \Delta m_t \\ mq_t \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_m \\ \pi_{mq} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \lambda_{r1} & 0 & 0 & \lambda_{r2} \\ 0 & 0 & \lambda_{m1} & 0 & 0 & \lambda_{m2} \\ 0 & 0 & \phi_1 & 0 & 0 & \phi_2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r_t \\ \Delta m_t \\ mq_t \\ r_{t-1} \\ \Delta m_{t-1} \\ mq_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon^r_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (39)$$

or equivalently,

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}. \quad (40)$$

The indicator vectors are defined as follows,

$$\mathbf{e}_r \equiv (1,0,0,0,0,0)', \quad (41)$$

$$\mathbf{e}_m \equiv (0,1,0,0,0,0)', \quad (42)$$

$$\mathbf{e}_{mq} \equiv (0,0,1,0,0,0)', \quad (43)$$

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

$$\boldsymbol{\Sigma} = \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}_t') = \begin{pmatrix} \mathrm{var}(r_t) & \mathrm{Cov}(r_t, \Delta m_t) & \mathrm{Cov}(r_t, mq_t) & 0 & 0 & 0 \\ \mathrm{Cov}(r_t, \Delta m_t) & \mathrm{var}(\Delta m_t) & \mathrm{Cov}(\Delta m_t, mq_t) & 0 & 0 & 0 \\ \mathrm{Cov}(r_t, mq_t) & \mathrm{Cov}(\Delta m_t, mq_t) & \mathrm{var}(mq_t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (44)$$

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(1) case, the t-ratios for the long-horizon coefficients are obtained by applying the delta method.

The estimation results presented in Table 4 show that mq_{t-1} helps to forecast r_{t+1}, as the corresponding coefficient is significantly negative. On the other hand, the slope estimates associated with mq_{t-1} are clearly insignificant in the regressions associated with Δm_{t+1} and mq_{t+1}. Hence, the estimate of ϕ_1 (0.48) is very similar to the corresponding estimate in the first-order VAR. Moreover, the R^2 estimates in the marginal profits growth and mq regressions are quite close to the corresponding values under the VAR(1) case.

The VAR(2) estimation results suggest that the variance decomposition for mq should be qualitatively similar to that estimated under the restricted VAR(1). The results largely confirm that proposition. Specifically, the long-run return and marginal profits growth slopes are -0.38 and -1.39, respectively. This indicates that the share of marginal profits growth predictability is somewhat higher than estimated under the restricted VAR(1), bet less extreme than that associated with the unrestricted VAR(1).

5.3 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 predictability mix in the restricted-VAR case is almost symmetric to the corresponding mix obtained with the benchmark data, as shown in Table 2, Panel D. Specifically, the long-run return and marginal profit growth slope estimates are 0.74 and -0.26, respectively, which indicates that return predictability drives most of the variation in mq. However, the corresponding standard errors are large as none of the long-run coefficient estimates is statistically significant at the 10% level.

When the slope estimates are based on the unrestricted VAR (Table 3, Panel D), the predictability mix looks almost opposite to that estimated under the restricted VAR and resembles very much the benchmark results discussed in the previous section. That is, what drives the variation in mq is the predictability of future profitability growth, with return predictability assuming a secondary role. Specifically, the long-run estimates of the r and Δm slopes are 0.16 and -0.83, respectively. The marginal profit growth coefficient estimate is not significant at the 10% level. However, by employing single-sided p-values, there is significance (t-ratio of -1.51). Using single-sided p-values is a less conservative approach than the usual double-sided p-values, yet, it is appropriate for the variance decomposition for mq, as the signs of the predictive coefficients are constrained by theory (i.e., the present-value relation).

The results of this subsection show that, when relying on investment series produced by the Liu–Whited–Zhang approach, the restricted and unrestricted VAR methods produce quite different decompositions for mq. This stems from an important mis-specification of the restricted VAR and is consistent with the evidence in Maio and Xu (2020) showing that such method can provide quite misleading long-run variance decompositions for aggregate stock market valuation ratios.[18] In related work, Yashiv (2016) derives a long-run variance decomposition for the log Q-to-capital productivity ratio, which is approximately equal to the symmetric of mq.[19] By relying on a first-order restricted VAR, he finds a largely dominant role for investment return predictability, which is in line with our evidence above for mq under the restricted VAR case.[20] However, following the discussion above, the findings obtained in Yashiv (2016) are likely not robust by employing superior statistical processes to produce variance decompositions, such as the unrestricted first-order VAR approach (or the direct method discussed below). Therefore, the combined evidence from this subsection suggests that the log profits-to-Q ratio, or its symmetric, moves mainly in reaction to cash-flow (profitability of capital) shocks rather than investment return shocks.

5.4 Direct Approach

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

[18] Maio and Xu (2020) provide evidence that a valid alternative to the unrestricted VAR (estimated by OLS) is to use the restricted VAR model, but estimated with the Projection Minimum Distance method of Jordà and Kozicki (2011) (rather than by OLS).

[19] Yashiv's investment return equation and associated present-value relation is different than ours. This implies that his long-run decomposition is also different than ours, involving four predictability sources, compared to two sources in our long-run decomposition for mq. For example, there is a term related to the predictability of the depreciation rate of capital, which is absent from our decomposition since we assume that the depreciation rate is constant over time.

[20] Yashiv (2016) relies on quarterly data and uses a different sample than ours. He estimates a share of long-run return predictability above 100% (115%), as reported in Table 5 of his paper. However, we cannot judge the statistical significance of these implied long-run coefficient estimates, as he does not report standard errors (or t-ratios) for those estimates.

Following Cochrane (2008, 2011) and Maio and Santa-Clara (2015), we estimate weighted long-horizon regressions of future cumulative log investment returns, cumulative log growth in marginal profits, and the log profits-to-Q ratio on the current profitability ratio:

$$\sum_{h=1}^{H} \rho^{h-1} r_{t+h} = a_r^H + b_r^H mq_t + \varepsilon_{t+H}^r, \tag{45}$$

$$\sum_{h=1}^{H} \rho^{h-1} \Delta m_{t+h} = a_m^H + b_m^H mq_t + \varepsilon_{t+H}^m, \tag{46}$$

$$\rho^H mq_{t+H} = a_{mq}^H + b_{mq}^H mq_t + \varepsilon_{t+H}^{mq}. \tag{47}$$

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), which 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 mq in Equation (12) with the predictive regressions presented above, we obtain an approximate identity involving the predictability coefficients associated with mq_t, at each forecasting horizon H:

$$1 \approx b_r^H - b_m^H + b_{mq}^H. \tag{48}$$

There is a key advantage related with the direct method. If the first-order VAR does not fully capture the dynamics of the data generating process for r, mq, and Δm, it follows that the resulting variance decomposition will be a poor approximation of the true decomposition for mq, 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). However, one possible problem associated with 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.

The horizon-specific direct variance decomposition for mq are displayed in Figure 4. We can see that at all horizons the driving source of variation in mq is predictability of marginal profits growth, with the respective slopes being significant (at the 5% level) in all cases. In fact, apart from the one-year horizon, it turns out that the slopes associated with Δm become larger than one in magnitude. This indicates that the predictability of future marginal profits growth accounts for more than 100% of the variation in current mq at nearly all horizons. The reason for such pattern is that the investment return slopes have the wrong sign (negative) at all forecasting horizons, although these estimates are not statistically significant (at the 5% level) in most cases. At long horizons ($H = 20$), the estimates of b_r^H, b_m^H, and b_{mq}^H are -0.36, -1.36, and 0.02, respectively, which are in line with the corresponding long-run estimates obtained under the restricted VAR(1). The direct variance decomposition is slightly less accurate than under the VAR approach. However, apart from very long horizons, the deviations from one are less than 2% at all horizons.

Overall, the predictability mix associated with mq obtained under the direct method is qualitatively similar to those obtained under the two indirect methods discussed in the previous section.

6 Comparative Statics

In this section, we conduct a comparative statics exercise. This exercise is important and constitutes a robustness check because of the parameter estimation error.

Specifically, we estimate a range of variance decompositions associated with mq for a set of artificial series of the key investment variables, r, Δm, and mq. 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 mq, that we obtained in the previous sections, holds for a reasonable range of those two underlying parameters. This is the more pertinent as the results of the last section show that the structural estimation approach employed in Liu, Whited, and Zhang (2009) yields quite different parameter estimates than the benchmark estimation method used in the paper. This implies, for example, that the variance decompositions (in particular, those based on the restricted VAR) can potentially differ substantially by using these two different sets of the investment variables.

The simulation results are presented in Table 5. 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

[21] Notice that we use the same series of investment and sales as in the data. However, different parameter values will yield different series of r, Δm, and mq.

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. To preserve space and keep the focus, we report only the long-run (infinite horizon) variance decomposition for mq. We use both the restricted and unrestricted VAR methods to produce the decompositions.

The first key pattern that emerges from Table 5 is that the share of marginal profits growth predictability declines with the magnitude of α. Specifically, at high values of this parameter, that is $\alpha = 0.50, 0.70$, it turns out that the cash-flow shares are below 50% (in magnitude) in most cases. The main exception occurs for the pair $\alpha = 0.50, a = 27.38$. However, even in that case, the Δm slope estimates obtained from both the restricted and unrestricted VARs are insignificant at the 10% level (based on two-sided p-values). At the other end of the spectrum, for $\alpha = 0.05, 0.15$, the estimates of b_m^{lr} are above (or fairly close to) one (in magnitude), and strongly significant (1% level), in most cases. The sole exception is when $\alpha = 0.15, a = 1.37$, in which case the long-run marginal profits coefficient estimate is -0.42 under both the restricted and unrestricted VAR approaches, which is significant at the 5% (unrestricted VAR) or 10% (restricted VAR) level. Untabulated results show that ρ declines with α, that is, we obtain estimates of ρ above 0.85 for $\alpha = 0.05$ and estimates below 0.70 for $\alpha = 0.70$. This means that higher shares of the cash-flow channel are associated with higher values of the log-linearization parameter.

High values of α imply larger returns to scale in production and therefore a larger marginal value of capital. Higher returns to scale will reduce the impact of a given change in investment on the future values of marginal profits. Therefore the mq ratio will have a lower correlation with the future growth rate of m. Moreover, higher values of the share of capital in profits renders the marginal profits more persistent. Therefore, a rise in mq stemming from a shock to m has less predictive power for the future growth rates of m.

The second key result from Table 5 is that, for a given value of α, the shares of cash-flow predictability tend to increase with the magnitude of the adjustment cost parameter (a). The likely reason for this patten is that marginal profits becomes substantially more volatile for higher values of a. However, importantly, even for small values of a our main result holds. For example, when α is 0.05 and a is in the range of 1.37 to 6.85 (parameter values similar to those in Gonçalves *et al.*, 2020) cash-flow predictability is strong whereas return predictability is with the "wrong" sign. We also observe across the board that the weights of Δm predictability are almost always larger under the unrestricted VAR than under the restricted VAR method. This result is consistent with the evidence obtained in Section 4 based in our sample estimates.

All in all, the results of this section show that the dominant role of marginal profits growth predictability in terms of driving variation in mq 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 investment return predictability under less plausible values for those structural parameters.

7 Forecasting Economic Activity

The results in the previous sections show that mq has strong forecasting power for the future growth in marginal profits. Since marginal profits represent a dimension of aggregate economic activity, it is relevant to investigate whether mq also helps predicting other dimensions of future business conditions. We pursue such analysis in this section.[22]

Specifically, we run bivariate long-horizon regressions of future economic activity on the current mq (e.g., Keim and Stambaugh, 1986; Fama and French, 1988, 1989),

$$\sum_{h=1}^{H} f_{t+h} = \varpi_H + \psi_H f_t + \eta_H mq_t + u_{t+H}, \tag{49}$$

where f denotes the log growth in a variable measuring economic activity. In these regressions, we assess the forecasting role of mq by controlling for current economic activity (f_t), which represents a standard practice in the related literature (e.g., Stock and Watson, 2003). We use forecasting horizons of 1, 5, 10, 15, and 20 years ahead. To assess the statistical significance of the coefficient estimates, we use Newey and West (1987) t-ratios (with $H-1$ lags).

We employ three variables as proxies of broad economic activity: S&P earnings (Δe), industrial production total index (Δip), and employees on non-farm payrolls (Δnfp). The data on the last two variables are retrieved from St. Louis Fed data library. The data on aggregate earnings are

[22] There is a voluminous literature on forecasting future economic activity (see the survey in Stock and Watson, 2003). A list of recent studies includes Gilchrist and Zakrajsek (2012), Hong and Yogo (2012), Faust *et al.* (2013), Choi *et al.* (2017), Mian *et al.* (2017), Bakshi *et al.* (2019), Cooper and Maio (2019), and Faccini *et al.* (2019).

obtained from Amit Goyal's webpage.

The results for the predictive regressions associated with future economic activity are reported in Table 6. We can see that mq forecasts a decline in future aggregate business conditions with the respective slope estimates being strongly significant (5% or 1% level) in most cases. This is consistent with the predictive performance for Δm documented in the previous sections. The few exceptions for that pattern occur at the longest forecasting horizon (20 years) when the economic activity proxies are Δip and Δnfp. Moreover, such forecasting power is economically significant, as the R^2 estimates are around or above 20% for most horizons. We also observe that mq tends to have stronger forecasting power for future earnings growth in comparison to the other two measures of economic activity.

Overall, the evidence from this section is that mq helps to forecast a decline in future economic activity at several horizons, which is in line with the evidence concerning future marginal profits growth documented in the previous sections.

8 Dividend-to-Price Ratio

In this section, we conduct a variance decomposition associated with the aggregate dividend yield. This enables us to put in perspective the results associated with mq, which were discussed in the previous sections. As shown in Section 2, the investment return is related to the stock return. The first variable (and its components) is associated with the supply side of the stock market, while the second variable (and its components) is related to the demand side of the equity market. Hence, it makes sense to compare the variance decompositions associated with these two connected variables. To save space, in this section, we only report the long-run decompositions for the dividend-to-price ratio.

Following Campbell and Shiller (1988), the dynamic accounting identity for the log dividend-to-price ratio (dp) can be represented as

$$dp_t \approx const. + \sum_{h=1}^{\infty} \rho_s^{h-1} r_{t+h}^s - \sum_{h=1}^{\infty} \rho_s^{h-1} \Delta d_{t+h}, \tag{50}$$

where $const.$ is a constant term that is irrelevant for the forthcoming analysis; $r_{t+h}^s \equiv \ln(R_{t+h}^s)$

denotes the log stock return between $t+h-1$ and $t+h$; and $\Delta d_{t+h} \equiv \ln(D_{t+h}/D_{t+h-1})$ represents the log dividend growth between $t+h-1$ and $t+h$. ρ_s is a (log-linearization) discount coefficient that depends on the mean of dp. This present-value relation is similar to that derived for mq: r^s plays the role of r, Δd is analog to Δm, and dp plays the role of mq. ρ_s is analog to ρ.

Following Cochrane (2008), the long-run variance decomposition associated with the dividend yield based on the restricted VAR is given by

$$
\begin{aligned}
1 &\approx b_r^{lr} - b_d^{lr}, \\
b_r^{lr} &\equiv \frac{\lambda_r}{1-\rho_s\phi}, \\
b_d^{lr} &\equiv \frac{\lambda_d}{1-\rho_s\phi},
\end{aligned} \tag{51}
$$

in which the one-year slopes are obtained from the following restricted first-order VAR:

$$r_{t+1}^s = \pi_r + \lambda_r dp_t + \varepsilon_{t+1}^r, \tag{52}$$

$$\Delta d_{t+1} = \pi_d + \lambda_d dp_t + \varepsilon_{t+1}^d, \tag{53}$$

$$dp_{t+1} = \pi_{dp} + \phi dp_t + \varepsilon_{t+1}^{dp}. \tag{54}$$

The unrestricted VAR(1) associated with dp can be presented in matrix form,

$$
\begin{pmatrix} r_{t+1}^s \\ \Delta d_{t+1} \\ dp_{t+1} \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_d \\ \pi_{dp} \end{pmatrix} + \begin{pmatrix} \gamma_r & \theta_r & \lambda_r \\ \gamma_d & \theta_d & \lambda_d \\ \gamma_{dp} & \theta_{dp} & \phi \end{pmatrix} \begin{pmatrix} r_t^s \\ \Delta d_t \\ dp_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+1}^r \\ \varepsilon_{t+1}^d \\ \varepsilon_{t+1}^{dp} \end{pmatrix}, \tag{55}
$$

which is equivalent to

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{56}$$

with $\mathbf{z}_t \equiv (r_t^s, \Delta d_t, dp_t)'$.

Given the indicator vectors, $\mathbf{e}_r \equiv (1,0,0)'$, $\mathbf{e}_d \equiv (0,1,0)'$, and $\mathbf{e}_{dp} \equiv (0,0,1)'$, and the covariance matrix of the state variables, $\mathbf{\Sigma} \equiv \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}_t')$, we define the long-run variance decomposition for

dp associated with the unrestricted VAR:

$$\begin{aligned} 1 &\approx b_r^{lr} - b_d^{lr}, \\ b_r^{lr} &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I} - \rho_s\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{dp}}{\mathbf{e}_{dp}'\boldsymbol{\Sigma}\mathbf{e}_{dp}}, \\ b_d^{lr} &\equiv \frac{\mathbf{e}_d'\mathbf{A}(\mathbf{I} - \rho_s\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{dp}}{\mathbf{e}_{dp}'\boldsymbol{\Sigma}\mathbf{e}_{dp}}. \end{aligned} \tag{57}$$

To construct the series for r^s, Δd, and dp, we use the same approach as that employed in Cochrane (2008). This consists of combining the series of the total return and the return excluding dividends associated with the value-weighted market portfolio. However, our market portfolio differs from the CRSP market portfolio in order to be consistent with the investment series used in the rest of the paper. The sample period is the same as in the previous sections (1965–2018).

The descriptive statistics associated with r^s, Δd, and dp are displayed in Table 7. We can see that the stock market return is slightly more volatile than aggregate dividend growth. As usual with stock returns, r^s is not serially correlated, whereas Δd appears to be mean-reverting, as indicated by the negative autocorrelation (-0.32). In contrast, the log dividend yield is a persistent variable, with an autoregressive slope of 0.95, which is in line with the existing evidence in the empirical asset pricing literature. The estimate of ρ_s is 0.98, which is clearly above the estimate for ρ.

The correlations among the three variables are presented in Panel B of Table 7. One noticeable feature is that those variables are much less mutually correlated than the investment variables analyzed in the previous sections (r, Δm, and mq). For example, the highest correlation, which occurs for r^s and Δd, is only 0.43. Table 7 also displays the correlation between the stock market variables and the investment variables. The correlations are typically modest in magnitude (below 0.30). The most relevant result is the negative correlation between dp and mq (-0.43). On the other hand, stock and investment returns are very weakly correlated (0.19). Further, marginal profits growth and dividend growth show a weak negative linear association (-0.22).

The VAR estimation results associated with dp are presented in Table 8. It turns out that, under both the restricted and unrestricted VAR approaches, the discount rate channel drives all the variation in dp from an economic viewpoint. Indeed, the long-run return slopes are above 100% in both cases, which implies that the slopes for dividend growth have the wrong sign (positive).

Critically, we cannot reject the null hypothesis (at the 10% level) that all the variation in dp is driven by long-run stock return predictability (t-ratios below one). The prominent role of the discount rate channel stems from the substantially larger magnitudes of the estimates of λ_r in comparison to λ_d under both VAR systems. These results are consistent with previous evidence on the predictability mix for the dividend yield associated with the value-weighted U.S. market portfolio (e.g., Cochrane, 2008, 2011; Rangvid, Schmeling, and Schrimpf, 2014; Maio and Santa-Clara, 2015, among others).[23]

The predictability pattern for dp represents basically the "mirror image" of the long-run predictability mix for mq estimated in the previous sections. What drives the sharp difference in results across the two variables? The descriptive statistics discussed above shed some light on such discrepancy. For example, the two predictors in the systems (mq and dp) are negative correlated, while the two return variables are only weakly positively correlated. Moreover, dp is more than twice as persistent than mq. We also observe that the stock return is more volatile than the investment return (volatilities of 0.16 and 0.10, respectively), while the opposite holds for dividend growth in relation to the growth in the marginal profitability of capital (volatilities of 0.12 and 0.23, respectively). A possible explanation is that stock returns are more subjective to short-term noise associated with investor behavioral biases in comparison to investment returns being affected by firm managers' behavioral biases. On the other hand, dividends are typically smoothed at the aggregate level (e.g., Fama and French, 2001; Brav, Graham, Harvey, and Michaely, 2005; Leary and Michaely, 2011), which might negatively affect the amount of dividend growth predictability emanated from the dividend yield. In sum, these results suggest a clear decoupling between the two groups of variables, which clearly affects the respective predictability results in the variance decompositions.

9 Conclusion

This paper derives a production-based present value relation in which fluctuations in the marginal profit to marginal Q ratio stem from changes in expected marginal profit growth, changes in expected investment returns, or both. We employ a parsimonious model with a standard pro-

[23] We get qualitatively similar results (that is, long-run return shares around or above one) by using either the value-weighted CRSP index (e.g., Cochrane, 2008, 2011; Maio and Santa-Clara, 2015) or the S&P 500 index (e.g., Campbell and Shiller, 1988; Campbell and Vuolteenaho, 2004; Maio and Philip, 2015; Maio and Xu, 2020).

duction and adjustment cost functions and estimate investment returns, marginal profits of capital, and the marginal values of capital for the aggregate of firms on the Compustat database. Following Gonçalves, Xue, and Zhang (2020), we correct for aggregation bias when conducting a GMM estimation of the share of capital in profits and the adjustment cost parameter. Correcting for aggregation bias is of utmost importance because aggregation bias can affect the main variance decomposition results produced in the paper.

Subsequently, we derive a present-value relation and show that variations in the ratio of the logarithm of the marginal profit of capital to the logarithm of marginal Q (mq) must emanate from shocks to expected future growth of the marginal profit of capital, shocks to expected investment returns, future values of mq, or any combination of these variables. We conduct predictability tests and find that the variation of the mq ratio is driven almost entirely by shocks to expected marginal profit growth, whereas shocks to expected investment returns play only a minor role. Thus, marginal profits growth is highly predictable at both short and long horizons. This finding is strikingly different from the findings in the asset pricing literature that all variations in the dividend-to-price ratio stem from discount rate fluctuations.

We also show that mq can forecast additional variables representing economic activity. Specifically, mq predicts negatively the growth rates of aggregate earnings, industrial production, and non-farm payrolls. Our finding can potentially guide future modeling in which the marginal profits of capital are highly time-varying.

Our main finding is driven by the combination of two sources. First, the profit function exhibits decreasing returns to scale with capital. This property of the profit function stems from declining marginal product of capital as well as any possible maintenance costs or other fixed costs that are proportional to the firm's stock of capital. Second, marginal profit is mean reverting. This property of marginal profit is consistent with the findings in other studies (e.g, Fama and French, 2000). Thus, a positive and mean reverting productivity shock entails a rise in both the marginal profit and in marginal Q. However marginal Q rises by less due to the adjustment costs of investment. The consequence is an increase in the mq ratio followed by a decline in the expected future growth of marginal profits.

We find that smaller values of the share of capital parameter and larger values of the adjustment cost parameter lead to a higher fraction of shocks to expected marginal profit growth as the source of variation in the mq ratio. The reason is that smaller values of the capital share in profits and larger adjustment cost parameter values render the marginal profit less persistent and hence it reverts faster to the mean, thus implying lower expected growth rates when mq rises.

We conduct several robustness checks, namely using portfolios sorted by Tobin's Q, as well as using portfolio medians in the GMM estimation, conducting simulation exercises, applying weighted long horizon regressions in the variance decomposition, as well as conducting the GMM estimation without correcting for aggregation bias. Our main qualitative results are robust in all those checks.

Overall, we find that marginal profit growth is highly predictable at both short and long horizons, implying that managers' assessments of the marginal value of capital differs from that of stock market investors.

Table 1: Descriptive Statistics

This table reports descriptive statistics for the log investment return (r), log growth in marginal profits (Δm), log profits-to-Q ratio (mq), log marginal profits (m), and log Q (q). The sample is 1965–2018. 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	0.11
Δm	0.01	0.23	−0.73	0.64	−0.08
mq	−1.88	0.14	−2.15	−1.26	0.46
m	−0.84	0.24	−1.30	0.03	0.54
q	1.03	0.10	0.85	1.29	0.61
Panel B (Correl.)					
	r	Δm	mq	m	q
r	1.00	0.92	0.63	0.64	0.59
Δm		1.00	0.53	0.48	0.39
mq			1.00	0.97	0.85
m				1.00	0.95
q					1.00

Table 2: Restricted VAR Estimates

This table reports the restricted VAR(1) estimation results when the predictor is the log profits-to-Q ratio (mq). The variables in the VAR are the log investment return (r), log growth in marginal profits (Δm), and mq. The results in Panel A correspond to the baseline set of investment variables. Results in Panels B to D correspond to alternative sets of the investment variables. λ, ϕ denote the VAR slopes associated with lagged mq, while t denotes the respective heteroskedasticity-robust t-statistics. 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 1965–2018. 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	−0.14	*−1.68*	0.04	−0.24	<u>−2.47</u>	**−12.81**
Δm	−0.75	**−3.48**	0.22	−1.24	**−2.58**	**−13.17**
mq	0.46	**2.88**	0.21			
			Panel B (Q Deciles)			
	λ, ϕ	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = 1)$
r	−0.17	<u>−1.97</u>	0.05	−0.27	**−2.99**	**−14.13**
Δm	−0.78	**−3.77**	0.23	−1.27	**−3.12**	**−14.52**
mq	0.45	**2.90**	0.20			
			Panel C (Median firm)			
	λ, ϕ	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = 1)$
r	−0.15	*−1.86*	0.05	−0.25	**−2.81**	**−14.28**
Δm	−0.76	**−3.66**	0.23	−1.25	**−2.94**	**−14.72**
mq	0.45	**2.90**	0.21			
			Panel D (Alternative series)			
	λ, ϕ	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = 1)$
r	0.22	1.50	0.04	0.74	1.22	−0.42
Δm	−0.08	−0.39	0.00	−0.26	1.22	−0.42
mq	0.85	**12.67**	0.72			

Table 3: Unrestricted VAR(1) Estimates

This table reports the unrestricted VAR(1) estimation results. The variables in the VAR are the log investment return (r), log growth in marginal profits (Δm), and the log profits-to-Q ratio (mq). The results in Panel A correspond to the baseline set of investment variables. Results in Panels B to D correspond to alternative sets of the investment variables. γ, θ, and $\lambda(\phi)$ denote the VAR slopes associated with lagged r, Δm, and mq, respectively. t denotes the respective heteroskedasticity-robust t-statistics. 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 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	γ	t	θ	t	$\lambda(\phi)$	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = 1)$
					Panel A					
r	0.02	0.07	0.17	1.12	−0.29	$\underline{-2.19}$	0.14	−0.51	**−5.11**	**−15.06**
Δm	0.40	0.57	0.09	0.25	−1.00	**−3.22**	0.27	−1.53	**−5.63**	**−16.37**
mq	0.45	0.93	−0.10	−0.41	0.35	1.60	0.23			
					Panel B (Q Deciles)					
r	0.03	0.10	0.17	1.16	−0.32	**−2.71**	0.16	−0.55	**−5.84**	**−16.48**
Δm	0.42	0.68	0.09	0.27	−1.05	**−3.86**	0.28	−1.56	**−6.38**	**−17.75**
mq	0.47	1.11	−0.10	−0.43	0.32	*1.73*	0.22			
					Panel C (Median firm)					
r	0.05	0.15	0.16	1.09	−0.30	$\underline{-2.38}$	0.16	−0.51	**−5.68**	**−16.85**
Δm	0.46	0.65	0.08	0.22	−1.02	**−3.42**	0.28	−1.52	**−6.32**	**−18.46**
mq	0.50	1.00	−0.10	−0.41	0.33	1.54	0.23			
					Panel D (Alternative series)					
r	−1.75	**−3.92**	1.40	**4.23**	0.48	**2.59**	0.31	0.16	0.30	−1.51
Δm	−2.46	**−4.41**	1.84	**4.36**	0.35	1.49	0.26	−0.83	0.30	−1.51
mq	−0.87	**−4.60**	0.54	**3.56**	1.07	**14.07**	0.78			

Table 4: Restricted VAR Estimates: VAR(2)

This table reports the restricted VAR(2) estimation results when the predictor is the log profits-to-Q ratio (mq). The variables in the VAR are the log investment return (r), log growth in marginal profits (Δm), and mq. λ_1, ϕ_1 and λ_2, ϕ_2 denote the VAR slopes associated with mq_t and mq_{t-1}, respectively, 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 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	λ_1, ϕ_1	t	λ_2, ϕ_2	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = 1)$
r	−0.03	−0.29	−0.24	$\underline{-2.04}$	0.13	−0.38	**−4.05**	**−14.59**
Δm	−0.61	**−2.60**	−0.28	−1.00	0.25	−1.39	**−4.32**	**−15.30**
mq	0.48	**3.02**	−0.03	−0.17	0.23			

Table 5: Long-Run Variance Decomposition for mq: Simulation

This table reports the simulation results for the long-run variance decomposition associated with the log profits-to-Q ratio (mq). The simulated series for the log investment return (r), log growth in marginal profits (Δm), and mq 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 either a restricted or an unrestricted VAR(1). b^{lr} denote the long-run coefficients (infinite horizon), while t represent the corresponding t-statistics. The original sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	Rest. VAR				**Unrest. VAR**			
	b_r^{lr}	t	b_m^{lr}	t	b_r^{lr}	t	b_m^{lr}	t
$\alpha = 0.05, a = 1.37$	0.08	(0.58)	−0.92	(**−6.65**)	−0.01	(−0.04)	−1.01	(**−6.15**)
$\alpha = 0.05, a = 6.85$	−0.17	($\underline{-1.99}$)	−1.17	(**−14.33**)	−0.38	(**−4.31**)	−1.39	(**−17.13**)
$\alpha = 0.05, a = 13.69$	−0.24	(**−3.11**)	−1.25	(**−16.62**)	−0.48	(**−6.15**)	−1.50	(**−20.84**)
$\alpha = 0.05, a = 20.53$	−0.27	(**−3.55**)	−1.28	(**−17.08**)	−0.53	(**−6.85**)	−1.54	(**−21.80**)
$\alpha = 0.05, a = 27.38$	−0.29	(**−3.78**)	−1.30	(**−17.19**)	−0.56	(**−7.20**)	−1.57	(**−22.13**)
$\alpha = 0.15, a = 1.37$	0.57	(**2.60**)	−0.42	(*−1.89*)	0.57	(**2.71**)	−0.42	($\underline{-2.00}$)
$\alpha = 0.15, a = 6.85$	0.14	(0.46)	−0.86	(**−2.90**)	−0.02	(−0.07)	−1.02	(**−2.88**)
$\alpha = 0.15, a = 13.69$	−0.12	(−0.57)	−1.12	(**−5.47**)	−0.41	(*−1.69*)	−1.42	(**−5.98**)
$\alpha = 0.15, a = 20.53$	−0.22	(−1.39)	−1.22	(**−7.92**)	−0.55	(**−3.15**)	−1.56	(**−9.34**)
$\alpha = 0.15, a = 27.38$	−0.26	($\underline{-2.01}$)	−1.27	(**−9.81**)	−0.60	(**−4.30**)	−1.61	(**−12.12**)
$\alpha = 0.30, a = 1.37$	0.76	(**4.64**)	−0.22	(−1.34)	0.73	(**4.81**)	−0.25	(−1.60)
$\alpha = 0.30, a = 6.85$	0.63	(*1.72*)	−0.36	(−0.98)	0.56	(1.58)	−0.42	(−1.17)
$\alpha = 0.30, a = 13.69$	0.31	(0.71)	−0.69	(−1.58)	0.15	(0.29)	−0.85	(*−1.70*)
$\alpha = 0.30, a = 20.53$	0.06	(0.17)	−0.93	($\underline{-2.54}$)	−0.21	(−0.47)	−1.21	(**−2.70**)
$\alpha = 0.30, a = 27.38$	−0.08	(−0.28)	−1.08	(**−3.70**)	−0.42	(−1.18)	−1.43	(**−4.03**)
$\alpha = 0.50, a = 1.37$	0.83	(**6.52**)	−0.15	(−1.16)	0.80	(**6.74**)	−0.18	(−1.46)
$\alpha = 0.50, a = 6.85$	0.88	(**3.29**)	−0.10	(−0.39)	0.80	(**3.34**)	−0.18	(−0.74)
$\alpha = 0.50, a = 13.69$	0.77	(*1.75*)	−0.22	(−0.49)	0.67	(1.61)	−0.32	(−0.76)
$\alpha = 0.50, a = 20.53$	0.54	(1.04)	−0.45	(−0.85)	0.41	(0.75)	−0.58	(−1.05)
$\alpha = 0.50, a = 27.38$	0.32	(0.64)	−0.67	(−1.32)	0.12	(0.21)	−0.87	(−1.48)
$\alpha = 0.70, a = 1.37$	0.86	(**7.85**)	−0.12	(−1.10)	0.84	(**8.14**)	−0.15	(−1.39)
$\alpha = 0.70, a = 6.85$	0.96	(**4.66**)	−0.03	(−0.12)	0.88	(**4.74**)	−0.10	(−0.54)
$\alpha = 0.70, a = 13.69$	0.98	(**2.86**)	−0.01	(−0.02)	0.87	(**2.85**)	−0.12	(−0.38)
$\alpha = 0.70, a = 20.53$	0.88	(*1.85*)	−0.11	(−0.24)	0.76	(*1.75*)	−0.23	(−0.51)
$\alpha = 0.70, a = 27.38$	0.70	(1.27)	−0.29	(−0.52)	0.58	(1.05)	−0.41	(−0.75)

Table 6: Long-horizon Regressions for Economic Activity

This table reports the results for bivariate long-horizon regressions associated with future economic activity at horizons (H) of 1, 5, 10, 15, and 20 years ahead. The forecasting variables are the log profits-to-Q ratio (mq) and the current macro variable. The variables that proxy for broad economic activity are the log growth in aggregate earnings (Δe), log growth in industrial production (Δip), and low growth in non-farm payrolls (Δnfp). The original sample is 1965–2018. In each regression, η_H denotes the slope estimates associated with mq. The corresponding t-ratios, which are based on Newey and West (1987) standard errors (computed with $H-1$ lags), are displayed in parentheses. R^2 denotes the coefficient of determination of a given regression. Bold, underlined, and italics t-ratios indicate statistical significance at the 1%, 5%, and 10% level, respectively.

	$H=1$	$H=5$	$H=10$	$H=15$	$H=20$
		Panel A (Δe)			
η_H	−0.99	−0.86	−1.01	−0.71	−0.44
	(−**3.86**)	(<u>−2.36</u>)	(−**4.23**)	(−**4.17**)	(−**3.62**)
R^2	0.20	0.38	0.42	0.29	0.14
		Panel B (Δip)			
η_H	−0.10	−0.17	−0.32	−0.41	−0.27
	(−**2.86**)	(<u>−2.57</u>)	(−**3.53**)	(−**3.69**)	(−1.05)
R^2	0.15	0.11	0.18	0.28	0.18
		Panel C (Δnfp)			
η_H	−0.05	−0.13	−0.24	−0.29	−0.20
	(−**4.44**)	(−**4.41**)	(−**6.37**)	(−**5.62**)	(−1.45)
R^2	0.41	0.17	0.30	0.27	0.08

Table 7: Descriptive Statistics: Stock Market Variables

This table reports descriptive statistics for the log stock return (r^S), log dividend growth (Δd), and log dividend-to-price ratio (dp). The sample is 1965–2018. AR(1) designates the first-order autocorrelation. The correlations between the stock market variables and the investment variables (r, Δm, and mq) are presented in Panel B.

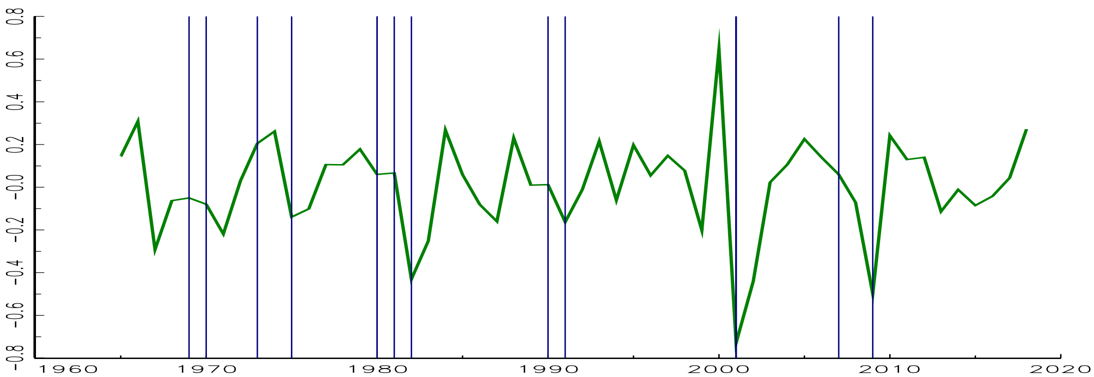
			Panel A			
	Mean	S.D.	Min.	Max.	AR(1)	
r^S	0.09	0.16	−0.30	0.49	−0.04	
Δd	0.06	0.12	−0.18	0.38	−0.32	
dp	−3.68	0.45	−4.80	−2.91	0.95	
			Panel B (Correl.)			
	r^S	Δd	dp	r	Δm	mq
r^S	1.00	0.43	0.04	0.19	0.27	0.17
Δd		1.00	0.19	−0.24	−0.22	−0.17
dp			1.00	−0.15	−0.06	−0.43

Table 8: VAR Estimates: dp

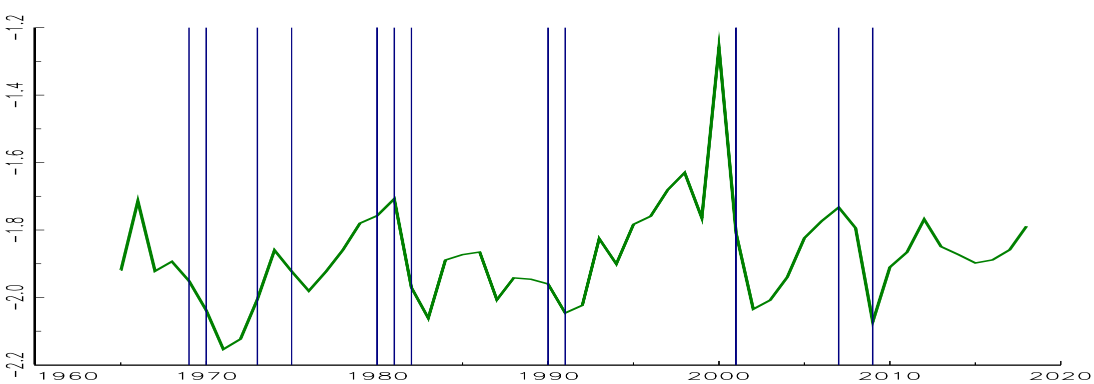
This table reports the estimation results for first-order VARs containing the stock market variables. The variables in the VARs are the log stock return (r^S), log dividend growth (Δd), and the log dividend-to-price ratio (dp). Panels A and B show the results for the restricted and unrestricted VAR, respectively. γ, θ, and $\lambda(\phi)$ denote the VAR slopes associated with lagged r^S, Δd, and dp, respectively. t denotes the respective heteroskedasticity-robust t-statistics. 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 ($b_r^{lr} = 0, b_d^{lr} = -1$) and ($b_r^{lr} = 1, b_d^{lr} = 0$), respectively. The original sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

	γ	t	θ	t	$\lambda(\phi)$	t	R^2	b^{lr}	$t(b_r^{lr} = 0)$	$t(b_r^{lr} = 1)$
					Panel A (Rest. VAR)					
r^S					0.10	*1.79*	0.08	1.25	<u>2.30</u>	0.45
Δd					0.02	0.53	0.00	0.25	<u>2.28</u>	0.45
dp					0.95	**17.31**	0.89			
					Panel B (Unrest. VAR)					
r^S	−0.13	−0.85	0.23	1.31	0.09	*1.66*	0.10	1.14	**3.14**	0.38
Δd	−0.28	<u>−2.06</u>	−0.19	−1.26	0.03	1.11	0.22	0.14	**3.10**	0.37
dp	−0.15	−0.93	−0.43	**−2.63**	0.97	**19.76**	0.91			

Panel A (r)



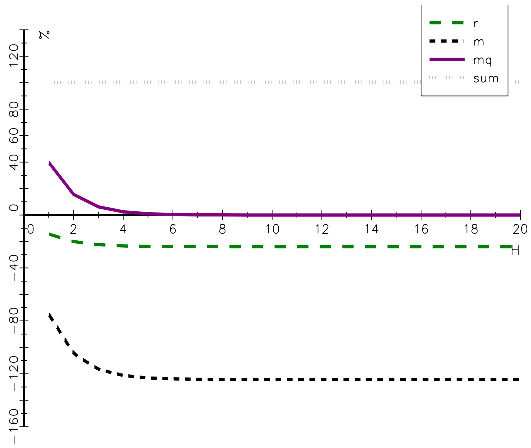
Panel B (Δm)



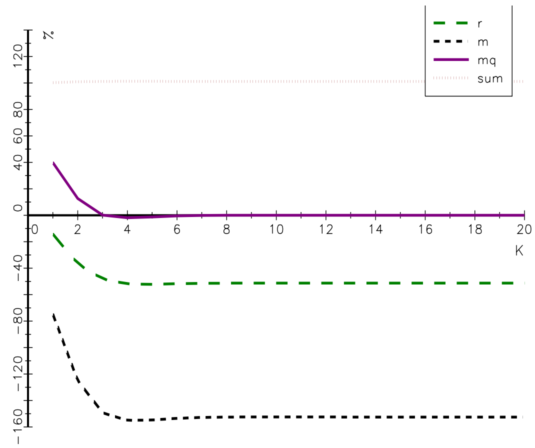
Panel C (mq)

Figure 1: Time-Series for r, Δm, and mq

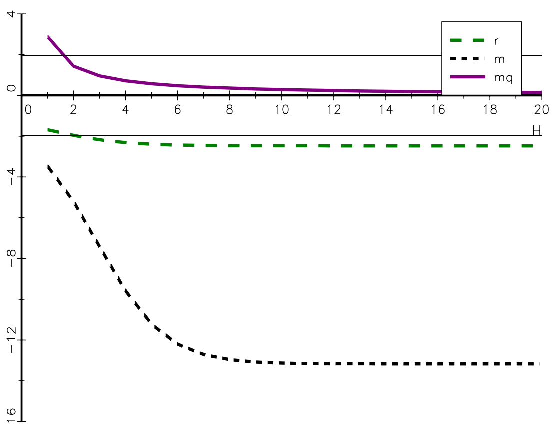
This figure plots the time-series for the log investment return (r), log growth in marginal profits (Δm), and log marginal profit-to-Q ratio (mq). The bars contain the years with NBER recessions (the 1980 and 2001 recessions are indicated by a single line). The sample is 1965 to 2018.



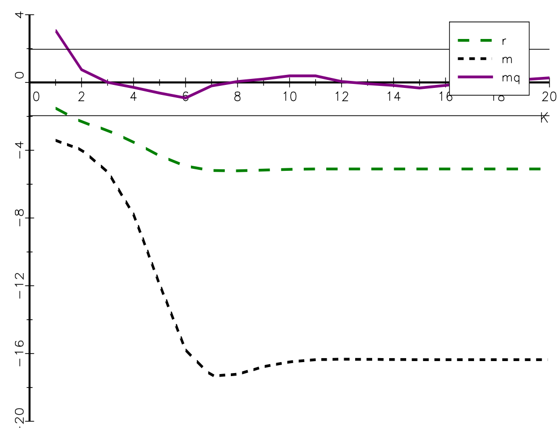
Panel A (Rest. VAR, slopes)



Panel B (Unrest. VAR, slopes)



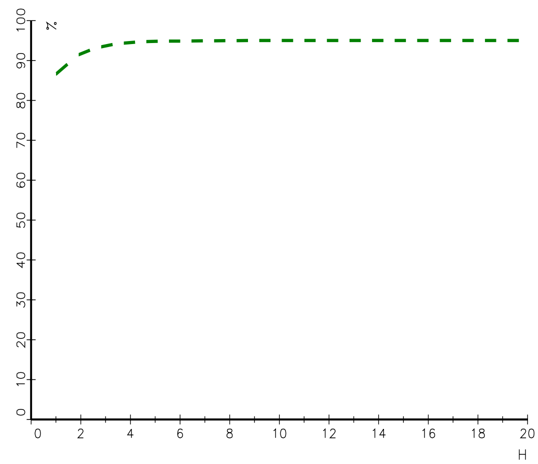
Panel C (Rest. VAR, t-stats)



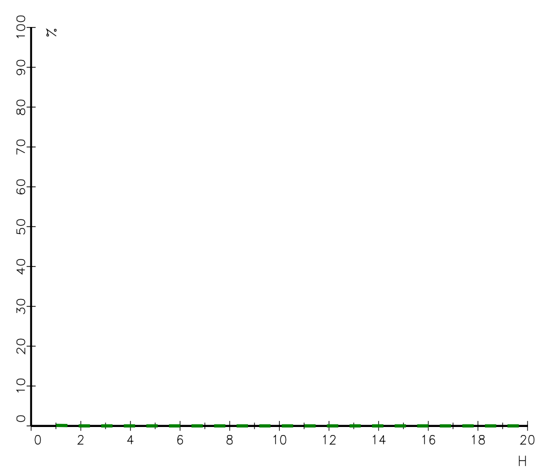
Panel D (Unrest. VAR, t-stats)

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 the log profits-to-Q ratio (mq). 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 growth in marginal profits (m), and future mq. The forecasting variable is mq 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 1965 to 2018.



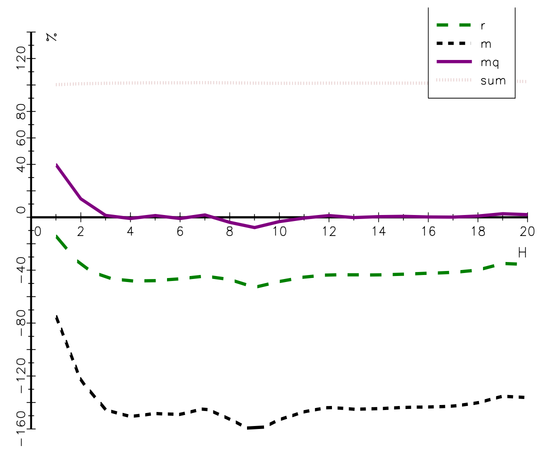
Panel A (r)



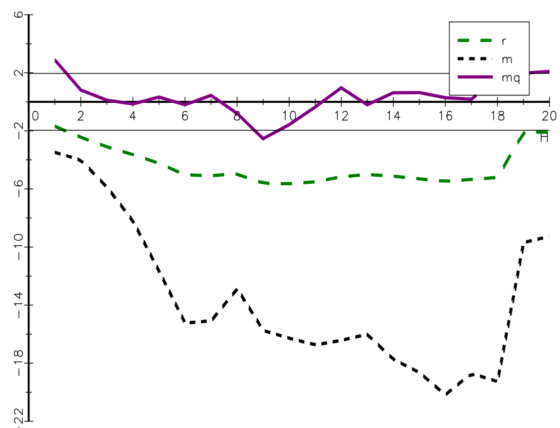
Panel B (m)

Figure 3: Bootstrap Simulation

This figure plots the simulated p-values for the restricted VAR-based return (r) and profitability growth (m) slopes from a Bootstrap simulation with 10,000 replications. The predictive variable is the log profits-to-Q ratio (mq). The numbers indicate the fraction of pseudo samples under which the return (profitability) coefficient is higher (lower) than the corresponding estimates from the original sample. H represents the number of years ahead. The original sample is 1965 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 the log profits-to-Q ratio (mq). The predictive slopes are obtained from weighted long-horizon regressions. The coefficients are associated with the log investment return (r), log growth in marginal profits (m), and future mq. The forecasting variable is mq 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 1965 to 2018.

References

Avino, D.E., Stancu, A., and Simen, C.W., 2020. The predictive power of the dividend risk premium, *Journal of Financial and Quantitative Analysis*, Forthcoming.

Bakke, T.E. and Whited, T.M., 2010. Which firms follow the market? an analysis of corporate investment decisions, *Review of Financial Studies*, 23, 1941–1980.

Bakshi, G., Gao, X., and Rossi, A.G., 2019. Understanding the sources of risk underlying the cross section of commodity returns, *Management Science*, 65, 619–641.

Belo, F., Gala, V., Salomao, J., and Vitorino, M.A., 2019. Decomposing firm value, *SSRN Working Paper*.

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

Binsbergen, J.H.V. and Koijen, R.S.J., 2010. Predictive regressions: A present-value approach, *Journal of Finance*, 65, 1439–1471.

Blanchard, O., Rhee, C., and Summers, L., 1993. The stock market, profit, and investment, *Quarterly Journal of Economics*, 108, 77–114.

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.

Brav, A., Graham, J.R., Harvey, C.R., and Michaely, R., 2005. Payout policy in the 21st century, *Journal of Financial Economics*, 77, 483–527.

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

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.

Chen, L., Da, Z., and Priestley, R., 2012. Dividend smoothing and predictability, *Management Science*, 58, 1834–1853.

Choi, H., Mueller, P., and Vedolin, A., 2017. Bond variance risk premiums, *Review of Finance*, 21, 987–1022.

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., 1992. Explaining the variance of price-dividend ratios, *Review of Financial Studies*, 5, 243–280.

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. and Maio, P., 2019. Asset growth, profitability, and investment opportunities, *Management Science*, 65, 3988–4010.

Cooper, I., Maio, P., and Yang, C., 2021. What drives marginal q fluctuations?, *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.

Doms, M. and Dunne, T., 1998. Capital adjustment patterns in manufacturing plants, *Review of Economic Dynamics*, 1, 409–429.

Engsted, T., Pedersen, T.Q., and Tanggaard, C., 2012. The log-linear return approximation, bubbles, and predictability, *Journal of Financial and Quantitative Analysis*, 47, 643–665.

Faccini, R., Konstantinidi, E., Skiadopoulos, G., and Sarantopoulou-Chiourea, S., 2019. A new predictor of U.S. real economic activity: The S&P 500 option implied risk aversion, *Management Science*, 65, 4927–4949.

Fama, E.F. and French, K.R., 1988. Dividend yields and expected stock returns, *Journal of Financial Economics*, 22, 3–25.

Fama, E.F. and French, K.R., 1989. Business conditions and expected returns on stocks and bonds, *Journal of Financial Economics*, 25, 23–49.

Fama, E.F. and French, K.R., 2000. Forecasting profitability and earnings, *Journal of Business*, 73, 161–175.

Fama, E.F. and French, K.R., 2001. Disappearing dividends: Changing firm characteristics or lower propensity to pay?, *Journal of Financial Economics*, 60, 3–43.

Faust, J., Gilchrist, S., Wright, J.H., and Zakrajsek, E., 2013. Credit spreads as predictors of real-time economic activity: a bayesian model-averaging approach, *Review of Economics and Statistics*, 95, 1501–1519.

Gilchrist, S. and Zakrajsek, E., 2012. Credit spreads and business cycle fluctuations, *American Economic Review*, 102, 1692–1720.

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

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.

Hong, H. and Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices?, *Journal of Financial Economics*, 105, 473–490.

Jordà, O. and Kozicki, S., 2011. Estimation and inference by the method of projection minimum distance: An application to the new keynesian hybrid phillips curve, *International Economic Review*, 52, 461–487.

Keim, D.B. and Stambaugh, R.F., 1986. Predicting returns in the stock and bond markets, *Journal of Financial Economics*, 17, 357–390.

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.

Leary, M.T. and Michaely, R., 2011. Determinants of dividend smmothing: Empirical evidence, *Review of Financial Studies*, 24, 3197–3249.

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., 2013a. The "Fed model" and the predictability of stock returns, *Review of Finance*, 17, 1489–1533.

Maio, P., 2013b. 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 Philip, D., 2015. Macro variables and the components of stock returns, *Journal of Empirical Finance*, 33, 287–308.

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.

Mian, A., Sufi, A., and Verner, E., 2017. Household debt and business cycles worldwide, *Quarterly Journal of Economics*, 132, 1755–1817.

Morck, R., Shleifer, A., and Vishny, R.W., 1990. The stock market and investment: Is the market a sideshow?, *Brookings Papers on Economic Activity*, 2, 157–215.

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

Pettenuzo, D., Sabbatucci, R., and Timmermann, A., 2020. Cash flow news and stock price dynamics, *Journal of Finance*, 75, 2221–2270.

Piatti, I. and Trojani, F., 2019. Dividend growth predictability and the price-dividend ratio, *Management Science*, Forthcoming.

Rangvid, J., Schmeling, M., and Schrimpf, A., 2014. Dividend predictability around the world, *Journal of Financial and Quantitative Analysis*, 49, 1255–1277.

Stock, J.H. and Watson, M.W., 2003. Forecasting output and inflation: the role of asset prices, *Journal of Economic Literature*, 41, 788–829.

Tobin, J., 1969. A general equilibrium approach to monetary theory, *Journal of Money, Credit and Banking*, 1, 15–29.

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.

Warusawitharana, M. and Whited, T.M., 2016. Equity market misvaluation, financing, and investment, *Review of Financial Studies*, 29, 603–654.

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

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

A Online Appendix

A.1 Variance Decompositions

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

A.1.1 Restricted VAR(1)

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

$$\mathrm{var}(mq_t) \approx -\,\mathrm{Cov}\left(\sum_{h=1}^{H} \rho^{h-1} \Delta m_{t+h}, mq_t\right) + \mathrm{Cov}\left(\sum_{h=1}^{H} \rho^{h-1} r_{t+h}, mq_t\right) + \mathrm{Cov}\left(\rho^H mq_{t+H}, mq_t\right). \tag{A.1}$$

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

$$1 \approx -\beta\left(\sum_{h=1}^{H} \rho^{h-1} \Delta m_{t+h}, mq_t\right) + \beta\left(\sum_{h=1}^{H} \rho^{h-1} r_{t+h}, mq_t\right) + \beta\left(\rho^H mq_{t+H}, mq_t\right), \tag{A.2}$$

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

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

$$1 \approx -\sum_{h=1}^{H} \rho^{h-1} \beta(\Delta m_{t+h}, mq_t) + \sum_{h=1}^{H} \rho^{h-1} \beta(r_{t+h}, mq_t) + \rho^H \beta(mq_{t+H}, mq_t). \tag{A.3}$$

Under the restricted first-order VAR, we have,

$$mq_{t+h-1} = \phi^{h-1} mq_t + \phi^{h-1} \sum_{l=1}^{h-1} \phi^{-l} (\pi_{mq} + \varepsilon^{mq}_{t+l}), \tag{A.4}$$

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

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

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

$$r_{t+h} = \pi_r + \phi^{h-1}\lambda_r mq_t + \phi^{h-1}\lambda_r \sum_{l=1}^{h-1} \phi^{-l}(\pi_{mq} + \varepsilon_{t+l}^{mq}) + \varepsilon_{t+h}^{r}. \quad \text{(A.6)}$$

Since $\text{Cov}(\varepsilon_{t+l}^{mq}, mq_t) = 0, l > 0$ and $\text{Cov}(\varepsilon_{t+h}^{r}, mq_t) = 0$, by construction, it follows that

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

Similarly, we have,

$$\beta(\Delta m_{t+h}, mq_t) = \phi^{h-1}\lambda_m. \quad \text{(A.8)}$$

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

$$mq_{t+H} = \phi^H mq_t + \phi^H \sum_{l=1}^{H} \phi^{-l}(\pi_{mq} + \varepsilon_{t+l}^{mq}), \quad \text{(A.9)}$$

we have

$$\beta(mq_{t+H}, mq_t) = \phi^H, \quad \text{(A.10)}$$

which leads to

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

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

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

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

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

The matrix of derivatives is given by

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

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 mq_t is given by

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

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

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

$$\begin{aligned}
\sum_{h=1}^{H} \rho^{h-1}\beta(r_{t+h}, mq_t) &= \sum_{h=1}^{H} \frac{\mathbf{e}_r' \rho^{h-1}\mathbf{A}^h \mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}} \\
&= \frac{\mathbf{e}_r'}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}} \left(\sum_{h=1}^{H} \rho^{h-1}\mathbf{A}^h \right) \mathbf{\Sigma}\mathbf{e}_{mq} \\
&= \frac{\mathbf{e}_r'}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}\rho} \left(\sum_{h=1}^{H} \rho^{h}\mathbf{A}^h \right) \mathbf{\Sigma}\mathbf{e}_{mq} \\
&= \frac{\mathbf{e}_r'(\rho\mathbf{A} - \rho^{H+1}\mathbf{A}^{H+1})(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_{mq}}{\rho\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}} \\
&= \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I} - \rho^{H}\mathbf{A}^{H})(\mathbf{I} - \rho\mathbf{A})^{-1}\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}. \qquad \text{(A.18)}
\end{aligned}$$

The H-period Δm slope is defined in a similar way. The slope associated with future mq at $t + H$ is derived as follows:

$$\beta(mq_{t+H}, mq_t) = \frac{\text{Cov}(\mathbf{e}_{mq}'\mathbf{z}_{t+H}, \mathbf{e}_{mq}'\mathbf{z}_t)}{\text{var}(\mathbf{e}_{mq}'\mathbf{z}_t)} = \frac{\text{Cov}(\mathbf{e}_{mq}'\mathbf{A}^H\mathbf{z}_t, \mathbf{e}_{mq}'\mathbf{z}_t)}{\text{var}(\mathbf{e}_{mq}'\mathbf{z}_t)} = \frac{\mathbf{e}_{mq}'\mathbf{A}^H\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}, \qquad \text{(A.19)}$$

which implies that

$$\rho^H \beta(mq_{t+H}, mq_t) = \frac{\rho^H\mathbf{e}_{mq}'\mathbf{A}^H\mathbf{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\mathbf{\Sigma}\mathbf{e}_{mq}}. \qquad \text{(A.20)}$$

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

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

where $\mathbf{b} \equiv (\gamma_m, \theta_m, \lambda_m, \gamma_r, \theta_r, \lambda_r, \gamma_{mq}, \theta_{mq}, \phi)'$. The derivatives are obtained from numerical methods.[24]

A.2 Bootstrap Simulation

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

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

1. We estimate the first-order restricted VAR,

$$
\begin{aligned}
r_{t+1} &= \pi_r + \lambda_r mq_t + \varepsilon^r_{t+1}, \\
\Delta m_{t+1} &= \pi_m + \lambda_m mq_t + \varepsilon^m_{t+1}, \\
mq_{t+1} &= \pi_{mq} + \phi mq_t + \varepsilon^{mq}_{t+1},
\end{aligned}
$$

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

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

$$
(\varepsilon^{r,u}_t, \varepsilon^{m,u}_t, \varepsilon^{mq,u}_t)', t = v^u_1, ..., v^u_T,
$$

where the time indices $v^u_1, ..., v^u_T$—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),

$$
r_{u,t+1} = -\rho\varepsilon^{mq,u}_{t+1} + \varepsilon^{m,u}_{t+1},
$$

for Δm (no-profit predictability null),

$$
\Delta m_{u,t+1} = \varepsilon^{m,u}_{t+1},
$$

and for mq:

$$
mq_{u,t+1} = \phi mq_{u,t} + \varepsilon^{mq,u}_{t+1}.
$$

The value of mq_u for the base period ($mq_{u,1}$) is picked at random from one of the observations of mq_t.

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

$$
\begin{aligned}
r_{u,t+1} &= \pi_{r,u} + \lambda_{r,u} mq_{u,t} + \upsilon_{t+1}^{r,u}, \\
\Delta m_{u,t+1} &= \pi_{m,u} + \lambda_{m,u} mq_{u,t} + \upsilon_{t+1}^{m,u}, \\
mq_{u,t+1} &= \pi_{mq,u} + \phi_u mq_{u,t} + \upsilon_{t+1}^{r,u},
\end{aligned}
$$

and estimate the implied long-horizon slopes,

$$
\begin{aligned}
b_{r,u}^H &\equiv \lambda_{r,u} \frac{1-\rho_u^H \phi_u^H}{1-\rho_u \phi_u}, \\
b_{m,u}^H &\equiv \lambda_{m,u} \frac{1-\rho_u^H \phi_u^H}{1-\rho_u \phi_u},
\end{aligned}
$$

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

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

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

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