# What drives Q and investment fluctuations?

Ilan Cooper Paulo Maio Chunyu Yang[1]

This version: April 2020[2]

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

[2]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, as well as seminar participants at Ben Gurion University, University of Haifa, Hebrew University, Monash University, University of New South Wales, University of Sydney, and conference participants at the New Economic School 25th Anniversary Conference, 2018 Finance Down Under Conference, 2018 Midwest Finance Association (San Antonio), 2018 Cavalcade (Yale School of Management), 2018 Asian Meeting of the Econometric Society (Sogang University), 2018 European Finance Association (Warsaw), 2018 European Economic Association (Cologne), 2018 SAFE Asset Pricing Workshop (Goethe University Frankfurt), and the 2018 Corporate Policies and Asset Pricing conference at Cass Business School for helpful comments. Any remaining errors are our own.

**Abstract**

A dynamic present-value relation implies that fluctuations in the marginal profit of capital to the marginal Q ratio ($mq$) are driven by shocks to expected growth of the marginal profit of capital (cash-flow channel), investment return shocks (discount-rate channel), or both. We find that, in contrast to the dividend-to-price ratio, $mq$ strongly predicts expected marginal profitability growth at both short and long forecasting horizons, but not investment returns. This result is driven by both investment's response to expected marginal profit shocks and by mean reversion of profitability growth. Firms' technology parameters play a role in the variance decomposition of $mq$.

Keywords: Asset pricing; Tobin's $q$; Present-value model; Investment return; Variance decomposition; VAR implied predictability; Aggregation bias; Marginal profit of capital; Long-horizon regressions;

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 profitability entailed by installing an additional unit of capital. Thus, similar to stock prices, variations in marginal Q are driven by shocks to expected cash flows as well as by discount rate shocks.

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 $mq$. We employ the supply approach to valuation as in Belo, Xue, and Zhang (2013). That is, we use firms' investment decisions to infer the marginal value of capital. The $mq$ ratio is an analog of the earnings yield or the dividend-to-price ratio variables from the stock market. However, while the stock price is determined by investors in the stock market, it is firms' investment decisions which inform us about their assessment of the marginal value of capital.

As in Liu, Whited, and Zhang (2009) we employ the q-theory of investment. In the estimation of the model's parameters, namely the share of capital in profit and the adjustment cost parameter, we follow Gonçalves, Xue, and Zhang (2019) and correct for aggregation bias when conducting the GMM estimation of the technology parameters (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 weighted average of firm-level investment returns (the investment return moment). Second, we match weighted average Tobin's marginal Q in the data to a weighted average model-implied Tobin's Q (the valuation moment).

We derive a dynamic present-value relation for $mq$, in which $mq$ is negatively correlated with the future multi-period log growth rate in the marginal profit of capital ($m$), and positively correlated with future log investment returns ($r$). This present-value relation is analogous to the present-value relation associated with the 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, marginal profits growth, and future $mq$ at a terminal date. 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.[1]

We use two methods to estimate empirically the variance decomposition for $mq$: first-order restricted VAR (as in Cochrane, 2008) and an unrestricted VAR (as in Larrain and Yogo, 2008 and Maio and Xu, 2019). The two methods produce qualitatively similar variance decompositions at both intermediate and long horizons. In all cases, the bulk of variation in $mq$ turns out to be marginal profit growth predictability, with predictability of investment returns assuming only a minor role at intermediate and long horizons. Specifically, the shares of return (marginal profits growth) predictability at the 20-year horizon are -0.24 (-1.24) and -0.51 (-1.53) under the restricted VAR 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 a rather marginal role. This finding stands in stark contrast to the finding in the asset pricing literature that the aggregate dividend-to-price ratio can forecast stock market returns but not aggregate dividend growth (see Cochrane, 2008, 2011).

These findings are robust to using median stock returns and investment returns instead of value-weighted stock returns and investment returns, as well as to conducting the GMM estimation of the structural investment model based on deciles sorted by marginal Q portfolios. 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

[1] For example, Stigler (1963) (p. 54) states: "important proposition in economic theory than that, under competition, the rate of return on investment tends toward equality in all industries. Entrepreneurs will seek to leave relatively unprofitable industries and enter relatively profitable industries."

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 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.

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 rise 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.[2] 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.[3]

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 given in Bloom, 2009). We find the following. First, higher values of the share of capital parameter are associated with lower fraction of marginal profit 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 fraction of the marginal

[2]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.

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

profit growth predictability. The reason is that the marginal profit becomes considerably more volatile when this parameter rises.[4]

In the last part of the paper, we conduct a variance decomposition associated with the aggregate dividend yield ($dp$) 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% in both cases, 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 yield 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. Answering this question is beyond the scope of the paper. One possibility 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 (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. The issue of the effect of investor sentiment on investment is still debated though. We do recognize that our estimation involves the valuation moment and hence the average marginal Q implied by the model to the average marginal Q in the data (which involves the market value). However, equality of the averages does not imply a period-by-period equality. Other explanations for our findings surely exist.

Our work is related to Abel and Blanchard (1986), who derive and estimate an approximation for the two components of marginal Q, namely expected returns and expected

[4] 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.

marginal profit (rather than marginal profit growth rate). They find that most of the variability in their measure for marginal Q is generated by variability in the cost of capital. Importantly, Abel and Blanchard (1986) use expected stock returns as the discount rate, whereas we predict investment returns, not stock returns, with $mq$. 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 is preferable. In related work, Cochrane (1991) finds that the aggregate investment-to-capital ratio predicts the future stock market return as well as future investment returns. Cochrane's result suggests that the cost of capital is a driver of investment, which seems to be somewhat at odds with the conclusion reached by Abel and Blanchard (1986).

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 $mq$. 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 $mq$ under the restricted VAR case. 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), produces a severe misspecification in the estimated variance decomposition, as suggested by our results. Furthermore, Yashiv's model includes labor as a quasi-fixed input in production, whereas our model (as the one in Abel and Blanchard, 1986; Liu, Whited, and Zhang, 2009; Belo, Xue, and Zhang, 2013; among many others) includes only capital as a costly adjustable input in production.

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). There are several key differences between the two papers. First, the focus of our analysis is $mq$ rather than $q$, which represent quite different variables. Second, 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 driving log Q and investment. Third, 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. 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 not about explaining the forecasting power of investment growth for future cash flows and returns, unlike our decomposition that is centered on the predictive power emanating from $mq$. Second, their driving forces of investment growth are equity-based variables (e.g., 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), while our approach critically relies exclusively on investment-based variables that are estimated from a structural model.

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 by computing variance decompositions for the financial ratios. Most of the work is concentrated on the predictability by the dividend yield (Cochrane, 1992, 2008, 2011; Chen, 2009; Binsbergen and Koijen, 2010; Engsted, Pedersen, and Tanggaard, 2012; Rangvid, Schmeling, and Schrimpf, 2014; Maio and Santa-Clara, 2015). On the other hand, several studies compute variance decompositions for other financial ratios like 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, 2013; Maio and Xu, 2019).

The rest of the paper is organized as follows. In Section 2, we present a model of a firm's optimal investment decisions. Section 3 describes the data and the econometric methodology for estimating the production and adjustment costs parameters and the components of investment returns. We derive a variance decomposition for the log marginal profits-to-Q ratio in Section 4, while Section 5 represents the main empirical analysis conducted in the paper. In Section 6, we provide a sensitivity analysis. Section 7 shows the results for a comparative statics exercise. In Section 8, we estimate a variance decomposition for the dividend-to-price ratio. The paper concludes in Section 9.

# 2 The Background Model

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

## 2.1 Theoretical Model

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

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

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

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

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

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

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

The firm's investment return is given by

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

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

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

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

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

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

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

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

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

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

## 2.2 A Model of Aggregation and Aggregate Investment Return

Given the firm-level parameters: the capital share ($\alpha$) and the adjustment cost parameter ($a$), using Equation (4) to (6), for each firm, we can construct a time series of the firm's investment returns and the associated two decomposition components: the firm's marginal profitability of capital ($M$) and the firm's marginal value of capital ($Q$). At the aggregate market level, we need to construct the aggregate investment returns and a similar representation as Equation (6) but with two aggregate components, namely aggregate marginal profitability of capital and aggregate marginal value of capital. This is needed in order to derive the present-value relation connecting the aggregate marginal value of capital to future aggregate marginal profitability of capital and future aggregate investment returns.

Let $N$ be the number of firms in the market. Each firm optimizes by equating the marginal adjustment costs of investment to the marginal value of capital. Each firm makes an investment $I_{i,t}$ at time $t$, and exiting time $t$ with a level of capital stock $K_{i,t+1}$. Given the constant returns to scale assumption, applying the result of Hayashi (1982) implies that the marginal value of capital is equal to the average value of capital, and hence the firm's value at the end of time $t$ is given by $K_{i,t+1}Q_{i,t}$ where $Q_{i,t}$ is the marginal value of capital at the end of time $t$.[5] The aggregate market value is therefore $\sum_{i=1}^{N} K_{i,t+1}Q_{i,t}$. We measure the aggregate marginal value of capital at the end of time $t$, denoted by $\overline{Q}_t$, by assuming that $\overline{Q}_t$ can 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:

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

Equivalently, $\overline{Q}_t$ can be expressed as

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

Thus, $\overline{Q}_t$ 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,

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

invested proportionally to the fraction of capital of each firm from the total capital in the economy will indeed have a value of $\overline{Q}_t$. [6]

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

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

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

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

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

$$R_{t+1} \equiv \frac{\overline{M}_{t+1} + (1-\delta)\,\overline{Q}_{t+1}}{\overline{Q}_t}. \tag{11}$$

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

[6] 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.

# 3 Estimating the Investment Return and its Components

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

## 3.1 Methodology to Estimate Firm-level Parameters

We follow Gonçalves, Xue, and Zhang (2019) and estimate the firm-level parameters, the capital share ($\alpha$) and the adjustment cost parameter ($a$), using one-step GMM to fit the investment Euler equation moment for each testing portfolio jointly with an additional moment, namely the valuation moment as in Belo, Xue, and Zhang (2013). We include the valuation moment in the estimation because in the benchmark estimation we consider the aggregate market portfolio as the testing portfolio and using only the investment Euler equation leads to an unidentified estimation with one moment but two parameters: the capital share and the adjustment cost parameter.[7] With two moments and two parameters, the estimation is exactly identified and the two moments fit perfectly. Specifically, for a given set of testing portfolios (indexed by $j$), the first set of moment conditions corresponds to testing whether the average stock return equals the average levered investment return for each testing portfolio $j$,

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

where $\mathrm{E}_T(\cdot)$ denotes the sample moment, $R^S_{j,t+1}$ is the portfolio stock return, and $R^{Iw}_{j,t+1}$ is the portfolio levered investment return that depends on parameters $\alpha$ and $a$. The second set of moment conditions tests whether the average Tobin's $Q$ in the data equals the average $Q$ predicted by the model:

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

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

where $A_{j,t}$ is the total assets and $\widetilde{Q}_{j,t}$ is the Tobin's $Q$ in the data defined as $\widetilde{Q}_{j,t} \equiv (P_{j,t} + B_{j,t+1}) / A_{j,t}$.

Firm-level accounting variables and, thus, firm-level investment returns are subject to the issue of outliers. The outliers in firm-level investment returns can contaminate the aggregate portfolio-level investment returns and lead to noisy parameter estimates from the GMM estimation. To alleviate the impact of outliers, we follow Gonçalves, Xue, and Zhang (2019) and construct firm-level investment returns using winsorized firm-level accounting variables[8], then compute equal- or value-weighted portfolio levered investment returns to match with equal- or value-weighted portfolio stock returns. Instead of winsorization, we also follow an alternative approach in Belo, Gala, Salomao, and Vitorino (2019), where portfolio median is used to to aggregate firm-level investment returns to portfolio level since median is known to be robust to outliers. More details about the GMM estimation methodology are provided in the online appendix.

## 3.2 Data

We largely follow Gonçalves, Xue, and Zhang (2019) and Belo, Xue, and Zhang (2013) in measuring accounting variables and in aligning their timing with the timing of stock returns. Our sample consists of all common stocks on NYSE, Amex, and Nasdaq from 1963 to 2018. The firm-level data are from the merged CRSP and COMPUSTAT industrial database. We include all firms with fiscal year ending in the second half of the calendar year. We exclude firms with primary standard industrial classifications between 4900 and 4999 (utilities) and between 6000 and 6999 (financials). We also delete firm-year observations for which total assets, capital stock, or sales are either zero or negative.

Capital stock ($K_{i,t}$) is net property, plant, and equipment (Compustat annual item PPENT). Investment ($I_{i,t}$) is capital expenditures (Compustat annual item CAPX) minus sales of property, plant, and equipment (Compustat annual item SPPE, zero if missing). Total debt ($B_{i,t+1}$) is long-term debt (Compustat annual item DLTT, zero if missing) plus short-term debt (Compustat annual item DLC, zero if missing). $A_{i,t}$ is total assets (Compustat annual item AT). Market equity ($P_{i,t}$) is the stock price per share (CRSP item prc) times the number of shares outstanding (CRSP item shrout). Market leverage ($w_{i,t}$) is the

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

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. We follow Cochrane (1991) and assume a depreciation rate ($\delta$) equal to 0.1.[9] Output ($Y_{i,t}$) is sales (Compustat annual item SALE). Market leverage, $w_{i,t}$, is the ratio of total debt to the sum of total debt and the market value of equity. In the benchmark estimation (i.e. the estimation that does not correct for aggregation-bias) we aggregate firm-specific characteristics to aggregate- and portfolio-level characteristics as in Fama and French (1995). For example, $Y_{i,t+1} = K_{i,t+1}$ is the sum of sales in year $t$ of all the firms in our sample divided by the sum of capital stocks at the beginning of year $t+1$ for the same set of firms. We proceed in a similar way for the aggregation of the remaining characteristics. We measure the tax rate ($\tau_t$) as the statutory corporate income tax (from the Commerce Clearing House, annual publications). The after-tax corporate bond returns ($R^{Ba}_{i,t+1}$) are computed from $R^{B}_{i,t+1}$ using the average of tax rates in year $t$ and $t+1$. For the pre-tax corporate bond returns ($R^{B}_{i,t+1}$) we use the ratio of total interest and related expenses (Compustat annual item XINT) scaled by the total debt ($B_{i,t+1}$).[10]

At the end of June of year $t$, we construct the aggregate "market" portfolio. That is, a portfolio whose value is the value of the aggregate capital stocks and whose return is the aggregate investment returns. Alternatively, we sort all stocks on Tobin's $Q$ at the end of June of year $t$ into deciles based on the NYSE breakpoints. For each testing portfolio, we compute annual value-weighted stock returns from July of year $t$ to June of year $t+1$. We construct annual levered investment returns to match with annual stock returns and annual valuation ratios to match with annual Tobin's $Q$. To construct the matching annual levered investment returns, we use capital at the end of fiscal year $t-1$ ($K_{i,t}$), the tax rate, investment, and capital at the end of year $t$ ($\tau_t$, $I_{i,t}$, and $K_{i,t+1}$), as well as other variables at the end of year $t+1$ ($\tau_{t+1}$, $Y_{i,t+1}$, and $I_{i,t+1}$). To match with $\widetilde{Q}_{i,t}$ for portfolios formed at the end of June of year $t$, we take $I_{i,t}$ from the fiscal year ending in calendar year $t$ and $K_{i,t}$ from the fiscal year ending in year $t-1$.[11]

[9]In the data, the mean value of the annual depreciation rate is equal to 0.1089.

[10]As shown in Gonçalves, Xue, and Zhang (2019), doing so increases the sample coverage by 12.7% as compared to the prior studies that use credit rating imputation such as Liu, Whited, and Zhang (2009).

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

### 3.3 Structural Parameter Estimates

In the basecase GMM esitmation, the testing portfolio is the aggregate stock market portfolio and the portfolio returns are measured as the value-weighed returns. The estimation is exactly identified and the two moments fit perfectly. The estimate of capital share ($\alpha$) is 0.08, which is similar to the results in Gonçalves, Xue, and Zhang (2019).[12] The estimate of adjustment cost parameter ($a$) is 15.18, which is higher than the results in Gonçalves, Xue, and Zhang (2019) based on testing portfolios different from our setup, but similar to the results in Liu, Whited, and Zhang (2009).[13] The estimated magnitude of the adjustment costs is 11.09% of sales, which is in line with those reported in prior studies.[14] Given parameters $\alpha$ and $a$ estimated at the firm level, we compute the aggregate investment return and its components, by plugging these parameter estimates in Equations (9) to (11), together with firm-level accounting variables.

## 4 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.[15] 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 ($R$) derived in the last section,

$$R_{t+1} = \frac{(1-\delta)Q_{t+1} + M_{t+1}}{Q_t}, \tag{14}$$

over the year, such as $I_{i,2003}$, from the 2003 income or cash flow statement.

[12] Gonçalves, Xue, and Zhang (2019) report estimates of capital share varying from 5.04% to 7.53% across different testing portfolios in Table 5 Panel B.

[13] Gonçalves, Xue, and Zhang (2019) report estimates of adjustment cost parameter varying from 0.72 to 5.66 in Table 5 Panel B and from 1.63 to 8.11 in Table 3 Panel B. Liu, Whited, and Zhang (2009) report estimates varying from 11.5 to 28.9 when matching two moments: expected returns and variances.

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

[15] The ratio of the marginal profitability of capital to the marginal value of capital (Q) is roughly proportional to the output-to-investment ratio plus investment growth (in case capital is slow moving). See the discussions in Cochrane (1991) and Chen, Da, and Larrain (2016).

where $Q$ represents the aggregate marginal value of capital and $M$ stands for aggregate marginal profits of capital. This definition is analog to the usual definition of the gross stock return with $Q$ 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 (14), and proceeding along the lines of Campbell and Shiller (1988), we obtain the following difference equation in the log profits-to-Q ratio,

$$mq_t \approx const. + \rho mq_{t+1} + r_{t+1} - \Delta m_{t+1}, \tag{15}$$

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.

$\rho$ plays an important role in the analysis, representing a (log-linearization) discount coefficient that depends on the mean of $mq$,

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

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

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

$$mq_t \approx const. + \sum_{h=1}^{H} \rho^{h-1} r_{t+h} - \sum_{h=1}^{H} \rho^{h-1} \Delta m_{t+h} + \rho^H mq_{t+H}. \tag{16}$$

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 time $t + H$ ($mq_{t+H}$), and negatively correlated with future multi-period log growth in marginal profits ($\Delta m_{t+h}$). 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, both future $\Delta m$ 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 present-value relation:

$$mq_t \approx const. + \sum_{h=1}^{\infty} \rho^{h-1} r_{t+h} - \sum_{h=1}^{\infty} \rho^{h-1} \Delta m_{t+h}. \tag{17}$$

Hence, at very long horizons, only predictability of future investment returns and profits growth drives the variation in the current $mq$ ratio.[16] Which of these two components matters most in terms of driving the dynamics of $mq$ remains an empirical question, which will be addressed in the following sections.

Table 1 (Panel A) presents the descriptive statistics for the variables in the present-value relation for $mq$. 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.[17] On the other hand, $\Delta m$ 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 marginally higher persistence compared to $mq$ itself, especially in the case of $q$. 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 $r$, $\Delta m$, and $mq$. All three variables appear to be mean-reverting to a large degree, and hence, stationary. Both $r$ and $\Delta m$ are clearly procyclical variables, as they tend to decline around most recession periods. This pattern is especially

[16] Lettau and Ludvigson (2002) derive a related dynamic accounting decomposition for the log Tobin's Q, rather than $mq$. Further, their present-value relation is based on a second-order Taylor expansion.

[17] 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).

notable during the last two recession periods (2001 and 2007–2009). On the other hand, $mq$ appears as being substantially less correlated with the business cycle. Regarding the critical parameter $\rho$, which is a function of the average $mq$ ratio, we obtain an estimate of 0.85. This estimate is somewhat smaller than corresponding estimates for the analogue parameter in present-value relations associated with stock returns (typically above 0.90), which are based on the average log dividend-to-price ratio (e.g., Campbell and Vuolteenaho, 2004; Cochrane, 2008, 2011; Maio and Santa-Clara, 2015).

Panel B of Table 1 shows the correlations among the three variables mentioned above. The correlation between $\Delta m$ 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 (14)). 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 positive association between marginal profits growth and $mq$ (correlation around 0.50). Table 1 also displays the correlations with both $m$ and $q$. It turns out that the main variables of interest ($r$, $\Delta m$, and $mq$) are somewhat more positively correlated with $m$ than $q$. However, there is a very high degree of association between $m$ and $q$ (correlation of 0.95). Thus, at times of high marginal profitability of capital, investment is also high. This is consistent with the well known cash-flow sensitivity of investment.

# 5 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.

## 5.1 Restricted VAR

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

$$
\begin{aligned}
r_{t+1} &= \pi_r + \lambda_r mq_t + \varepsilon^{r}_{t+1}, && (18)\\
\Delta m_{t+1} &= \pi_m + \lambda_m mq_t + \varepsilon^{m}_{t+1}, && (19)\\
mq_{t+1} &= \pi_{mq} + \phi mq_t + \varepsilon^{mq}_{t+1}, && (20)
\end{aligned}
$$

where the $\varepsilon$s represent error terms. This VAR system is estimated by multiple-equation OLS (see Hayashi, 2000), with Newey and West (1987) $t$-statistics (computed with one lag).

By combining the VAR above with the present-value relation in Equation (16), we obtain an approximate identity involving the predictability coefficients associated with $mq_t$, at each forecasting horizon $H$:

$$
\begin{aligned}
1 &\approx b^H_r - b^H_m + b^H_{mq}, && (21)\\
b^H_r &\equiv \lambda_r \frac{(1-\rho^H\phi^H)}{1-\rho\phi},\\
b^H_m &\equiv \lambda_m \frac{(1-\rho^H\phi^H)}{1-\rho\phi},\\
b^H_{mq} &\equiv \rho^H\phi^H.
\end{aligned}
$$

This equation can be interpreted as a variance decomposition for $mq$. The predictive coefficients, $b^H_r$, $-b^H_m$, and $b^H_{mq}$, represent the fraction of the variance of current $mq$ attributed to the predictability of future investment returns, growth in marginal profits, and $mq$, respectively. Hence, these slopes measure the weight (of the predictability) of each of these variables ($\sum_{h=1}^{H}\rho^{h-1}r_{t+h}$, $\sum_{h=1}^{H}\rho^{h-1}\Delta m_{t+h}$, 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 above.

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

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

$$\begin{aligned} 1 &\approx b_r^{lr} - b_m^{lr}, \\ b_r^{lr} &\equiv \frac{\lambda_r}{1-\rho\phi}, \\ b_m^{lr} &\equiv \frac{\lambda_m}{1-\rho\phi}. \end{aligned} \tag{22}$$

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

Following Cochrane (2008), we compute $t$-statistics for two joint null hypotheses of long-run predictability: the first null assumes that there is only 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 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 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 $\Delta m$ 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 $\Delta m$ (22%) in comparison to the return regression (4%). Such different forecasting performance for future $r$ and $\Delta m$ 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.

### 5.2 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,

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \begin{pmatrix} 0 \\ \rho\phi - 1 \\ \phi \end{pmatrix} mq_t + \begin{pmatrix} \varepsilon^m_{t+1} - \rho\varepsilon^{mq}_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \end{pmatrix}, \tag{23}$$

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 $\Delta m$. This means that all the variation in $mq$ comes from predicting future investment returns:

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \begin{pmatrix} 1 - \rho\phi \\ 0 \\ \phi \end{pmatrix} mq_t + \begin{pmatrix} \varepsilon^m_{t+1} - \rho\varepsilon^{mq}_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \end{pmatrix}. \tag{24}$$

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 $mq_t$. One key advantage of the bootstrap simulation relative to a monte-carlo simulation is that we can skip the normally-distributed assumption for the variables in the system. 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 slopes (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.

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

The results associated with the bootstrap simulation are presented in Figure 3. The simulation confirms the strong statistical significance of the $\Delta m$ 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 coefficients associated with future profits growth present robust statistical significance across methods.

### 5.3 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):

$$
\begin{aligned}
r_{t+1} &= \pi_r + \gamma_r r_t + \theta_r \Delta m_t + \lambda_r mq_t + \varepsilon^r_{t+1}, && (25)\\
\Delta m_{t+1} &= \pi_m + \gamma_m r_t + \theta_m \Delta m_t + \lambda_m mq_t + \varepsilon^m_{t+1}, && (26)\\
mq_{t+1} &= \pi_{mq} + \gamma_{mq} r_t + \theta_{mq} \Delta m_t + \phi mq_t + \varepsilon^{mq}_{t+1}. && (27)
\end{aligned}
$$

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 (2019) 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,

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_m \\ \pi_{mq} \end{pmatrix} + \begin{pmatrix} \gamma_r & \theta_r & \lambda_r \\ \gamma_m & \theta_m & \lambda_m \\ \gamma_{mq} & \theta_{mq} & \phi \end{pmatrix} \begin{pmatrix} r_t \\ \Delta m_t \\ mq_t \end{pmatrix} + \begin{pmatrix} \varepsilon^r_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \end{pmatrix}. \tag{28}$$

Equivalently, the VAR can be defined as

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{29}$$

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 Newey-West $t$-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 $\mathbf{\Sigma} \equiv \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}'_t)$. Given these definitions, and following Larrain and Yogo (2008) and Maio and Xu (2019), we derive the following variance decomposition for $mq$ at each horizon $H$,

$$\begin{aligned} 1 &\approx b^H_r - b^H_m + b^H_{mq}, \\ b^H_r &\equiv \frac{\mathbf{e}'_r \mathbf{A}(\mathbf{I} - \rho^H \mathbf{A}^H)(\mathbf{I} - \rho \mathbf{A})^{-1} \mathbf{\Sigma} \mathbf{e}_{mq}}{\mathbf{e}'_{mq} \mathbf{\Sigma} \mathbf{e}_{mq}}, \\ b^H_m &\equiv \frac{\mathbf{e}'_m \mathbf{A}(\mathbf{I} - \rho^H \mathbf{A}^H)(\mathbf{I} - \rho \mathbf{A})^{-1} \mathbf{\Sigma} \mathbf{e}_{mq}}{\mathbf{e}'_{mq} \mathbf{\Sigma} \mathbf{e}_{mq}}, \\ b^H_{mq} &\equiv \frac{\rho^H \mathbf{e}'_{mq} \mathbf{A}^H \mathbf{\Sigma} \mathbf{e}_{mq}}{\mathbf{e}'_{mq} \mathbf{\Sigma} \mathbf{e}_{mq}}, \end{aligned} \tag{30}$$

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

Further details on the derivation of this variance decomposition are available on the online appendix. The expressions above show that the relative shares of predictability (e.g., $b_r^H/b_m^H$) change with the forecasting horizon, in contrast to the restricted VAR case. In other words, the relative importance of the different predictability channels is not a mechanical function of the one-year VAR slope estimates, that is, 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

$$\begin{aligned} 1 &\approx b_r^{lr} - b_m^{lr}, \\ b_r^{lr} &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I}-\rho\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}}, \\ b_m^{lr} &\equiv \frac{\mathbf{e}_m'\mathbf{A}(\mathbf{I}-\rho\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}}. \end{aligned} \tag{31}$$

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

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

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \end{pmatrix} = \widehat{\boldsymbol{\pi}} + \begin{pmatrix} 0.02(0.08) & 0.17(1.16) & -0.29(\underline{-2.39}) \\ 0.40(0.62) & 0.09(0.26) & -1.00(\mathbf{-3.55}) \\ 0.45(1.04) & -0.10(-0.43) & 0.35(\mathit{1.79}) \end{pmatrix} \begin{pmatrix} r_t \\ \Delta m_t \\ mq_t \end{pmatrix} + \begin{pmatrix} \widehat{\varepsilon}^{r}_{t+1} \\ \widehat{\varepsilon}^{m}_{t+1} \\ \widehat{\varepsilon}^{mq}_{t+1} \end{pmatrix},$$

with $R^2$ estimates of 0.14, 0.27, and 0.23, respectively. The numbers in parentheses represent the $t$-ratios, with bold, underlined, and italic numbers denoting significance at the 1%, 5%, and 10%, respectively.

These results show that the slopes associated with lagged $r$ and $\Delta m$ are largely insignificant in all three VAR equations. We also note that the estimates for both $\lambda_r$ and $\lambda_m$ increase in magnitude relative to the corresponding estimates in the restricted VAR. On the other hand, the estimate for $\phi$ is lower than the corresponding estimate in the benchmark VAR, being only marginally significant (10% level). Nevertheless, because the estimates for the $\gamma$s and $\theta$s 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). The results are indeed qualitatively similar to those in the decomposition corresponding to the benchmark VAR. Specifically, the shares of 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 $\Delta m$ 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 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. However, 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.

All in all, the punch line of these results is that predictability of future profitability growth (the cash-flow channel) is the major driving force of variation in $mq$.

### 5.4 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 associated with $mq$. This analysis seems 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,

$$\begin{aligned}
\lambda_r &= \frac{\mathrm{Cov}(r_{t+1}, mq_t)}{\mathrm{var}(mq_t)} \\
&= \frac{\mathrm{Cov}(r_{t+1}, m_t)}{\mathrm{var}(mq_t)} - \frac{\mathrm{Cov}(r_{t+1}, q_t)}{\mathrm{var}(mq_t)} \\
&= \frac{\mathrm{Cov}(r_{t+1}, m_t)}{\mathrm{var}(m_t)} \frac{\mathrm{var}(m_t)}{\mathrm{var}(mq_t)} - \frac{\mathrm{Cov}(r_{t+1}, q_t)}{\mathrm{var}(q_t)} \frac{\mathrm{var}(q_t)}{\mathrm{var}(mq_t)} \\
&= \mu_r \omega_m - \eta_r \omega_q, \qquad (32)
\end{aligned}$$

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

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

$$\begin{aligned}
\lambda_m &= \frac{\mathrm{Cov}(\Delta m_{t+1}, mq_t)}{\mathrm{var}(mq_t)} \\
&= \frac{\mathrm{Cov}(\Delta m_{t+1}, m_t)}{\mathrm{var}(m_t)} \frac{\mathrm{var}(m_t)}{\mathrm{var}(mq_t)} - \frac{\mathrm{Cov}(\Delta m_{t+1}, q_t)}{\mathrm{var}(q_t)} \frac{\mathrm{var}(q_t)}{\mathrm{var}(mq_t)} \\
&= \mu_m \omega_m - \eta_m \omega_q. \qquad (33)
\end{aligned}$$

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

$$\begin{aligned}
r_{t+1} &= \tau_r + \mu_r m_t + \nu^r_{t+1}, \qquad (34) \\
\Delta m_{t+1} &= \tau_m + \mu_m m_t + \nu^m_{t+1}, \qquad (35)
\end{aligned}$$

while the slopes corresponding to $q_t$ are retrieved from:

$$\begin{aligned}
r_{t+1} &= \varpi_r + \eta_r q_t + \varsigma^r_{t+1}, \qquad (36) \\
\Delta m_{t+1} &= \varpi_m + \eta_m q_t + \varsigma^m_{t+1}. \qquad (37)
\end{aligned}$$

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

$$b_r^{lr} = \frac{\lambda_r}{1-\rho\phi} = \frac{\mu_r}{1-\rho\phi}\omega_m - \frac{\eta_r}{1-\rho\phi}\omega_q, \tag{38}$$

$$b_m^{lr} = \frac{\lambda_m}{1-\rho\phi} = \frac{\mu_m}{1-\rho\phi}\omega_m - \frac{\eta_m}{1-\rho\phi}\omega_q. \tag{39}$$

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,

$$\begin{aligned} r_{t+1} &= -0.03(-0.67) - 0.11(-2.07)m_t, R^2 = 0.06, \\ \Delta m_{t+1} &= -0.38(-3.08) - 0.46(-3.49)m_t, R^2 = 0.22, \end{aligned}$$

where the numbers in parentheses denote the Newey–West $t$-ratios.

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

$$\begin{aligned} r_{t+1} &= 0.35(2.76) - 0.29(-2.23)q_t, R^2 = 0.08, \\ \Delta m_{t+1} &= 1.01(3.17) - 0.97(-3.04)q_t, R^2 = 0.18. \end{aligned}$$

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 $m$.[18] Given these estimates and the estimated variance weights, $\omega_m = 2.69$ and $\omega_q = 0.49$, the long-run return and profits growth coefficients associated with $mq$ are decomposed as follows:

$$\begin{aligned} b_r^{lr} &= \frac{\mu_r}{1-\rho\phi}\omega_m - \frac{\eta_r}{1-\rho\phi}\omega_q = -0.47 - (-0.23) = -0.24 \\ b_m^{lr} &= \frac{\mu_m}{1-\rho\phi}\omega_m - \frac{\eta_m}{1-\rho\phi}\omega_q = -2.03 - (-0.79) = -1.24. \end{aligned}$$

[18] Why would $q$ forecast negative $\Delta m$? The reason is that $q$ rises when there is a temporary (albeit quite persistent) rise in $m$. In untabulated results, we find that $q$ predicts positively the level, not the growth rate, of $m$. Interestingly, $m$ forecasts negatively future investment returns. This is likely due to the fact that when $m$ rises so does $q$ (the two are highly positively correlated as is evident in Table 1). A rise in $q$ stems from a rise in the investment-to-capital ratio and this rise is associated with lower future investment returns.

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 other words, the estimate of $\eta_r$ ($-0.29$) is not negative enough in order to dominate the negative estimate of $\mu_r$ ($-0.11$) and produce an estimate of $\lambda_r$ with the correct sign (positive). On the other hand, the estimate of $\eta_m$ ($-0.97$) is not negative enough in order to attenuate the estimate of $\mu_m$ ($-0.46$) so that the resulting estimate of $\lambda_m$ has a lower magnitude (i.e., lower than one in magnitude), that is, 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.

# 6 Sensitivity Analysis

In this section, we provide a sensitivity analysis to the empirical results discussed in the previous section. To save space, the results are presented in the online appendix.

## 6.1 Direct Approach

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

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

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

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

$$\rho^H mq_{t+H} = a_{mq}^H + b_{mq}^H mq_t + \varepsilon_{t+H}^{mq}. \tag{42}$$

The estimation is conducted by equation-by-equation OLS and the $t$-statistics for the direct predictive slopes are based on Newey and West (1987) standard errors with $H$ lags (i.e., the Bartlett Kernel with a bandwidth of $H+1$), 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 (16) with the predictive regressions presented above, we obtain an approximate identity involving the predictability coefficients associated with $mq_t$, at each forecasting horizon $H$:

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

If the first-order VAR does not fully capture the dynamics of the data generating process for $r$, $mq$, and $\Delta m$, it follows that the variance decomposition will be a poor approximation of the true decomposition for $q$, as discussed in Cochrane (2008) and Maio and Xu (2019).[19] This problem does not exist under the direct approach, which a priori should yield the most correct estimates for the variance decomposition (see Cochrane, 2008, 2011; Maio and Santa-Clara, 2015). The minus side of the direct approach is that with small or moderate samples, the statistical power of the long-horizon regressions is negatively affected at very long horizons, given the decline in the number of usable observations.

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 $\Delta m$ become larger than one in magnitude. This indicates that the predictability of future marginal profits

[19] In a broader scope, there are several examples in the economics literature showing an inconsistency between the short-run dynamics of stochastic processes and the implied long-run dynamics (e.g., Campbell and Shiller, 1988; Bandi, Perron, Tamoni, and Tebaldi, 2019; Gonçalves, 2019).

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. The 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% in 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.

## 6.2 Alternative Investment Series I

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 basecase. The estimate of capital share ($\alpha$) is 0.07 when matching value-weighted portfolio returns and 0.08 when matching median portfolio returns. The estimate of adjustment cost parameter ($a$) is respectively 20.31 and 20.89 and the corresponding ratio of adjustment-cost-to-sales is respectively 14.84% and 15.27%. The model is not rejected by the $\chi^2$-test, with a $p$-value around 0.44.

The predictability mixes are very similar to those estimated with the benchmark data. Specifically, the long-run return and profits growth slopes are $-0.27$ and $-1.27$, respectively, in the case of the restricted VAR. When the estimation is based on the unrestricted VAR, 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 predictability patterns are very similar to those estimated in the benchmark case. In particular, the long-run return and profits growth slopes are $-0.25$ and $-1.25$, respectively, in the case of the restricted VAR. The corresponding estimates in the case of the unrestricted VAR are $-0.51$ and $-1.52$, respectively.

## 6.3 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

$$\begin{aligned} r_{t+1} &= \pi_r + \lambda_{r1} mq_t + \lambda_{r2} mq_{t-1} + \varepsilon^r_{t+1}, && (44) \\ \Delta m_{t+1} &= \pi_m + \lambda_{m1} mq_t + \lambda_{m2} mq_{t-1} + \varepsilon^m_{t+1}, && (45) \\ mq_{t+1} &= \pi_{mq} + \phi_1 mq_t + \phi_2 mq_{t-1} + \varepsilon^{mq}_{t+1}. && (46) \end{aligned}$$

The VAR(2) is estimated by multiple-equation OLS, with Newey–West $t$-statistics (computed with two lags).

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

$$\begin{pmatrix} r_{t+1} \\ \Delta m_{t+1} \\ mq_{t+1} \\ r_t \\ \Delta m_t \\ mq_t \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_m \\ \pi_{mq} \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & \lambda_{r1} & 0 & 0 & \lambda_{r2} \\ 0 & 0 & \lambda_{m1} & 0 & 0 & \lambda_{m2} \\ 0 & 0 & \phi_1 & 0 & 0 & \phi_2 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} r_t \\ \Delta m_t \\ mq_t \\ r_{t-1} \\ \Delta m_{t-1} \\ mq_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon^r_{t+1} \\ \varepsilon^m_{t+1} \\ \varepsilon^{mq}_{t+1} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (47)$$

or equivalently,

$$\mathbf{z}_{t+1} = \boldsymbol{\pi} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}. \quad (48)$$

The indicator vectors as defined as follows,

$$\mathbf{e}_r \equiv (1,0,0,0,0,0)', \quad (49)$$

$$\mathbf{e}_m \equiv (0,1,0,0,0,0)', \quad (50)$$

$$\mathbf{e}_{mq} \equiv (0,0,1,0,0,0)', \quad (51)$$

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

$$\mathbf{\Sigma} = \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}_t') = \begin{pmatrix} \mathrm{var}(r_t) & \mathrm{Cov}(r_t, \Delta m_t) & \mathrm{Cov}(r_t, mq_t) & 0 & 0 & 0 \\ \mathrm{Cov}(r_t, \Delta m_t) & \mathrm{var}(\Delta m_t) & \mathrm{Cov}(\Delta m_t, mq_t) & 0 & 0 & 0 \\ \mathrm{Cov}(r_t, mq_t) & \mathrm{Cov}(\Delta m_t, mq_t) & \mathrm{var}(mq_t) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \quad (52)$$

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, the $t$-ratios for the long-horizon coefficients are obtained by applying the delta method.

The results 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 not significant in the regressions associated with $\Delta m_{t+1}$ and $mq_{t+1}$. Hence, the estimate of $\phi_1$ (0.48) is very similar to the corresponding estimate in the restricted VAR. Moreover, the $R^2$ estimates in the profits growth and $mq$ regressions are quite close to the corresponding values under the restricted VAR.

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. Indeed, the cash-flow channel plays the dominant role in terms of driving variation in $mq$ at all forecasting horizons, and the corresponding slope estimates are strongly significant (1% level). Specifically, the long-run return and profits growth slopes are $-0.38$ and $-1.39$, respectively. This indicates that the share of profits growth predictability is somewhat higher than under the restricted VAR(1), bet less extreme than that associated with the unrestricted VAR(1).

## 6.4 Alternative Investment Series II

We conduct the variance decomposition for $mq$ by using other time series of the investment variables. In contrast to the rest of the paper, we rely on the structural estimated 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 ($\alpha$) 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 ($Phi/Y$) of 11.76%.

The predictability mix in the restricted-VAR case is almost symmetric to the corresponding mix obtained with the benchmark data: At intermediate and long horizons, it is the case that return predictability drives most of the variation in $mq$. Specifically, the long-run return and profit growth estimates are 0.74 and $-0.26$, respectively. However, the corresponding standard errors are large as none of the return slope estimates is statistically significant at the 10%. When the slope estimates are based on the unrestricted VAR, 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: except at short forecasting horizons, what drives the variation in $mq$ is the predictability of future profitability growth, with return predictability assuming a secondary role. At an infinite horizon, the estimates of the $r$ and $\Delta m$ slopes are 0.16 and $-0.83$, respectively. The profit growth coefficient estimates are not significant at the 10% level. However, by employing single-sided $p$-values, those estimates are significant at most forecasting horizons. Using single-sided $p$-values is a less conservative approach, 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 (2019) showing that such method can provide quite misleading long-run variance decompositions for aggregate stock market valuation ratios.[20] In related work, Yashiv (2016) derives a long-run variance

[20] Maio and Xu (2019) 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

decomposition for the log Q-to-capital productivity, which is approximately equal to the symmetric of $mq$.[21] 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.[22] However, following the discussion above, the findings obtained in Yashiv (2016) are likely not robust by employing better methods to produce variance decompositions, such as the unrestricted first-order VAR approach (or the direct method). 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.

# 7 Comparative Statics

In this section, we conduct a comparative statics exercise.

Specifically, we estimate a range of variance decompositions associated with $mq$ for a set of artificial series of the key investment variables, $r$, $\Delta m$, and $mq$. The artificial time-series are obtained from calibration of the two key structural parameters of the theoretical model presented in Section 2, $\alpha$ and $a$.[23] 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 last section show that the empirical 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 based on the restricted VAR can differ substantially by using these two different sets of the investment variables.

The simulation results are presented in Table 3. We calibrate five different values for $\alpha$

Jordà and Kozicki (2011) (rather than by OLS).

[21]His long-run decomposition involves four predictability sources, compared to two sources in our long-run decomposition for $mq$. For example, he has 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.

[22]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.

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

(0.05, 0.15, 0.30, 0.50, and 0.70), and five values for $a$ (1.37, 6.85, 13.69, 20.53, and 27.38). In the case of $a$, these values are associated with a calibration of the adjustment cost-to-output ratio ($\Phi/Y$) of 0.01, 0.05, 0.10, 0.15, and 0.20, respectively. Hence, we have a total of 25 ($5 \times 5$) different artificial data sets, which are used in the computation of the variance decomposition. To save space and keep the focus, we report only the long-run (infinite horizon) variance decompositions for $mq$. We use both the restricted and unrestricted VAR methods to produce the decompositions.

The first key pattern that emerges from Table 3 is that the share of profits growth predictability declines with the magnitude of $\alpha$. 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 sole exception occurs for the pair $\alpha = 0.50, a = 27.38$. However, even in that case, the profits growth slope is 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 profits slope is $-0.42$ under both the restricted and unrestricted VAR approaches, which is significant at the 5% (unrestricted VAR) or 10% (restricted VAR) level. High values of $\alpha$ 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 is that, for a given value of $\alpha$, the shares of cash-flow predictability tend to increase with the magnitude of $a$. This is especially evident when $\alpha$ assumes small or intermediate values ($0.05 \leq \alpha \leq 0.30$), in which cases the relation is monotonic. For example, when $\alpha = 0.30$, the cash-flow channel assumes a dominant role (shares above 50%) only for intermediate and high values of $a$ ($13.69 \leq a \leq 27.38$). For the $1.37 \leq a \leq 6.85$ range, it turns out that return predictability is the major driving force of $mq$. The likely reason for this patten is that $m$ becomes substantially more volatile for higher values of $a$. We also observe across the board that the weights of profits growth predictability tend to be larger under the unrestricted VAR than under the restricted VAR method. This result is consistent with the evidence obtained in Section 5.

All in all, the results of this section show that the dominant role of 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 return predictability under less plausible values for those parameters.

# 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 main 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. Hence, the first variable is associated with the supply side of the stock market, while the second variable is related to the demand side. Hence, it makes sense to compare the variance decompositions associated with these two 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

$$dp_t \approx const. + \sum_{j=1}^{\infty} \rho_s^{j-1} r_{t+j}^s - \sum_{j=1}^{\infty} \rho_s^{j-1} \Delta d_{t+j}, \tag{53}$$

where $const.$ is a constant term that is irrelevant for the forthcoming analysis; $r_{t+j}^s \equiv \ln(R_{t+j}^s)$ denotes the log stock return between $t+j-1$ and $t+j$; and $\Delta d_{t+j} \equiv \ln(D_{t+j}/D_{t+j-1})$ represents the log dividend growth between $t+j-1$ and $t+j$. $\rho_s$ 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 $r$, $\Delta d$ is analog to $\Delta m$, and $dp$ plays the role of $mq$. $\rho_s$ is analog to $\rho$.

Following Cochrane (2008), Engsted, Pedersen, and Tanggaard (2012), among others, the long-run variance decomposition associated with the dividend yield based on the restricted

VAR is given by

$$
\begin{aligned}
1 &\approx b_r^{lr} - b_d^{lr}, \\
b_r^{lr} &\equiv \frac{\lambda_r}{1-\rho_s \phi}, \\
b_d^{lr} &\equiv \frac{\lambda_d}{1-\rho_s \phi},
\end{aligned} \tag{54}
$$

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

$$r_{t+1}^s = \pi_r + \lambda_r dp_t + \varepsilon_{t+1}^r, \tag{55}$$

$$\Delta d_{t+1} = \pi_d + \lambda_d dp_t + \varepsilon_{t+1}^d, \tag{56}$$

$$dp_{t+1} = \pi_{dp} + \phi dp_t + \varepsilon_{t+1}^{dp}. \tag{57}$$

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

$$\begin{pmatrix} r_{t+1}^s \\ \Delta d_{t+1} \\ dp_{t+1} \end{pmatrix} = \begin{pmatrix} \pi_r \\ \pi_d \\ \pi_{dp} \end{pmatrix} + \begin{pmatrix} \gamma_r & \pi_r & \lambda_r \\ \gamma_d & \pi_d & \lambda_d \\ \gamma_{dp} & \pi_{dp} & \phi \end{pmatrix} \begin{pmatrix} r_t^s \\ \Delta d_t \\ dp_t \end{pmatrix} + \begin{pmatrix} \varepsilon_{t+1}^r \\ \varepsilon_{t+1}^d \\ \varepsilon_{t+1}^{dp} \end{pmatrix}, \tag{58}$$

which is equivalent to

$$\mathbf{z}_{t+1} = \boldsymbol{\alpha} + \mathbf{A}\mathbf{z}_t + \boldsymbol{\varepsilon}_{t+1}, \tag{59}$$

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

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, $\boldsymbol{\Sigma} \equiv \mathrm{Cov}(\mathbf{z}_t, \mathbf{z}_t')$, we define the long-run variance decomposition for $dp$ associated with the unrestricted VAR:

$$
\begin{aligned}
1 &\approx b_r^{lr} - b_d^{lr}, \\
b_r^{lr} &\equiv \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I}-\rho_s\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{dp}}{\mathbf{e}_{dp}'\boldsymbol{\Sigma}\mathbf{e}_{dp}}, \\
b_d^{lr} &\equiv \frac{\mathbf{e}_d'\mathbf{A}(\mathbf{I}-\rho_s\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{dp}}{\mathbf{e}_{dp}'\boldsymbol{\Sigma}\mathbf{e}_{dp}}.
\end{aligned} \tag{60}
$$

To construct the series for $r^s$, $\Delta d$, 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 previous sections. The sample period is the same as in the previous sections, 1965–2018.

The descriptive statistics associated with $r^s$, $\Delta d$, and $dp$ are displayed in Table 4. 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 $\Delta d$ appears to be mean-revering, 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 existent evidence in the asset pricing literature. The estimate of $\rho_s$ is 0.98, which is clearly above the estimate for $\rho$.

The correlations among the three variables are presented in Panel B of Table 4. One noticeable feature is that those variables are much less mutually correlated than the investment variables, $r$, $\Delta d$, and $mq$. For example, the highest correlation, which occurs for $r^s$ and $\Delta d$, is only 0.43. Table 4 also displays the correlation between the stock market variables and the investment variables. The correlations are typically small in magnitude (below 0.30). The most relevant result is the negative correlation between $dp$ and $mq$ ($-0.43$). On the other hand, stock returns and investment returns are very weakly correlated (0.19). Further, profits growth and dividend growth show a weak negative association ($-0.22$).

The long-run variance decomposition results associated with $dp$ are presented in Table 5. 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 stock return predictability ($t$-ratios below one). 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).[24]

[24]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) or the S& P 500 index (e.g., Campbell and

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 shed some light on this 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 stock returns are more volatile than investment returns, while the opposite holds for dividend growth in relation to the growth in the marginal profitability of capital. This may arise because 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 (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 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.

# 9 Conclusion

This paper explores the sources of fluctuations in the ratio of the marginal profit of capital to the marginal value of capital (marginal Q). We employ a parsimonious model with a standard production 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 (2019), 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.

Shiller (1988) and Maio and Xu (2019)).

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. 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.

Our 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, our findings are consistent with the conjecture 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 ($\Delta m$), 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 |
| $\Delta m$ | 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.)** | | | | | |
| | $r$ | $\Delta m$ | $mq$ | $m$ | $q$ |
| $r$ | 1.00 | 0.92 | 0.63 | 0.64 | 0.59 |
| $\Delta m$ | | 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: VAR Estimates

This table reports the 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 ($\Delta m$), and $mq$. $\lambda, \phi$ denote the VAR slopes associated with lagged $mq$, 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, in %. $b^{lr}$ 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 sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

| | $\lambda, \phi$ | $t$ | $R^2$ | $b^{lr}$ | $t(b_r^{lr} = 0)$ | $t(b_r^{lr} = 1)$ |
|---|---|---|---|---|---|---|
| $r$ | −0.14 | *−1.77* | 0.04 | −0.24 | <u>−2.55</u> | **−13.21** |
| $\Delta m$ | −0.75 | **−3.75** | 0.22 | −1.24 | **−2.62** | **−13.39** |
| $mq$ | 0.46 | **3.19** | 0.21 | | | |

Table 3: 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 ($\Delta m$), and $mq$ are based on different pairs of the calibrated structural parameters $\alpha$ 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 sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

| | **Rest. VAR** | | | | **Unrest. VAR** | | | |
|---|---|---|---|---|---|---|---|---|
| | $b^{lr}_r$ | $t$ | $b^{lr}_m$ | $t$ | $b^{lr}_r$ | $t$ | $b^{lr}_m$ | $t$ |
| $\alpha = 0.05, a = 1.37$ | 0.08 | (0.63) | −0.92 | (−**7.16**) | −0.01 | (−0.05) | −1.01 | (−**7.00**) |
| $\alpha = 0.05, a = 6.85$ | −0.17 | ($\underline{-2.09}$) | −1.17 | (−**14.80**) | −0.38 | (−**4.47**) | −1.39 | (−**17.50**) |
| $\alpha = 0.05, a = 13.69$ | −0.24 | (−**3.20**) | −1.25 | (−**16.84**) | −0.48 | (−**6.03**) | −1.50 | (−**20.17**) |
| $\alpha = 0.05, a = 20.53$ | −0.27 | (−**3.62**) | −1.28 | (−**17.27**) | −0.53 | (−**6.57**) | −1.54 | (−**20.70**) |
| $\alpha = 0.05, a = 27.38$ | −0.29 | (−**3.84**) | −1.30 | (−**17.38**) | −0.56 | (−**6.83**) | −1.57 | (−**20.84**) |
| $\alpha = 0.15, a = 1.37$ | 0.57 | (**2.60**) | −0.42 | (*−1.87*) | 0.57 | (**2.87**) | −0.42 | ($\underline{-2.10}$) |
| $\alpha = 0.15, a = 6.85$ | 0.14 | (0.49) | −0.86 | (−**3.03**) | −0.02 | (−0.07) | −1.02 | (−**3.10**) |
| $\alpha = 0.15, a = 13.69$ | −0.12 | (−0.59) | −1.12 | (−**5.64**) | −0.41 | (*−1.79*) | −1.42 | (−**6.24**) |
| $\alpha = 0.15, a = 20.53$ | −0.22 | (−1.44) | −1.22 | (−**8.09**) | −0.55 | (−**3.24**) | −1.56 | (−**9.49**) |
| $\alpha = 0.15, a = 27.38$ | −0.26 | ($\underline{-2.07}$) | −1.27 | (−**9.97**) | −0.60 | (−**4.31**) | −1.61 | (−**11.99**) |
| $\alpha = 0.30, a = 1.37$ | 0.76 | (**4.43**) | −0.22 | (−1.26) | 0.73 | (**4.94**) | −0.25 | (−1.64) |
| $\alpha = 0.30, a = 6.85$ | 0.63 | (*1.77*) | −0.36 | (−1.00) | 0.56 | (*1.69*) | −0.42 | (−1.26) |
| $\alpha = 0.30, a = 13.69$ | 0.31 | (0.73) | −0.69 | (−1.63) | 0.15 | (0.30) | −0.85 | (*−1.77*) |
| $\alpha = 0.30, a = 20.53$ | 0.06 | (0.18) | −0.93 | (−**2.62**) | −0.21 | (−0.48) | −1.21 | (−**2.77**) |
| $\alpha = 0.30, a = 27.38$ | −0.08 | (−0.29) | −1.08 | (−**3.79**) | −0.42 | (−1.22) | −1.43 | (−**4.12**) |
| $\alpha = 0.50, a = 1.37$ | 0.83 | (**6.11**) | −0.15 | (−1.08) | 0.80 | (**6.84**) | −0.18 | (−1.48) |
| $\alpha = 0.50, a = 6.85$ | 0.88 | (**3.27**) | −0.10 | (−0.39) | 0.80 | (**3.56**) | −0.18 | (−0.79) |
| $\alpha = 0.50, a = 13.69$ | 0.77 | (*1.79*) | −0.22 | (−0.51) | 0.67 | (*1.67*) | −0.32 | (−0.78) |
| $\alpha = 0.50, a = 20.53$ | 0.54 | (1.08) | −0.45 | (−0.88) | 0.41 | (0.77) | −0.58 | (−1.07) |
| $\alpha = 0.50, a = 27.38$ | 0.32 | (0.66) | −0.67 | (−1.35) | 0.12 | (0.21) | −0.87 | (−1.49) |
| $\alpha = 0.70, a = 1.37$ | 0.86 | (**7.30**) | −0.12 | (−1.01) | 0.84 | (**8.22**) | −0.15 | (−1.40) |
| $\alpha = 0.70, a = 6.85$ | 0.96 | (**4.52**) | −0.03 | (−0.12) | 0.88 | (**5.01**) | −0.10 | (−0.57) |
| $\alpha = 0.70, a = 13.69$ | 0.98 | (**2.89**) | −0.01 | (−0.02) | 0.87 | (**2.98**) | −0.12 | (−0.40) |
| $\alpha = 0.70, a = 20.53$ | 0.88 | (*1.90*) | −0.11 | (−0.24) | 0.76 | (*1.78*) | −0.23 | (−0.52) |
| $\alpha = 0.70, a = 27.38$ | 0.70 | (1.30) | −0.29 | (−0.54) | 0.58 | (1.06) | −0.41 | (−0.75) |

Table 4: Descriptive Statistics: Stock Market Variables

This table reports descriptive statistics for the log stock return ($r^S$), log dividend growth ($\Delta d$), 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 ($r$, $\Delta m$, and $mq$) are presented in Panel B.

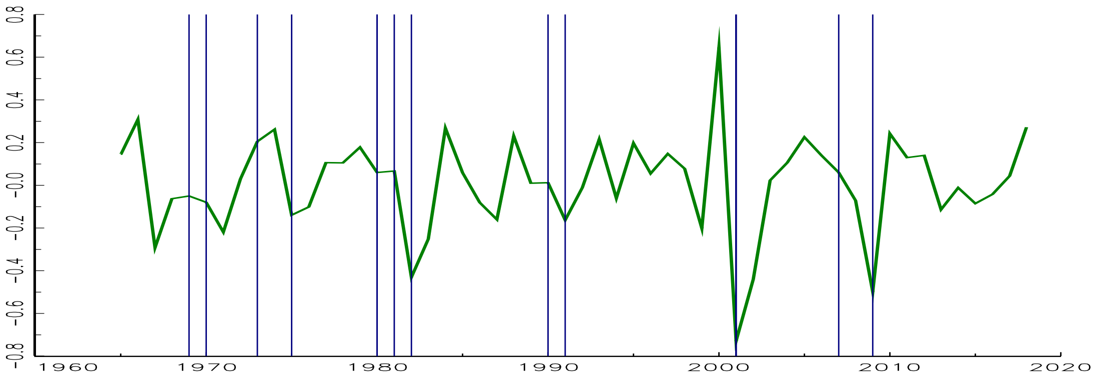
| | | | | | | |
|---|---|---|---|---|---|---|
| | | | **Panel A** | | | |
| | Mean | S.D. | Min. | Max. | AR(1) | |
| $r^S$ | 0.09 | 0.16 | −0.30 | 0.49 | −0.04 | |
| $\Delta d$ | 0.06 | 0.12 | −0.18 | 0.38 | −0.32 | |
| $dp$ | −3.68 | 0.45 | −4.80 | −2.91 | 0.95 | |
| | | | **Panel B (Correl.)** | |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

Table 5: Long-Run Variance Decomposition for $dp$

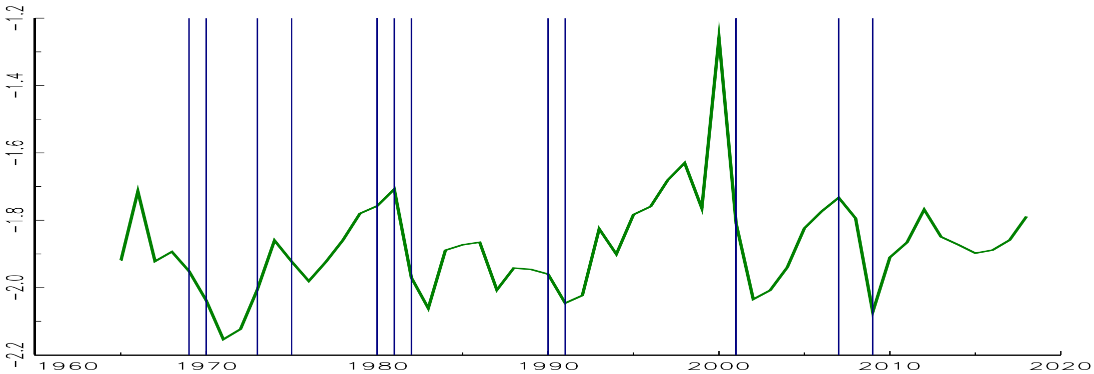
This table reports the long-run variance decomposition for the log dividend-to-price ratio ($dp$) based on a first-order VAR. The variables in the VAR are the log stock return ($r^S$), log dividend growth ($\Delta d$), and $dp$. $b^{lr}$ denote the long-run coefficients (infinite horizon). $t(b_r^{lr} = 0)$ 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. Panels A and B report the implied estimates based on a restricted and unrestricted VAR, respectively. The sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

| | $b^{lr}$ | $t(b_r^{lr} = 0)$ | $t(b_r^{lr} = 1)$ |
|---|---|---|---|
| | **Panel A (Rest. VAR)** | | |
| $r^S$ | 1.25 | <u>2.52</u> | 0.50 |
| $\Delta d$ | 0.25 | <u>2.49</u> | 0.49 |
| | **Panel B (Unrest. VAR)** | | |
| $r^S$ | 1.14 | **3.15** | 0.38 |
| $\Delta d$ | 0.14 | **3.10** | 0.37 |


Panel A ($r$)



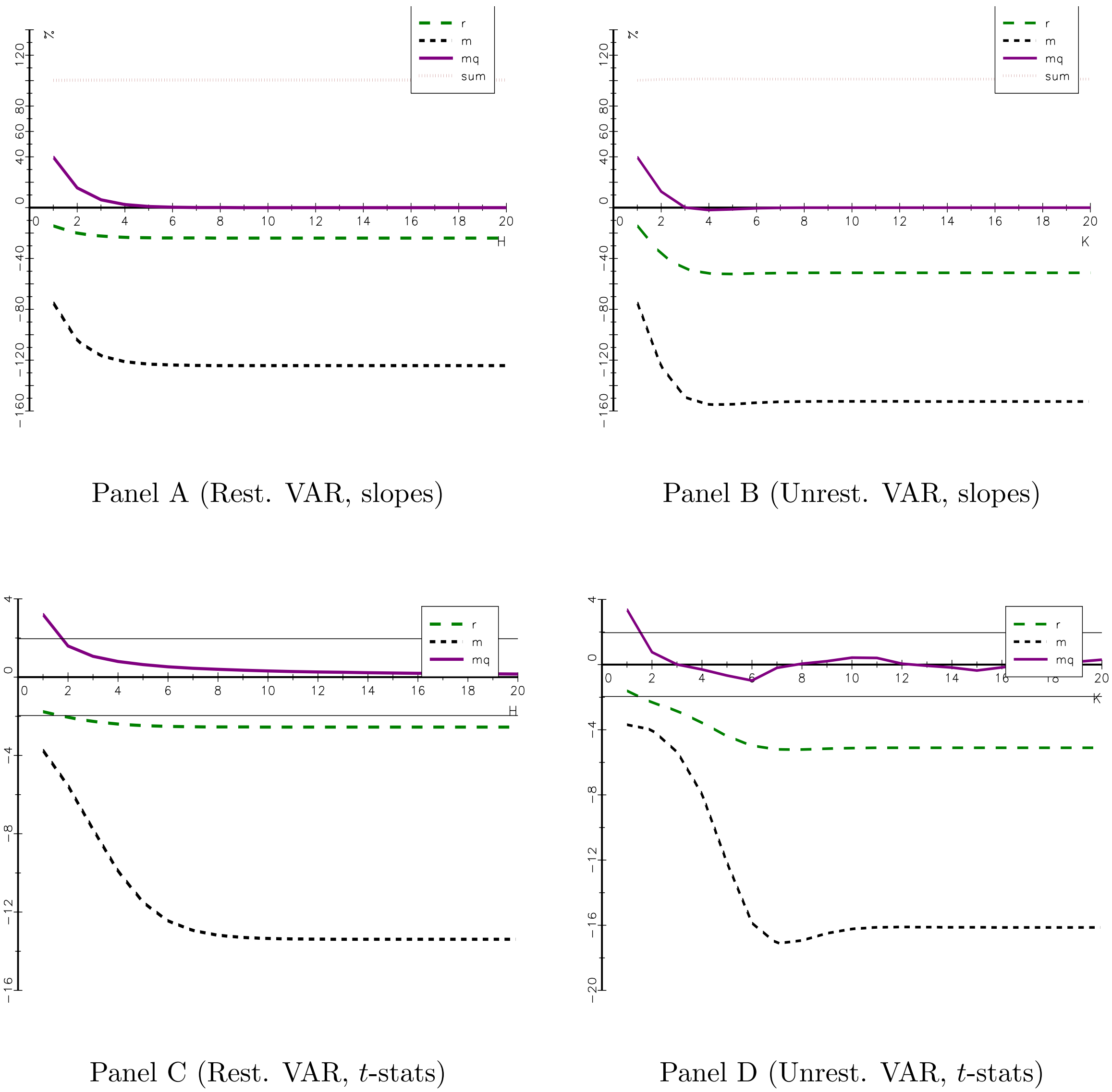

Panel B ($\Delta m$)




Panel C ($mq$)

Figure 1: Time-Series for $r$, $\Delta m$, and $mq$

This figure plots the time-series for the log investment return ($r$), log growth in marginal profits ($\Delta m$), 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.

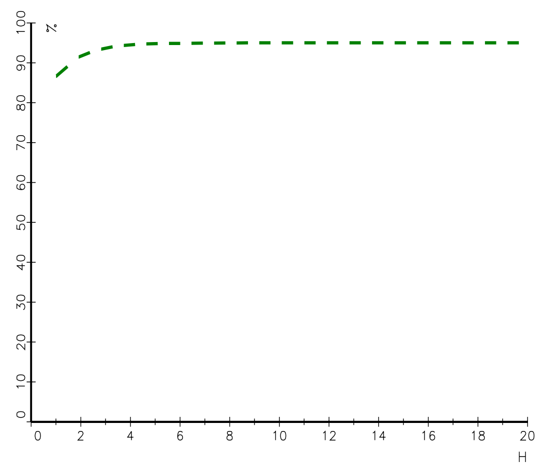



Panel A (Rest. VAR, slopes) Panel B (Unrest. VAR, slopes)

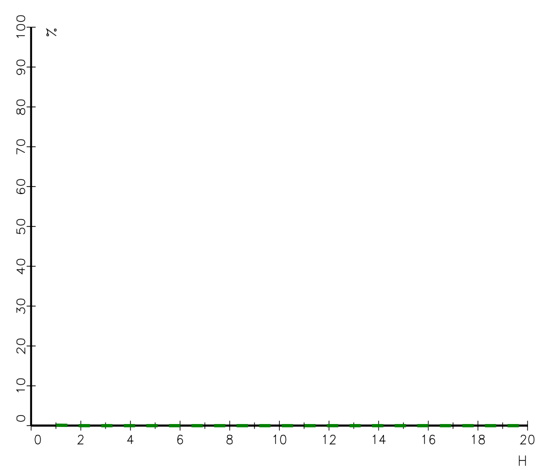
Panel C (Rest. VAR, $t$-stats) Panel D (Unrest. VAR, $t$-stats)

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 1964 to 2018.

# References

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

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.

Bandi, F.M., Perron, B., Tamoni, A., and Tebaldi, C., 2019. The scale of predictability, *Journal of Econometrics*, 208, 120–140.

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., 2009. On the reversal of return and dividend growth predictability: A tale of two periods, *Journal of Financial Economics*, 92, 128–151.

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.

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

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.

Fama, E.F. and French, K.R., 1995. Size and book-to-market factors in earnings and returns, *Journal of Finance*, 50, 131–155.

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.

Gonçalves, A.S., 2019. Can reinvestment risk explain the dividend and bond term structures?, *SSRN Working Paper*.

Gonçalves, A.S., Xue, C., and Zhang, L., 2019. Aggregation, capital heterogeneity, and the investment capm, *Review of Financial Studies*, forthcoming.

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.

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.

Lagarias, J.C., Reeds, J.A., Wright, M.H., and Wright, P.E., 1998. Convergence properties of the nelder-mead simplex method in low dimensions, *SIAM Journal on Optimization*, 9, 112–147.

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

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., 2019. Cash-flow or return predictability at long horizons? the case of earnings yield, *SSRN Working Paper*.

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.

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

Stigler, G.J., 1963. *Capital and Rates of Return in Manufacturing Industries*, Princeton University Press Princeton, NJ.

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.

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

# A Online Appendix: Not for Publication

## A.1 Estimating the Investment Return and its Components

### A.1.1 GMM Estimation Methodology

We estimate the parameters, $\mathbf{b} \equiv (\alpha, a)$, by minimizing a weighted average of the sample moments $e_i^r$ and $e_i^q$, denoted by $\mathbf{g}_T$. The GMM objective function is a weighted sum of squares of the model errors, that is, $\mathbf{g}_T'\mathbf{W}\mathbf{g}_T$, in which $\mathbf{W}$ is the (adjusted) identity matrix.[25] Following Belo *et al.* (2013), we adjust the weighting matrix $\mathbf{W}$ such that the two sets of errors, $e_i^r$ and $e_i^q$, are comparable in magnitude.[26] Let $\mathbf{D} = \partial \mathbf{g}_T / \partial \mathbf{b}$ and $\mathbf{S}$ equal a consistent estimate of the variance-covariance matrix of the sample errors $\mathbf{g}_T$. We estimate $\mathbf{S}$ using a standard Bartlett kernel with a window length of two.[27] The estimate of $\mathbf{b}$, denoted $\widehat{\mathbf{b}}$, is asymptotically normal with variance-covariance matrix given by

$$\text{var}(\widehat{\mathbf{b}}) = \frac{1}{T}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}. \tag{A.1}$$

To construct standard errors for the model errors, we use

$$\text{var}(\mathbf{g}_T) = \frac{1}{T}[\mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}]\mathbf{S}[\mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}], \tag{A.2}$$

which is the variance-covariance matrix for the model errors, $\mathbf{g}_T$.

To evaluate the investment Euler equation moment for each testing portfolio, Liu *et al.* (2009) aggregate firm-level accounting variables to portfolio-level variables, from which portfolio levered investment returns are constructed and matched with portfolio stock returns. Gonçalves *et al.* (2019) argue this approach has important drawbacks as follows. First, this approach assumes that all firms within a portfolio have the same investment returns. Second, the approach ignores substantial amount of heterogeneity across firms' accounting variables that can help identify structural parameters. We also note that non-linearities in the firm-

[25] As in Liu *et al.* (2009), we conduct a numerical optimization based on the simplex search method of Lagarias *et al.* (1998).

[26] We multiply the valuation moments by a factor of $|e_i^r| \,/\, |e_i^q|$, in which $|e_i^q|$ is the mean absolute valuation error from estimating only the valuation moments and $|e_i^r|$ is the mean absolute return error from estimating only the expected return moments across the same testing assets.

[27] The GMM estimation is robust to the choice of window length.

level production and adjustment costs functions imply that the estimated share of capital and adjustment cost parameters do not represent the true firm-level parameters. Intuitively, the aggregate production, for example, depends on the cross sectional distribution of capital. By using artificial data simulated from known firm-level parameters, Belo *et al.* (2019) show that the parameter estimates obtained from the aggregation procedure in Liu *et al.* (2009) and Belo *et al.* (2013) are subject to an aggregation bias, and hence do not have a structural interpretation. Thus, portfolio-level and aggregate investment returns implied by the aggregation approach can potentially be different from a value-weighted investment returns of individual firms. Existing studies of the sources of fluctuations in aggregate investment, for example Abel and Blanchard (1986), Chen *et al.* (2016), and Yashiv (2016), abstract from the issue of aggregation bias. To treat aggregation bias, we follow Gonçalves *et al.* (2019) and Belo *et al.* (2019) in the benchmark estimation, where firm-level accounting variables are used to construct firm-level levered investment returns, which are then aggregated to the portfolio level to match with portfolio stock returns.

Firm-level accounting variables and, thus, firm-level investment returns are subject to the issue of outliers. The outliers in firm-level investment returns can contaminate the aggregate portfolio-level investment returns and lead to noisy parameter estimates from the GMM estimation. To alleviate the impact of outliers in the firm-level GMM estimation, Gonçalves *et al.* (2019) construct firm-level investment returns using winsorized firm-level accounting variables, then compute equal- or value-weighted portfolio levered investment returns to match with equal- or value-weighted portfolio stock returns. Instead of winsorization, Belo *et al.* (2019) 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 levered 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. For robustness, we try both the winsorization approach and the median approach in our GMM estimations. In the winsorization approach, we winsorize firm-level accounting variables at the 1–99% level, then construct value-weighted portfolio levered investment returns to match with value-weighted portfolio stock returns.[28] In the median approach, we do not winsorize data but use portfolio

[28]Equal-weighted portfolio mean are still sensitive to outliers even after winsorizing firm-level data at

median to aggregate both firm-level investment returns and stock returns to the portfolio level.

For the valuation moment, we need to aggregate firm-level $Q$ to the portfolio level, which is similar to the task of aggregating firm-level investment returns to the portfolio level. Again, aggregation bias may be introduced if one aggregates firm-level accounting data to portfolio-level variables, from which portfolio $Q$ is constructed and matched with portfolio $Q$ in the data. To avoid aggregation bias, we construct firm-level $Q$ using firm-level accounting variables, then aggregate firm-level $Q$ to portfolio level. That is, the portfolio's $Q$ is a weighted average of the individual firms' $Q$ values. When aggregating firm-level $Q$ to portfolio level, we use each firm's total assets as weight and calculate weighted average portfolio $Q$ both in the model and in the data. This aggregation approach is equivalent to a valuation moment that matches the total portfolio firm value in the data with the total portfolio firm value predicted by the model.[29]

For robustness, we also conduct the GMM estimation following Liu *et al.* (2009), where firm-level accounting variables are aggregated to portfolio-level variables, from which portfolio-level levered investment returns are constructed to match with portfolio-level stock returns, either equal- or value-weighted. For the valuation moment, we follow Belo *et al.* (2013) and construct both the portfolio $Q$ in the data ($\widetilde{Q}$) and the portfolio $Q$ predicted by the model using portfolio-level accounting variables.

## A.2 Variance Decompositions

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

1-99% level, thus the results are not tabulated.

[29]At firm level, the market value of firm $i$ satisfies: $P_{i,t} + B_{i,t+1} = \left(1 + (1-\tau_t)a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right) K_{i,t+1}$, thus $\frac{P_{i,t}+B_{i,t+1}}{A_{i,t+1}} = \left(1 + (1-\tau_t)a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right)\frac{K_{i,t+1}}{A_{i,t+1}}$. Aggregating individual firms in portfolio $j$ using weights $\frac{A_{i,t+1}}{A_{j,t+1}}$, where $A_{j,t+1} = \sum_{i\in j} A_{i,t+1}$, we get $\sum_{i\in j}\left(\frac{A_{i,t+1}}{A_{j,t+1}}\frac{P_{i,t}+B_{i,t+1}}{A_{i,t+1}}\right) = \sum_{i\in j}\left[\left(1 + (1-\tau_t)a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right)\frac{K_{i,t+1}}{A_{i,t+1}}\frac{A_{i,t+1}}{A_{j,t+1}}\right]$. It can be simplified as $\sum_{i\in j}\left(P_{i,t} + B_{i,t+1}\right) = \sum_{i\in j}\left[\left(1 + (1-\tau_t)a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right) K_{i,t+1}\right]$.

### A.2.1 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 $mq_t$,

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

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

$$1 \approx -\beta\left(\sum_{h=1}^{H} \rho^{h-1}\Delta m_{t+h}, mq_t\right) + \beta\left(\sum_{h=1}^{H} \rho^{h-1} r_{t+h}, mq_t\right) + \beta\left(\rho^{H} mq_{t+H}, mq_t\right), \tag{A.4}$$

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:

$$1 \approx -\sum_{h=1}^{H} \rho^{h-1}\beta(\Delta m_{t+h}, mq_t) + \sum_{h=1}^{H} \rho^{h-1}\beta(r_{t+h}, mq_t) + \rho^{H}\beta(mq_{t+H}, mq_t). \tag{A.5}$$

Under the restricted first-order VAR, we have,

$$mq_{t+h-1} = \phi^{h-1} mq_t + \phi^{h-1}\sum_{l=1}^{h-1} \phi^{-l}(\pi_{mq} + \varepsilon_{t+l}^{mq}), \tag{A.6}$$

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

$$r_{t+h} = \pi_r + \lambda_r mq_{t+h-1} + \varepsilon_{t+h}^{r}, \tag{A.7}$$

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

$$r_{t+h} = \pi_r + \phi^{h-1}\lambda_r mq_t + \phi^{h-1}\lambda_r \sum_{l=1}^{h-1} \phi^{-l}(\pi_{mq} + \varepsilon_{t+l}^{mq}) + \varepsilon_{t+h}^{r}. \tag{A.8}$$

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

$$\beta(r_{t+h}, mq_t) = \phi^{h-1}\lambda_r. \tag{A.9}$$

Similarly, we have,

$$\beta(\Delta m_{t+h}, mq_t) = \phi^{h-1}\lambda_m. \tag{A.10}$$

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

$$mq_{t+H} = \phi^H mq_t + \phi^H \sum_{l=1}^{H} \phi^{-l}(\pi_{mq} + \varepsilon^{mq}_{t+l}), \tag{A.11}$$

we have

$$\beta(mq_{t+H}, mq_t) = \phi^H, \tag{A.12}$$

which leads to

$$1 \approx -\sum_{h=1}^{H} \rho^{h-1}\phi^{h-1}\lambda_m + \sum_{h=1}^{H} \rho^{h-1}\phi^{h-1}\lambda_r + \rho^H\phi^H. \tag{A.13}$$

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

$$\begin{aligned} 1 &\approx -b^H_m + b^H_r + b^H_{mq}, \\ b^H_m &\equiv \frac{\lambda_m(1-\rho^H\phi^H)}{1-\rho\phi}, \\ b^H_r &\equiv \frac{\lambda_r(1-\rho^H\phi^H)}{1-\rho\phi}, \\ b^H_{mq} &\equiv \rho^H\phi^H. \end{aligned} \tag{A.14}$$

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

$$\text{var}\left(\mathbf{b}^H\right) = \frac{\partial \mathbf{b}^H}{\partial \mathbf{b}'}\,\text{var}\left(\mathbf{b}\right)\frac{\partial \mathbf{b}^H}{\partial \mathbf{b}}. \tag{A.15}$$

The matrix of derivatives is given by

$$\frac{\partial \mathbf{b}^H}{\partial \mathbf{b}'} \equiv \begin{bmatrix} \frac{\partial b_m^H}{\partial \lambda_m} & \frac{\partial b_m^H}{\partial \lambda_r} & \frac{\partial b_m^H}{\partial \phi} \\ \frac{\partial b_r^H}{\partial \lambda_m} & \frac{\partial b_r^H}{\partial \lambda_r} & \frac{\partial b_r^H}{\partial \phi} \\ \frac{\partial b_{mq}^H}{\partial \lambda_m} & \frac{\partial b_{mq}^H}{\partial \lambda_r} & \frac{\partial b_{mq}^H}{\partial \phi} \end{bmatrix} = \begin{bmatrix} \frac{1-\rho^H\phi^H}{1-\rho\phi} & 0 & \frac{-H\lambda_m\rho^H\phi^{H-1}(1-\rho\phi)+\rho\lambda_m(1-\rho^H\phi^H)}{(1-\rho\phi)^2} \\ 0 & \frac{1-\rho^H\phi^H}{1-\rho\phi} & \frac{-H\lambda_r\rho^H\phi^{H-1}(1-\rho\phi)+\rho\lambda_r(1-\rho^H\phi^H)}{(1-\rho\phi)^2} \\ 0 & 0 & H\rho^H\phi^{H-1} \end{bmatrix}. \tag{A.16}$$

### A.2.2 Unrestricted VAR(1)

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

$$\mathbf{z}_{t+h} = (\mathbf{I} + \mathbf{A} + ... + \mathbf{A}^{h-1})\boldsymbol{\pi} + \mathbf{A}^h\mathbf{z}_t + \mathbf{A}^{h-1}\boldsymbol{\varepsilon}_{t+1} + ... + \mathbf{A}\boldsymbol{\varepsilon}_{t+h-1} + \boldsymbol{\varepsilon}_{t+h}, \tag{A.17}$$

or equivalently,

$$\mathbf{z}_{t+h} = \mathbf{A}^h\mathbf{z}_t + \mathbf{A}^h\sum_{l=1}^{h}\mathbf{A}^{-l}(\boldsymbol{\pi} + \boldsymbol{\varepsilon}_{t+l}). \tag{A.18}$$

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

$$\beta(r_{t+h}, mq_t) = \frac{\mathrm{Cov}(\mathbf{e}_r'\mathbf{z}_{t+h}, \mathbf{e}_{mq}'\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}_{mq}'\mathbf{z}_t)} = \frac{\mathrm{Cov}(\mathbf{e}_r'\mathbf{A}^h\mathbf{z}_t, \mathbf{e}_{mq}'\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}_{mq}'\mathbf{z}_t)} = \frac{\mathbf{e}_r'\mathbf{A}^h\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}}, \tag{A.19}$$

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

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

$$\begin{aligned} \sum_{h=1}^{H}\rho^{h-1}\beta(r_{t+h}, mq_t) &= \sum_{h=1}^{H}\frac{\mathbf{e}_r'\rho^{h-1}\mathbf{A}^h\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}} \\ &= \frac{\mathbf{e}_r'}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}}\left(\sum_{h=1}^{H}\rho^{h-1}\mathbf{A}^h\right)\boldsymbol{\Sigma}\mathbf{e}_{mq} \\ &= \frac{\mathbf{e}_r'}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}\rho}\left(\sum_{h=1}^{H}\rho^{h}\mathbf{A}^h\right)\boldsymbol{\Sigma}\mathbf{e}_{mq} \\ &= \frac{\mathbf{e}_r'(\rho\mathbf{A} - \rho^{H+1}\mathbf{A}^{H+1})(\mathbf{I} - \rho\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\rho\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}} \\ &= \frac{\mathbf{e}_r'\mathbf{A}(\mathbf{I} - \rho^H\mathbf{A}^H)(\mathbf{I} - \rho\mathbf{A})^{-1}\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}_{mq}'\boldsymbol{\Sigma}\mathbf{e}_{mq}}. \end{aligned} \tag{A.20}$$

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

$$\beta(mq_{t+H}, mq_t) = \frac{\mathrm{Cov}(\mathbf{e}'_{mq}\mathbf{z}_{t+H}, \mathbf{e}'_{mq}\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}'_{mq}\mathbf{z}_t)} = \frac{\mathrm{Cov}(\mathbf{e}'_{mq}\mathbf{A}^H\mathbf{z}_t, \mathbf{e}'_{mq}\mathbf{z}_t)}{\mathrm{var}(\mathbf{e}'_{mq}\mathbf{z}_t)} = \frac{\mathbf{e}'_{mq}\mathbf{A}^H\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}'_{mq}\boldsymbol{\Sigma}\mathbf{e}_{mq}}, \quad \text{(A.21)}$$

which implies that

$$\rho^H\beta(mq_{t+H}, mq_t) = \frac{\rho^H\mathbf{e}'_{mq}\mathbf{A}^H\boldsymbol{\Sigma}\mathbf{e}_{mq}}{\mathbf{e}'_{mq}\boldsymbol{\Sigma}\mathbf{e}_{mq}}. \quad \text{(A.22)}$$

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,

$$\mathrm{var}\left(\mathbf{b}^H\right) = \frac{\partial\mathbf{b}^H}{\partial\mathbf{b}'}\,\mathrm{var}\left(\mathbf{b}\right)\frac{\partial\mathbf{b}^H}{\partial\mathbf{b}}, \quad \text{(A.23)}$$

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.[30]

## A.3 Bootstrap Simulation

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

1. We estimate the first-order restricted VAR,

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

   and save the time-series of residuals ($\varepsilon_{t+1}^r$, $\varepsilon_{t+1}^m$, and $\varepsilon_{t+1}^{mq}$), as well as the estimates of $\phi$ and $\rho$.

2. In each replication ($u = 1, ..., 10,000$), we construct pseudo VAR innovations by draw-

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

ing with replacement from the original VAR residuals:

$$(\varepsilon_t^{r,u}, \varepsilon_t^{m,u}, \varepsilon_t^{mq,u})', t = v_1^u, ..., v_T^u,$$

where the time indices $v_1^u, ..., v_T^u$—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),

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

for $\Delta m$ (no-profit predictability null),

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

and for $mq$:

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

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),

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

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

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

$$
\begin{aligned}
p(b^H_r) &= \#\left\{b^H_{r,u} > b^H_r\right\}/10000, \\
p(b^H_m) &= \#\left\{b^H_{m,u} < b^H_m\right\}/10000,
\end{aligned}
$$

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

## A.4 Sensitivity Analysis

Below we present the results associated with Section 6 in the paper. Figure A.1 corresponds to Subsection 6.1. The results discussed in Subsection 6.2 are presented in Table A.1 and in Figure A.2 (in what concerns the variables obtained from the structural estimation with Q deciles) and in Table A.2 and in Figure A.3 (in what concerns the variables obtained from structural estimation for the median firm). Table A.3 and Figure A.4 correspond to Subsection 6.3. Finally, the results discussed in Subsection 6.4 are displayed in in Table A.4 and in Figure A.5.

## Table A.1: VAR Estimates: Q Deciles

This table reports the 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 ($\Delta m$), and $mq$. $\lambda, \phi$ denote the VAR slopes associated with lagged $mq$, 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, in %. $b^{lr}$ 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 sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively. All the investment series are obtained from a GMM estimation of the structural investment model based on ten portfolios sorted on Q.

| | $\lambda, \phi$ | $t$ | $R^2$ | $b^{lr}$ | $t(b_r^{lr} = 0)$ | $t(b_r^{lr} = 1)$ |
|---|---|---|---|---|---|---|
| $r$ | −0.17 | $\underline{-2.06}$ | 0.05 | −0.27 | **−3.06** | **−14.48** |
| $\Delta m$ | −0.78 | **−4.06** | 0.23 | −1.27 | **−3.15** | **−14.70** |
| $mq$ | 0.45 | **3.22** | 0.20 | | | |

## Table A.2: VAR Estimates: Median Firm

This table reports the 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 ($\Delta m$), and $mq$. $\lambda, \phi$ denote the VAR slopes associated with lagged $mq$, 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, in %. $b^{lr}$ 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 sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively. All the investment series are associated with the median firm.

| | $\lambda, \phi$ | $t$ | $R^2$ | $b^{lr}$ | $t(b_r^{lr} = 0)$ | $t(b_r^{lr} = 1)$ |
|---|---|---|---|---|---|---|
| $r$ | −0.15 | *−1.95* | 0.05 | −0.25 | **−2.89** | **−14.69** |
| $\Delta m$ | −0.76 | **−3.96** | 0.23 | −1.25 | **−2.98** | **−14.94** |
| $mq$ | 0.45 | **3.22** | 0.21 | | | |

## Table A.3: VAR Estimates: VAR(2)

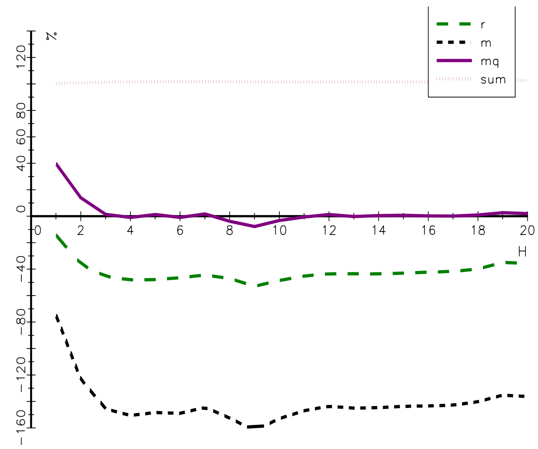
This table reports the VAR(2) estimation results when the predictor is the log profits-to-Q ratio ($mq$). The variables in the VAR are the log investment return ($r$), log growth in marginal profits ($\Delta m$), and $mq$. $\lambda_1, \phi_1$ and $\lambda_2, \phi_2$ denote the VAR slopes associated with $mq_t$ and $mq_{t-1}$, respectively, while $t$ denotes the respective Newey and West (1987) $t$-statistics (calculated with two lags). $R^2$ is the coefficient of determination for each equation in the VAR, in %. $b^{lr}$ 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 sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively.

| | $\lambda_1, \phi_1$ | $t$ | $\lambda_2, \phi_2$ | $t$ | $R^2$ | $b^{lr}$ | $t(b_r^{lr} = 0)$ | $t(b_r^{lr} = 1)$ |
|---|---|---|---|---|---|---|---|---|
| $r$ | −0.03 | −0.29 | −0.30 | $\underline{-2.30}$ | 0.13 | −0.38 | **−4.03** | **−14.51** |
| $\Delta m$ | −0.61 | **−2.71** | −0.28 | −1.14 | 0.25 | −1.39 | **−4.25** | **−15.05** |
| $mq$ | 0.48 | **3.18** | −0.03 | −0.19 | 0.23 | | | |

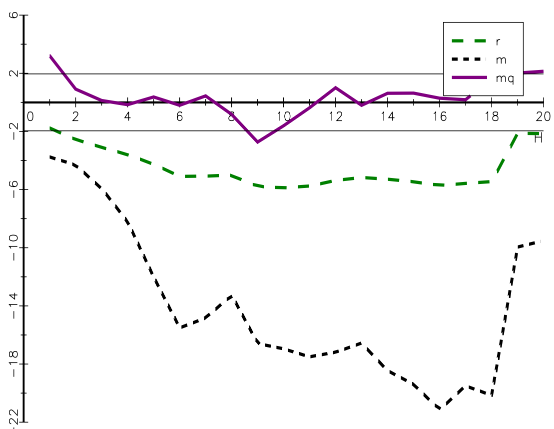
## Table A.4: VAR Estimates: Alternative Investment Series

This table reports the 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 ($\Delta m$), and $mq$. $\lambda, \phi$ denote the VAR slopes associated with lagged $mq$, 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, in %. $b^{lr}$ 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 sample is 1965–2018. Italic, underlined, and bold numbers denote statistical significance at the 10%, 5%, and 1% levels, respectively. All the investment series are obtained by employing the Liu–Whited–Zhang method.

| | $\lambda, \phi$ | $t$ | $R^2$ | $b^{lr}$ | $t(b_r^{lr} = 0)$ | $t(b_r^{lr} = 1)$ |
|---|---|---|---|---|---|---|
| $r$ | 0.22 | 1.51 | 0.04 | 0.74 | 1.21 | −0.42 |
| $\Delta m$ | −0.08 | −0.39 | 0.00 | −0.26 | 1.21 | −0.42 |
| $mq$ | 0.85 | **11.78** | 0.72 | | | |



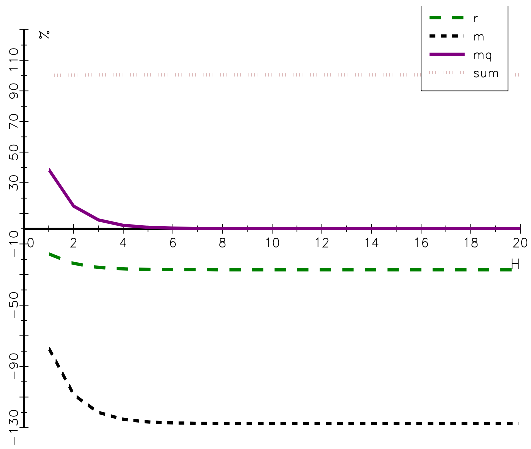

Panel A (slopes)



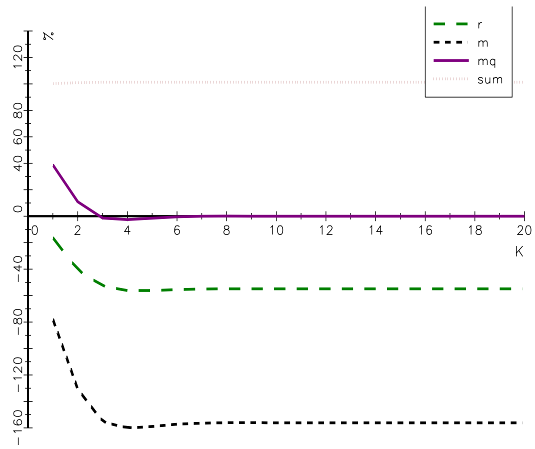

Panel B ($t$-stats)

# Figure A.1: Variance Decomposition: Direct Approach

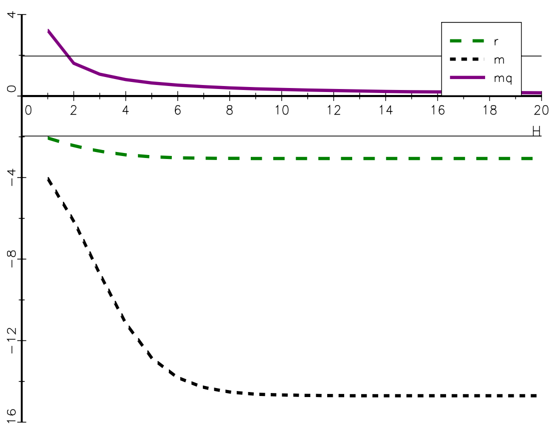
This figure plots the term structure of multiple-horizon predictive coefficients (in %), and respective $t$-statistics, corresponding to the variance decompositions for the log profits-to-Q ratio ($mq$). The predictive slopes are obtained from weighted long-horizon regressions. The coefficients are associated 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.



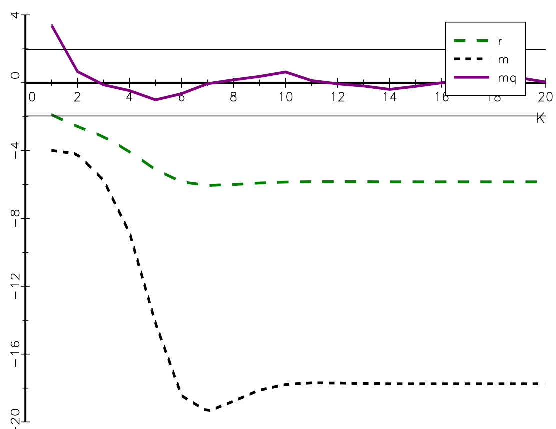

Panel A (Rest. VAR, slopes)




Panel B (Unrest. VAR, slopes)



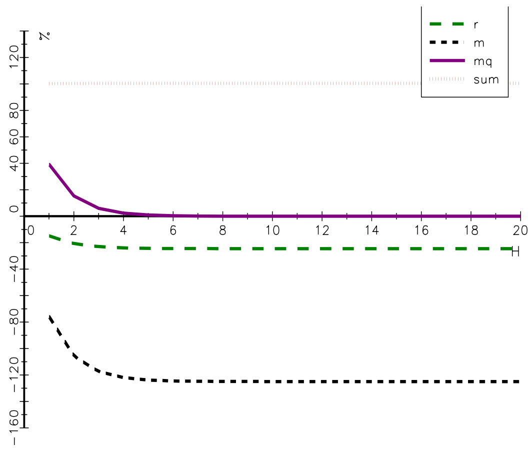

Panel C (Rest. VAR, $t$-stats)



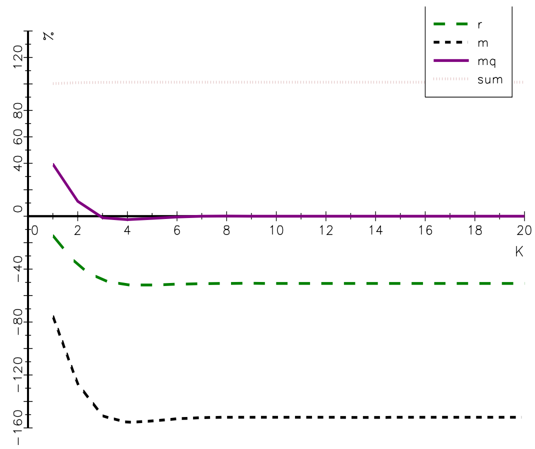

Panel D (Unrest. VAR, $t$-stats)

Figure A.2: Variance Decomposition: Q Deciles

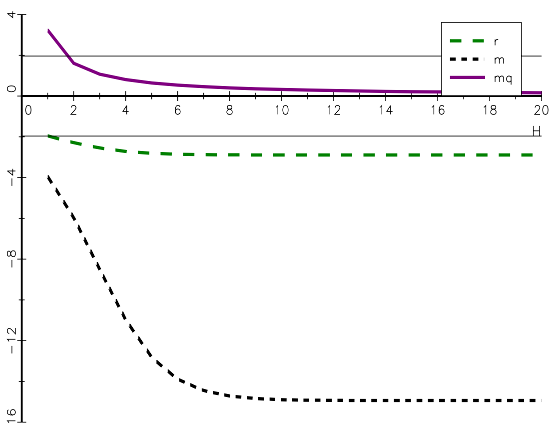
This figure plots the term structure of multiple-horizon predictive coefficients (in %), and respective $t$-statistics, corresponding to the variance decompositions for the log profits-to-Q ratio ($mq$). The predictive slopes are obtained from either a restricted or an unrestricted first-order VAR. The 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. All the investment series are obtained from a GMM estimation of the structural investment model based on ten portfolios sorted on Q.



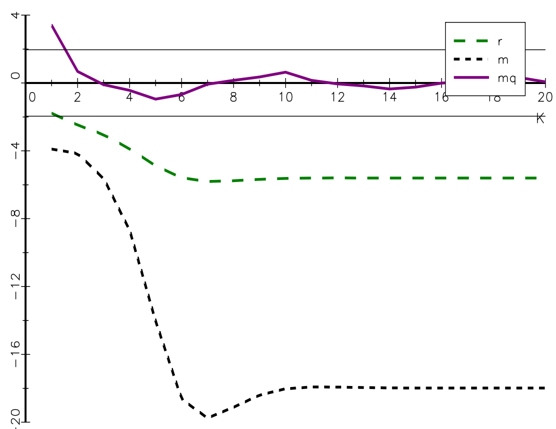

Panel A (Rest. VAR, slopes)




Panel B (Unrest. VAR, slopes)



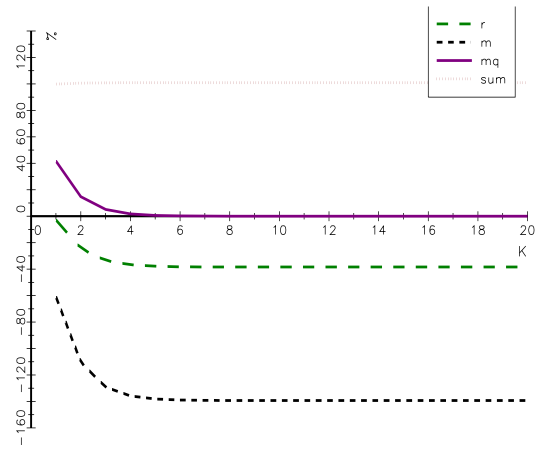

Panel C (Rest. VAR, $t$-stats)



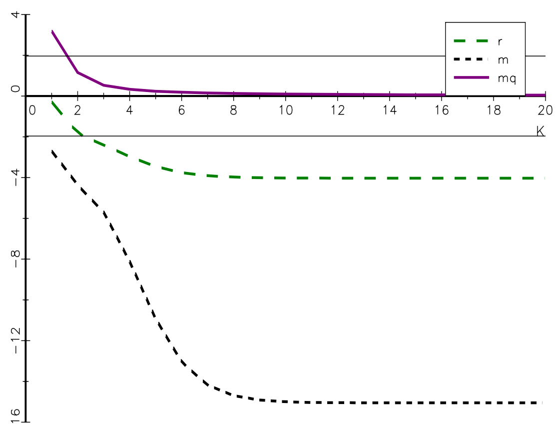

Panel D (Unrest. VAR, $t$-stats)

## Figure A.3: Variance Decomposition: Median firm

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. All the investment series are associated with the median firm.



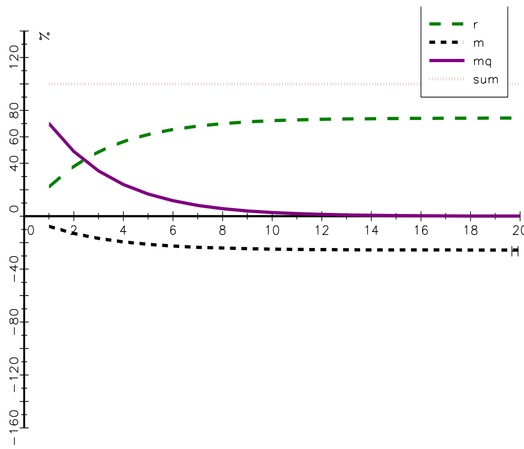

Panel A (slopes)



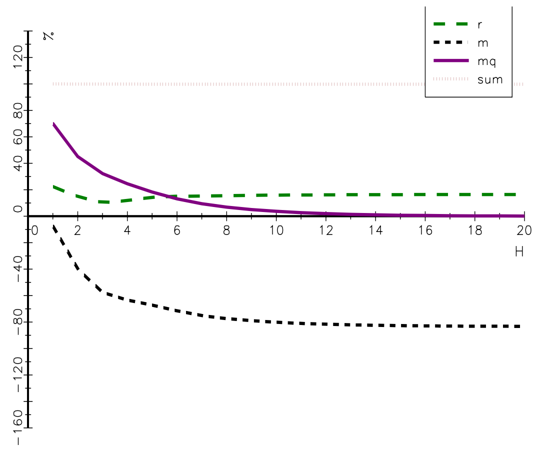

Panel B ($t$-stats)

# Figure A.4: Variance Decomposition: Restricted VAR(2)

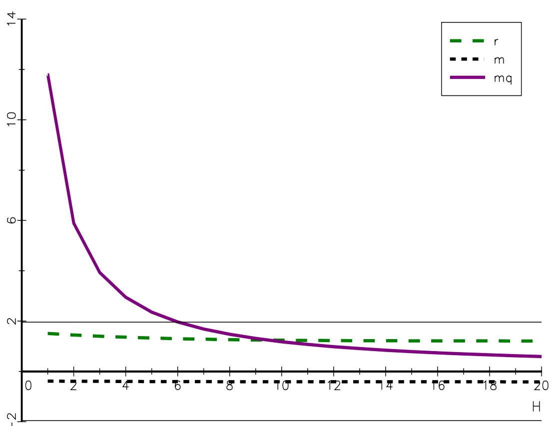
This figure plots the term structure of multiple-horizon predictive coefficients (in %), and respective $t$-statistics, corresponding to the variance decompositions for the log profits-to-Q ratio ($mq$). The predictive slopes are obtained from a restricted second-order VAR. The coefficients are associated 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.



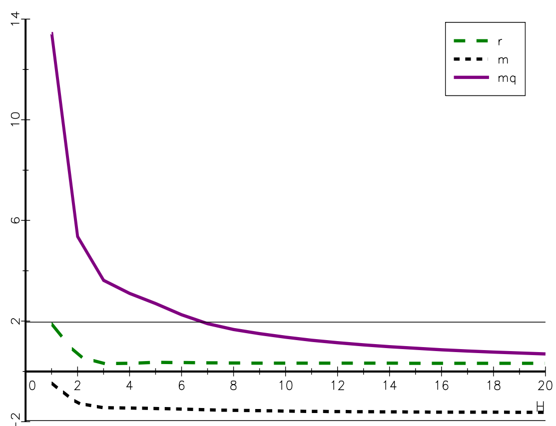

Panel A (Rest. VAR, slopes)




Panel B (Unrest. VAR, slopes)




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




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

Figure A.5: Variance Decomposition: Alternative Investment Series

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. All the investment series are obtained by employing the Liu–Whited–Zhang method.