Why Did The q Theory Of Investment Start Working?*

Daniel Andrei
UCLA

William Mann
UCLA

Nathalie Moyen
Colorado

November 19, 2017

Abstract

We show that the relationship between aggregate investment and Tobin's q has become remarkably tight in recent years, contrasting with earlier eras. We attribute this change to the growing empirical dispersion in Tobin's q, which we document in both in the cross-section and the time-series. To study the source of this dispersion, we augment a standard investment model with learning. Information acquisition endogenously amplifies volatility in the firm's value function. Perhaps counterintuitively, the investment-q regression works better for research-intensive industries, a growing segment of the economy, despite their lower level of tangible capital. We confirm the model's predictions in the data.

*Andrei: daniel.andrei@anderson.ucla.edu; Mann: william.mann@anderson.ucla.edu; Moyen: nathalie.moyen@colorado.edu. We thank Andrew Abel for helpful comments on an early version of this paper.

1 Introduction

The q theory of investment predicts a strong relationship between corporations' market values and their investment rates. Hayashi (1982) provides justification for measuring Tobin's marginal q with a valuation ratio, average q, so that a simple regression of investment on q should have a strong fit. Researchers have found that this regression in fact performs quite poorly. While the Hayashi model assumptions may not exactly hold in the data, a stark disconnect between investment and valuation is deeply puzzling to financial economists. A large literature investigates the reasons why Tobin's q does not work well in the data, pointing to the existence of financial constraints, decreasing returns to scale, inefficient equity-market valuations, and measurement problems, among other things.[1]

Curiously, even as this literature has continued to grow, the stylized fact has changed. Using data from the NIPA tables combined with the Fed Flow of Funds, we document that the aggregate investment-q regression has worked remarkably well in recent years. The simple regression achieves an R^2 of 70% during 2000–2015, comparable to the empirical performances of the bond price q regression proposed in Philippon (2009) and the total tangible and intangible asset q regression in Peters and Taylor (2016). If one were to test the simple theory using data from recent years, one would conclude that the q theory of investment is in fact an empirical success.

Yet this recent development only deepens the puzzle, as problems with q theory highlighted by the literature seem to have *worsened* in recent years. For example, Peters and Taylor (2016) focuses on the failure to measure intangible assets, which have grown substantially in the aggregate, and Philippon (2009) focuses on excess volatility in equity market valuations. We show that, counterintuitively, it is precisely this higher volatility in valuations, *especially* in intangibles-intensive industries, that has revived the empirical performance of the classic regression. To explain the puzzle, we develop a learning-based model of corporate investment. The model endogenously produces more variation in marginal q. Intuitively, the investment-q regression works better when there is more endogenous variation in the regressor q. This provides a simple, yet previously unexplored, explanation behind the poor fit of the regression. The culprit is the historically low variation in Tobin's q relative to residual factors

[1] For examples, see Fazzari, Hubbard, and Petersen (1988), Kaplan and Zingales (1997), Erickson and Whited (2000), Gomes (2001), Cooper and Ejarque (2003) Philippon (2009), and Peters and Taylor (2016), among many others.

affecting investment.

To motivate the intuition empirically, we establish several stylized facts. First, the volatility of aggregate q in the data is higher during precisely the years when the aggregate investment-q regression performs better. Second, the between- and within-firm variation of Tobin's q in Compustat have both risen steeply since the late 1990s. Third, the within-firm variation in Tobin's q varies by orders of magnitude over the time series, reflecting a changing composition of firms in Compustat. Finally, the panel version of the investment-q regression also fits much better when Tobin's q is more volatile. These stylized facts support our intuition: the empirical performance of the theory hinges critically on the amount of endogenous variation that one finds in Tobin's q.

Turning to the model, we study a standard q-theoretic investment framework, most closely resembling the model of Abel (2017). Our main innovation is to embed learning about uncertain cash flows. We allow the firm to acquire, at a cost, informative signals about the time-varying cash flow mean. These features provide a theoretical foundation for the stochastic variation in marginal q, which is typically considered exogenous in prior papers. We show that learning endogenously amplifies the volatility of marginal q, thereby also improving the fit of the investment-q regression.

The empirical implication is that firms investing more in learning—in the form of research—should feature a tighter fit between investment and Tobin's q. At first glance, this prediction seems counterintuitive because research creates an intangible asset, and therefore a measurement error when accounting only for tangible capital in Tobin's q as discussed in Peters and Taylor (2016). Our model abstracts from this measurement error, and our empirical findings point to a large offsetting effect.

In the cross-section of firms in Compustat, industries featuring greater investment in research and development, higher rates of patenting, and greater intangibility feature noticeably higher R^2 values compared to the average industry in the simple investment-q panel regression. Although this stylized fact is documented in Peters and Taylor (2016) Section 5.2, it is earmarked as a puzzle. Our model provides an explanation, by predicting that research-intensive firms exhibit greater volatility in Tobin's q. We confirm that the better fit in high-tech industries was present even before the aggregate regression fit began to improve, so it is not driven simply by the fact that these firms are more common later in the sample. As high-tech firms have become a larger segment of the economy, their greater endogenous volatility in Tobin's q has caused the aggregate regression to improve.

We investigate other predictions of the learning model. The model predicts that the investment-q regression works better in settings where Tobin's q is less correlated with cash flow. With learning, q becomes less responsive to cash flow because the firm chooses to pay more attention to other signals. The learning mechanism works in the opposite direction as misspecification issues, which have been the focus in much of the prior research. Consider a misspecified regression by omitting cash flow when an alternative theory (e.g., based on financial constraints) would predict that cash flow is an important variable. With such an omitted variable bias, the investment-q regression should work better in settings where Tobin's q is *more*, not less, correlated with cash flow, as the bias shrinks when q and cash flow are more highly correlated.

We test the relative importance of the learning mechanism against potential misspecification. To implement this test, we sort industries into quartiles based on their estimated correlations of Tobin's q with cash flow in Compustat data, and compare across these quartiles the R^2 from the investment-q regression. The pattern strongly supports the learning mechanism. The lowest quartile of correlation between q and cash flow features an R^2 of 18%, compared to 10% for the highest quartile. While this does not indicate that there is no misspecification in our model, it does suggest that the empirical effects of this misspecification are outweighed by the learning mechanism. Intuitively, when q is more volatile, it is highly correlated with investment but not with cash flow.

In another prediction of the model, we investigate the effect of decreasing interest rates over time on the empirical performance of q theory. In a low-interest-rate environment, firm valuations become more volatile relative to the model's residuals, suggesting a better empirical performance for the investment-q regression. The aggregate data support this prediction. The high-yield era of 1975-1995 coincides with the worst empirical performance of the aggregate investment-q regression.

The most closely related empirical paper to ours is Peters and Taylor (2016), who augment the simple investment-q regression by adjusting for intangible capital. Another related empirical paper is Gutiérrez and Philippon (2016), who highlight that aggregate investment has trended downward while aggregate Tobin's q has trended upward, a divergence they attribute to weakened competition and governance in the US. Our analysis is mostly silent on the levels of investment and q, and focuses instead on their correlations, which have improved in recent years. Finally, our paper builds on a long theoretical literature investigating the q theory of investment. The most closely related theory paper to ours is Abel (2017), from which we add a learning mechanism. Related papers that analyze learning in corporate investment frameworks are Alti (2003) and Moyen and Platikanov (2012).

The rest of the paper is organized as follows: Section 2 establishes the motivating empirical facts related to the empirical dispersion in Tobin's q and the fit of the investment-q regression. Section 3 builds an investment model with learning that endogenizes volatility in q and derives testable implications. Section 4 returns to the data and investigates the implications of the model. Section 5 concludes.

2 Stylized empirical facts

2.1 Improved fit of the aggregate regression

We first document that the aggregate investment-q regression has performed much better in recent years than in the past. Figure 1 plots and compares aggregate investment and lagged aggregate Tobin's q from 1955 to 2015. To construct the series, we use quarterly data from the Fed Flow of Funds and from NIPA tables, following Hall (2001) and Philippon (2009). See the Appendix for details.

[Figure 1 here]

The figure is divided into 20-year intervals. At the bottom of each section is the R^2 value that would be obtained from the simple regression of the investment rate on lagged q. If one had performed the aggregate investment-q regression using the data available from 1955–1975, one would conclude that the theory has moderate explanatory power, achieving an R^2 value of 32%, though much is left unexplained.

In the second interval, 1975-1995, the performance of the regression deteriorated dramatically, with an R^2 of only 8%. This fact has been widely confirmed, e.g., see Philippon (2009), Table III, top panel, second column. Recent papers often describe the investment-q regression as an empirical failure, and attempt to "fix" it in various ways, as discussed above.

In the third interval, 1995-2015, the investment-q regression has begun to work much better. Over the last twenty years, the R^2 of the basic regression has been nearly

70%. Looking only at this interval, one would conclude that the simple regression implementation of q theory is in fact a resounding success.

Figure 2 performs a similar analysis in differences. The blue and red series are the year-over-year differences of the series from Figure 1. The R^2 values from the regression within each 20-year period are listed at the bottom of the figure, and they yield largely the same conclusion as in Figure 1: The R^2 of the investment-q regression fell from 16% in 1955-1975 to less than 1% in 1975-1995, but then rose to over 48% in 1995-2015. (These values can be compared with Philippon (2009) Table III, bottom panel, second column, where he finds an R^2 of 10.2% from this regression, again using data from 1953–2007.)

[Figure 2 here]

Also listed at the bottom of each period are the volatilities of the explanatory variable in the regression, differenced Tobin's q, within each sub-period. These figures provide motivating evidence for the core mechanism on which we focus in the remaining analysis: The volatility of Tobin's q was lowest during the 20-year period in which the investment-q regression performed the worst, and the volatility was the highest during the 20-year period in which the regression performed the best. In the next section, we turn to a closer examination of trends in the dispersion of q.

2.2 Increased dispersion in Tobin's q

Under the null hypothesis that the model is true, the investment-q regression should yield a higher R^2 when there is more empirical dispersion in the key explanatory variable, Tobin's q. One possible explanation for the improved fit of the aggregate regression is that the theory has always been "true," but that Tobin's q has become more volatile relative to the model's residuals. In this section, we confirm that this conjecture holds in the data. We decompose the volatility of investment and Tobin's q along two dimensions: first between-firm, and then within-firm. These two dimensions are summarized in Figures 3 and 4, respectively.

In Figure 3 we plot, for each year, the cross-sectional standard deviation of investment and Tobin's q in Compustat, then we smooth it by simple averaging over a rolling five-year lag in order to focus on trends. Both of the cross-sectional dispersion measures have trended upward over time, with investment dispersion reaching about twice its 1960 level and q dispersion reaching about four times its 1960 level.

[Figure 3 here]

In Figure 4, we investigate how Compustat firms have changed over time by examining the within-firm volatility. This dimension is more relevant to both the standard panel regressions (which include a firm fixed effect), and also to the aggregate regressions (which sum across firms and thus does not account for the cross-sectional dispersion).

[Figure 4 here]

To create this figure, we proceed in two steps. First, we calculate for each Compustat firm the within-firm volatilities of its investment and Tobin's q during its entire lifetime in Compustat. This creates measures of investment and valuation volatility that are fixed at the firm level. Next, for each year, we average these fixed volatility numbers across all firms that are present in Compustat that year. Figure 4 reveals that within-firm volatilities of investment and Tobin's q have reached about two and three times their 1960 values, respectively. The increase in Tobin's q volatility (red line) is especially noticeable in the late 1990s and early 2000s. The resulting series thus show that the composition of Compustat has shifted towards firms that exhibit higher volatilities on both dimensions.

2.3 Better performance for firms with more volatile q

Motivated by the Compustat evidence above, which shows that firms exhibit greater volatility of q in recent years, we next show that the investment-q regression works better for these firms.

The first point to make is that the within-firm volatility of Tobin's q varies by orders of magnitude across firms. We sort Compustat firms into four bins of within-firm q volatility, and find that the average volatility in the lowest bin is 0.26, while in the highest bin it is 11.29.[2] Under the q theory of investment, the greater variation in q in the latter settings will be able to trace out a much tighter connection between valuation and investment. Intuitively, there is more opportunity for the investment-q regression to work in those settings.

[2] Related, Erickson and Whited (2000) observe that Tobin's q is highly skewed in the data, which intuitively aids the identification of their strategy based on higher-order moments.

Indeed, we find that the high-volatility firms are where the regression works the best. To show this, we run standard panel regressions of investment on lagged Tobin's q. The specification is

$$\frac{I_{i,t+1}}{K_{it}} = \alpha_i + \beta \frac{V_{it}}{K_{it}} + \epsilon_{it}, \tag{1}$$

where i indexes firms, t indexes years, I is capital expenditures, K is gross property, plant, and equipment, and V is defined as the market value of equity plus book value of debt minus current assets. All of these definitions follow Peters and Taylor (2016). Results are qualitatively unchanged if we do not subtract out current assets.

Table 1 performs this regression separately across the four bins of within-firm volatility in Tobin's q, with bin 1 as the lowest volatility and bin 4 as the highest volatility. The table confirms that the quality of the fit improves when Tobin's q is more volatile.

[Table 1 here]

Figure 5 illustrates visually the variation in the data that gives rise to these results: It samples 100 observations randomly from each of the four bins, and plots the investment rate against the value of Tobin's q for each observation, along with regression lines with slopes that correspond to the coefficients in Table 1. The left-most bin shows no particular relationship between q and the investment rate, while the right-most bin illustrates a fairly tight relationship. Note that investment is not much more dispersed in the highest bin than the others; the important effect is that q becomes more dispersed relative to the residuals of the regression.

[Figure 5 here]

Again, these results should not obtain if the q theory is severely misspecified, or if the large volatility in q is meaningless for investment. Greater variation in q provides the opportunity for the investment-q regression to work, but does not force it to do so. Instead, our findings suggest that the information reflected in equity market valuations is tightly connected to investment policies, and this relationship becomes the clearest when valuations move the most.

For robustness, Table 2 repeats the analysis of Table 1, after winsorizing q at the 1st and 99th percentiles, as is standard in the literature. The average volatility of

q in the highest-volatility bin falls to about 7 instead of 11, but the results of the regression are intuitively the same as in the previous table.

[Table 2 here]

Unreported results show that the effect is even more dramatic when we winsorize investment as well. Under this approach, the R^2 of the fourth regression is over 20%. However, inspecting Figure 5, one can see that winsorizing the outcome variable at the same cutoffs across the whole sample has unpredictable effects on a comparison of R^2 values across bins of variation in the explanatory variable, and thus we do not necessarily view this approach as our strongest evidence.

In sum, this section has demonstrated that the investment-q regression works better in settings with more dispersion in Tobin's q, both in the cross-section and in the time series. In Section 3, we explain these facts with a learning model that shows why the types of firms appearing in the data in recent years are likely to exhibit a tighter relationship between their investments and their valuations.

3 Model

We develop a model of firm investment and learning. The model extends the setup analyzed by Abel (2017) to account for cash flow uncertainty and learning about the expected long-term growth in cash flows.

3.1 Setup

Consider a competitive firm with capital K_t at time t, which accumulates according to

$$dK_t = (I_t - \delta K_t)dt, \tag{2}$$

where I_t denotes the firm's investment decision.

Similar to Erickson and Whited (2000), adjustments to the capital stock are linear homogenous in I and K

$$\psi(I_t, K_t, \nu_t) = \frac{a}{2}\left(\frac{I_t}{K_t}\right)^2 K_t + \nu_t I_t, \tag{3}$$

where a is a positive constant so that the adjustment cost function is strictly convex. The term ν_t represents a shock to the purchase price of capital. It follows a stochastic process with zero mean

$$d\nu_t = -\kappa \nu_t dt + \sigma_\nu dW_t^\nu. \tag{4}$$

While the firm knows the current value of ν_t, the econometrician does not. For the econometrician, ν_t is noise.

The firm produces cash flows according to a technology with constant returns to scale:

$$\Pi(K_t, \theta_t) = \theta_t K_t, \tag{5}$$

where we use the output price as numéraire. Without loss of generality, we abstract from describing the flexible labor decision.[3]

The cash flow per unit of capital θ_t follows a mean reverting process:

$$d\theta_t = \lambda(\mu_t - \theta_t)dt + \sigma_\theta dW_t^\theta. \tag{6}$$

While the instantaneous cash flow θ_t is observable, its long-term mean, μ_t, is not. The firm forms expectations over its future stream of cash flows, but cannot perfectly infer the process driving cash flows from past realizations because the long-term mean μ_t evolves stochastically as described below.

3.2 Learning

The long-term mean around which θ_t evolves, μ_t, is not observable. It is assumed to follow a mean-reverting process

$$d\mu_t = \eta(\bar{\mu} - \mu_t)dt + \sigma_\mu dW_t^\mu. \tag{7}$$

For simplicity of exposition, we fix $\bar{\mu} = \mu_0$. In the special case without volatility

[3] We can equivalently write the firm's problem to include a labor decision. In this case, the firm produces according to a Cobb-Douglas production function $A_t L_t^\alpha K_t^{1-\alpha}$, where $0 < \alpha < 1$ and $A_t > 0$. It pays a constant wage rate w per unit of labor, set to 1 for simplicity. The instantaneous cash flow of the firm is $\max_{L_t} \left[A_t L_t^\alpha K_t^{1-\alpha} - L_t\right] = \underbrace{\alpha^{\frac{\alpha}{1-\alpha}} A_t^{\frac{1}{1-\alpha}}}_{\equiv \theta_t} K_t \equiv \Pi(K_t, \theta_t)$.

$\sigma_\mu = 0$, the long-term mean would be observable. The firm could choose how much to invest at each point in time knowing all the necessary information. However, as soon as there is some volatility $\sigma_\mu > 0$, the firm will need to continuously update its beliefs about μ.

The firm learns about the long-term mean from two sources. The first source is free. The firm uses information from past cash flow realizations in order to infer the long term mean μ_t in the process (6). The second source is costly. The firm may purchase a signal s_t that is informative about changes in the long-term mean dW_t^μ:

$$ds_t = dW_t^\mu + \frac{1}{\sqrt{\Phi}} dW_t^s, \tag{8}$$

All Brownian motions (W_t^ν, W_t^θ, W_t^μ, and W_t^s) are independent. The parameter $\Phi \geq 0$ dictates the informativeness of the signal. For now, one may consider Φ as exogenously given, and subsection 3.5 below discusses how the signal informativeness Φ is optimally chosen ex ante at a cost by the firm.

The following proposition and its corollary follow from filtering theory (Liptser and Shiryayev, 1977), with the proof provided in Appendix C.1.

Proposition 1 ***(Learning)*** *The filtered variable $\widehat{\mu}_t$ evolves according to:*

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \frac{\sigma_\theta}{\lambda}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right) d\widehat{W}_t^\theta + \sigma_\mu\sqrt{\frac{\Phi}{1+\Phi}} d\widehat{W}_t^s, \tag{9}$$

where $d\widehat{W}_t^\theta \equiv dW_t^\theta + \frac{\lambda}{\sigma_\theta}(\mu_t - \widehat{\mu}_t)dt$ represents the "surprise" component of the change in cash flows per unit of capital and $d\widehat{W}_t^s \equiv \sqrt{\frac{\Phi}{1+\Phi}} ds_t$ is a scaled version of the signal (8), such that $\widehat{W}_t^s$ is a standard Brownian motion.

The standard Brownian motion $d\widehat{W}_t^\theta$ arises as follows. The firm expects a change in cash-flows per unit of capital of $\lambda(\widehat{\mu}_t - \theta_t)dt$, but instead observes the realization $d\theta_t$. The difference, $d\theta_t - \lambda(\widehat{\mu}_t - \theta_t)dt$, represents the unexpected change, i.e., the "surprise." Dividing this difference by σ_θ yields the standard Brownian motion $d\widehat{W}_t^\theta$. This Brownian motion is distinct from the true cash-flow shock dW_t^θ which is unobservable by the firm, because it incorporates firm's expectations of future cash-flow growth (see Appendix C.1).

Corollary 1.1 *The conditional variance of the filter $\widehat{\mu}$,*

$$\mathrm{Var}_t[\widehat{\mu}_t] = \sigma_\mu^2 - \frac{2\eta\sigma_\theta^2}{\lambda^2}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{10}$$

is strictly increasing in both σ_μ and Φ.

According to Corollary 1.1, the conditional variance of the filter is lower than the conditional variance of the unobserved process (which equals σ_μ^2). With learning, the filtered long-term mean $\widehat{\mu}_t$ is a projection of μ_t on the observation filtration of the firm and thus it remains a conservative estimate of μ_t. As such, learning dampens the fluctuations of the filter $\widehat{\mu}_t$ relative to the truth μ_t.

For the rest of the paper, we refer to $d\widehat{W}_t^\theta$ as "cash-flow shocks" and to $d\widehat{W}_t^s$ as "information shocks." Two key results arise from Proposition 1 and its Corollary, reflecting the two sources of information from which firms learn. First, learning from cash flow realizations induces a positive correlation between the filter $\widehat{\mu}_t$ and cash-flows θ_t, through cash-flow shocks $d\widehat{W}_t^\theta$. This extrapolative feature of learning (Brennan, 1998) amplifies the impact of cash-flow shocks.

Second, learning from the signal s_t causes the firm's estimate of the long-term cash flow mean $\widehat{\mu}_t$ to respond to information shocks $d\widehat{W}_t^s$. This increases the conditional volatility of $\widehat{\mu}_t$. We note that the learning that takes place does not change the conditional volatility of the process (6), which remains constant at σ_θ for any level of σ_μ. Learning, however, does change the volatility of the filter through the continuous updating of the long-term cash flow mean $\widehat{\mu}_t$.

3.3 The investment decision

The firm's objective is to maximize the expected discounted sum of future cash flows, net of investment costs,

$$V(K_t, \theta_t, \widehat{\mu}_t, \nu_t) = \max_I \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)}\left\{\theta_s K_s - I_s - \psi(I_s, K_s, \nu_s)\right\} dt\right], \tag{11}$$

subject to Equations (2) and (3), where r is the interest rate.

The Hamilton-Jacobi-Bellman equation associated with problem (11) is

$$rV = \max_I \left\{ \theta K - I - \psi(I, K, \nu) + \mathcal{D}V(K, \theta, \widehat{\mu}, \nu) \right\}, \tag{12}$$

where $\mathcal{D}$ is the differential operator. This leads to the first order condition for investment,

$$0 = V_K(K, \theta, \widehat{\mu}, \nu) - 1 - \psi_I(I, K, \nu). \tag{13}$$

In our model, as in Hayashi (1982), the shadow cost of capital, marginal q is equal to average $q = \frac{V}{K}$:

$$V(K, \theta, \widehat{\mu}, \nu) = q(\theta, \widehat{\mu}, \nu)K. \tag{14}$$

Replacing the adjustment cost function (3) yields the following relationship between the rate of investment and q:

$$\frac{I_t}{K_t} = -\frac{1}{a} + \frac{1}{a} q(\theta_t, \widehat{\mu}_t, \nu_t) - \frac{1}{a}\nu_t. \tag{15}$$

Using equation (14) and solving for the optimal investment, we obtain the following partial differential equation for q:

$$\begin{aligned} 0 = {} & \theta_t + \frac{(1+\nu_t)^2}{2a} - \frac{1 + a(r+\delta) + \nu_t}{a} q + \frac{1}{2a} q^2 \\ & + \lambda(\widehat{\mu}_t - \theta_t) q_\theta + \eta(\bar{\mu} - \widehat{\mu}_t) q_{\widehat{\mu}} - \kappa \nu_t q_\nu \\ & + \frac{\sigma_\theta^2}{2} q_{\theta\theta} + \left[\frac{\sigma_\mu^2}{2} - \frac{\eta \sigma_\theta^2}{\lambda^2} \left(\sqrt{\eta^2 + \frac{1}{1+\Phi} \frac{\lambda^2 \sigma_\mu^2}{\sigma_\theta^2}} - \eta \right) \right] q_{\widehat{\mu}\widehat{\mu}} + \frac{\sigma_\nu^2}{2} q_{\nu\nu} \\ & + \frac{\sigma_\theta^2}{\lambda^2} \left(\sqrt{\eta^2 + \frac{1}{1+\Phi} \frac{\lambda^2 \sigma_\mu^2}{\sigma_\theta^2}} - \eta \right) q_{\theta\widehat{\mu}}. \end{aligned} \tag{16}$$

We solve this equation numerically by approximating $q(\theta, \widehat{\mu}, \nu)$ with Chebyshev polynomials.[4]

[4] Since θ, $\widehat{\mu}$, and ν are all mean-reverting, we define a grid that is centered on $\{\bar{\mu}, \bar{\mu}, 0\}$. The algorithm yields a very accurate solution, with an approximation error of magnitude 10^{-23} obtained with four polynomials in each dimension. For a similar approach, see Alti (2003).

3.4 Learning and the relationship between investment and q

Without ν, the econometrician would observe a deterministic relationship between investment and q in equation (15) and, counterfactually, this relationship would always have a R^2 of one. The shock to the capital purchase price causes the R^2 to be below one:

$$R^2 = \frac{\mathrm{Var}[q(\theta_t, \widehat{\mu}_t, \nu_t)] \left(1 - \frac{\mathrm{Cov}[q(\theta_t, \widehat{\mu}_t, \nu_t), \nu_t]}{\mathrm{Var}[q(\theta_t, \widehat{\mu}_t, \nu_t)]}\right)^2}{\mathrm{Var}[q(\theta_t, \widehat{\mu}_t, \nu_t)] + \mathrm{Var}[\nu_t] - 2\,\mathrm{Cov}[q(\theta_t, \widehat{\mu}_t, \nu_t), \nu_t]}. \tag{17}$$

Equation (17) has two important implications. First, the econometrician obtains a R^2 lower than one. Second, if the covariance between q and ν is negligible, the R^2 coefficient increases with the variance of q.[5] Notice also that a stronger regression coefficient for q in Equation (15) does not mechanically improve the R^2, since the parameter a simplifies away from (17).

The firm's learning affects the econometrician's coefficient of determination. This can be seen from an application of Itô's lemma on $q(\theta_t, \widehat{\mu}_t, \nu_t)$:

$$dq = \xi_t dt + \left[q_\theta \sigma_\theta + q_{\widehat{\mu}} \frac{\sigma_\theta}{\lambda} \left(\sqrt{\eta^2 + \frac{1}{1+\Phi} \frac{\lambda^2 \sigma_\mu^2}{\sigma_\theta^2}} - \eta \right) \right] d\widehat{W}_t^\theta + q_{\widehat{\mu}} \sigma_\mu \sqrt{\frac{\Phi}{1+\Phi}} d\widehat{W}_t^s, \tag{18}$$

where the drift ξ_t is a function whose explicit form is not necessary for the arguments made here. When the firm learns about the unobservable productivity growth μ_t, q becomes more sensitive to cash-flow shocks $d\widehat{W}_t^\theta$ (through the second term in square brackets above), but also to information shocks $d\widehat{W}_t^s$ (through the last term above). Both these effects increase the volatility of $q(\theta_t, \widehat{\mu}_t, \nu_t)$ and, according to Equation (17), the R^2 of the q regression.

We illustrate by means of simulations the impact of learning on econometrician's relationship between investment and q. To this aim, we implement a discretization of continuous-time processes at quarterly frequency (see Appendix C.2). We then solve for the partial differential equation (16) for each simulated point $\{\theta_t, \widehat{\mu}_t, \nu_t\}$.

[5] In our numerical calibration, we ensure that the covariance between q_t and ν_t is virtually zero (in other words, q_t does not depend on ν_t). This occurs for large values of κ, i.e., when the persistence of ν_t is close to zero. When the persistence of ν_t is non-negligible, it creates temporal dependence through which q_t depends on ν_t. Even in this case, the covariance term in (17) is of small magnitude, and does not impact the intuition discussed here. See also Erickson and Whited (2000, page 1036).

The resulting value for q_t can then be replaced in Equation (15), yielding the investment to capital ratio I_t/K_t. This completes the dataset necessary for implementing investment-q regressions.

Figure 6 shows results from one simulation of 100 years of quarterly data. The horizontal axis in each panel is the marginal q. The vertical axis represents the optimal investment to capital ratio I_t/K_t. The calibration used is provided in the caption of the plot.

[Figure 6 here]

The left panel corresponds to the case of an unobservable μ without learning, that is, the firm sets $\mu_t = \bar{\mu}, \forall t$. In the middle panel, the firm learns about μ_t, but only using the observable process for θ_t, i.e., $\Phi = 0$. In the right panel, the firm also learns through the signal (8), with $\Phi = 100$. Changes in μ_t are not yet perfectly observed, but with $\Phi = 100$ the signal (8) is much more informative relative to the cash-flow signal (6). The three panels show that learning improves the fit of the regression. As elaborated above, this occurs through an increase in the volatility of the regressor q. The R^2 coefficients typically obtained with this particular calibration are of magnitudes of about 10% for the left panel, 40% for the middle panel, and 55% for the right panel.

3.5 Endogenous learning

In this section, we endogenize the information acquisition problem. Specifically, the firm can purchase a more informative signal (higher Φ) and learn more about μ_t, but the more informative signal is costlier (Detemple and Kihlstrom, 1987). The cost can be viewed as a research expense that firms incur. We consider here a static information acquisition decision, in which the firm makes a choice of Φ at time 0 and maintains this capacity of information acquisition over its lifetime.

The firm value immediately after the choice of informativeness Φ is defined as $\tilde{V}(\cdot)$, and its associated cost, $c(\Phi)$, is a strictly increasing and convex function with $c'(0) = 0$. With Φ as a parameter in $\tilde{V}(\cdot)$, the problem is equivalent to the earlier model without an endogenous Φ. The optimal Φ^* is defined by the first-order condition $\tilde{V}_\Phi(\cdot) = c'(\Phi^*)$, and there is an interior solution if and only if $\tilde{V}_{\Phi\Phi}(\cdot) - c''(\Phi^*) < 0$.

Differentiating the first-order condition with respect to σ_μ and rearranging, we get

$$\frac{d\Phi^*}{d\sigma_\mu} = \frac{\tilde{V}_{\Phi\sigma_\mu}(\cdot)}{c''(\Phi^*) - \tilde{V}_{\Phi\Phi}(\cdot)}. \tag{19}$$

The denominator is positive if the problem has an interior solution. The optimal amount of information acquisition, i.e., the optimal choice of Φ, increases in σ_μ if and only if the cross-derivative $\tilde{V}_{\Phi\sigma_\mu}(\cdot) > 0$. Intuitively, the optimal purchase of information increases with the uncertainty about μ_t, if and only if an increase in σ_μ increases the marginal benefit of purchasing information.

The problem reduces to evaluating $\tilde{V}(\cdot)$, which is just the value function from the problem without an endogenous signal informativeness choice Φ, and then demonstrating that $\tilde{V}_{\Phi\sigma_\mu}(\cdot) > 0$. While there is no closed-form proof of this result, it can be easily checked numerically. Figure 7 shows that this is indeed the case. The left panel plots the function $q(\theta_t, \widehat{\mu}_t, \nu_t)$ when Φ varies from 0 to 1 on the x-axis. Each line in the plot corresponds to a different value of σ_μ: for the bottom line σ_μ is equal to 0.01, whereas for the top line σ_μ is equal to 0.2. For these plots, the state variables are fixed at $\theta_t = \widehat{\mu}_t = \bar{\mu}$ and $\nu_t = 0$, but qualitative results do not change with different parameter values. The second derivative $q_{\Phi\sigma_\mu}$ is approximated using finite difference. The right panel of Figure 7 computes this derivative and shows that it is positive at all times, suggesting that the optimal information acquisition level Φ^* increases in σ_μ.

[Figure 7 here]

This implies that firms that operate in more uncertain environments (e.g., high-tech firms) choose optimally to invest more in research. Considering then the result of the previous section that more learning increases the R^2 of the q regression, this generates a cross-sectional implication: the investment-q regression performs better for firms that spend more on gathering information through research and development.

4 Empirical analysis of the model

The model in the previous section demonstrated that the investment-q regression performs better in settings with greater learning, and that this effect will be even stronger among firms that endogenously acquire more information. In this section, we turn back to the data to assess other testable implications of the model.

4.1 Better performance in high-tech industries

The most important result from the model was in Section 3.5, which showed that the investment-q regression performs better among firms that endogenously choose to expend greater resources on information acquisition.

Empirically, we interpret this decision as manifesting itself in spending on research. A firm that invests heavily in R&D is by definition attempting to learn, more so than a firm that only invests in physical capital to exploit existing technology. Therefore, our proposed learning mechanism should cause the investment-q regression to work better in industries featuring high investment in research. This insight provides both testable cross-sectional implications of the model, and broader insights about the future of the q theory of investment.

For an operational definition of a research-intensive industry, we select the following seven SIC codes: 283 (drugs); 357 (office and computing equipment); 366 (communications equipment); 367 (electronic components); 382 (scientific instruments); 384 (medical instruments); and 737 (software). We will refer to these as "research-intensive" or "high-tech" industries for the remainder of this paper. Our industry classification follows Brown, Fazzari, and Petersen (2009), who show that these seven industries account for nearly all the growth in aggregate R&D during the 1990s. Our interpretation of the model in terms of research expenditures makes this industry classification a natural proxy for the mechanism we have in mind. However, many other reasonable proxies are highly correlated with this classification: For example, the same industries feature high rates of patenting and low tangibility.

Turning to the data, we build up our analysis of research-intensive industries in several layers. First, we examine the empirical distribution of Tobin's q for these industries compared to the average Compustat firm, following our general intuition that focuses on the empirical variation in q. Figure 8 displays the empirical density of Tobin's q for firm-years in the high-tech and other industries in the sample.

[Figure 8 here]

Figure 9 calculates the within-firm volatility of Tobin's q (which is more relevant for panel regressions), retains one observation per firm, and plots the empirical density of this volatility, again separating out the high-tech and other industries.

[Figure 9 here]

As predicted by the model, for firms that spend more on research, the empirical distribution of Tobin's q is much higher, more dispersed, and more skewed. Intuitively, in these industries, the firm's value at any point in time is much more sensitive to the arrival of signals about future profitability.

On some level, it is not surprising that equity-market valuations of high-tech companies are particularly volatile. Assigning a valuation to these companies is difficult, and their gross PP&E (the denominator of Tobin's q) is smaller. However, this intuition suggests that the fluctuations in equity value are simply "noise" with respect to the firm's investment policy. In contrast, we now show that those fluctuations are highly predictive of investment, which is the core implication of the q theory of investment.

Tables 3 and 4 repeat the panel regressions of investment on lagged q with fixed effects, as specified earlier in equation (1) and implemented in Tables 1 and 2. The various columns separate out the high-tech from the other industries to demonstrate how the performance of the regression differs across these subsamples.

[Tables 3 and 4 here]

Columns 1 and 2 of Table 3 show that the standard investment-q panel regression fares much better among high-tech than other industries: The R^2 value from the regression more than doubles from 10.4% to 21.7% when we move from the non- to the high-tech subsample. One may object that, since we already have shown that the investment-q regression works better in recent years, this comparison simply captures the increasing importance of high-tech firms towards the end of the sample. To check this, in Columns 3 and 4 we restrict to the years before 2000. The same discrepancy holds across the subsamples: The R^2 of the panel regression is over 17% for the high-tech industries, and only 8.3% for the other industries. We conclude that the investment-q regression works better for high-tech than for other industries. In fact, we argue that the growth of these industries may therefore explain the improved fit of the aggregate regression.

Table 4 checks robustness to some alternative approaches. Column 1 shows that firm fixed effects are not critical to the performance of the regression in high-tech industries, as the (overall) R^2 from the pooled regression is similar to the (within) R^2 that was reported in Table 3. Column 2 shows that the fit of the regression improves even more when we add time fixed effects, as is done in some other papers. Columns

3 and 4 return to our main panel specification, with firm fixed effects and no time fixed effects, and adds in annual R&D expenditures as an "investment" in intangible assets. We find that the coefficient on lagged q increases greatly, especially in high-tech industries where it rises from 0.0119 (in Column 2 of Table 3) to 0.0245 (in Column 3 of Table 4). The R^2 values from these regressions also increase somewhat compared to the regressions ignoring R&D, but the primary conclusion is the same as before: The regression works better in high-tech industries ($R^2 = 27\%$) than in other industries ($R^2 = 12\%$).

The stylized fact that the investment-q regression works better in high-tech industries was already established in Peters and Taylor (2016) (see their Section 5.1). They move beyond our industry-based classification and use several different proxies for intangibility. They explore a few different explanations, but ultimately reject all of them, and simply conclude as follows: *"Why the classic q-theory fits the data better in high-intangible settings is also an interesting open question."* Our model suggests that this finding is not so puzzling.

In fact, the growth of high-tech industries is likely critical to understanding the improved fit of the aggregate investment-q relationship in recent years, and by extension the future empirical performance of the q theory of investment. Figure 10 shows that the firms in our high-tech industry classification represent a growing fraction of firms, of book assets, and of market value in Compustat. Similarly, Peters and Taylor (2016) show that their measure of intangible capital, which capitalizes past intangible investments such as R&D and SG&A, has also increased in both Compustat and in aggregate figures from the Fed Flow of Funds.

In conjunction with our cross-sectional findings, these trends suggest that the q theory of investment may have been the right theory at the wrong time. While the theory has traditionally not fared well for the capital-intensive firms that dominated the economy when the theory was first developed, it turns out to be well-suited for the new research-intensive economy that features wider swings in both valuation and investment.

4.2 Better performance with low cash flow-q correlation

We next explore a subtler implication of the model. The model predicts that the investment-q regression works better in settings where Tobin's q is less correlated with cash flow. With learning, q becomes less responsive to cash flow because the firm chooses to pay more attention to other signals.

The learning mechanism works in the opposite direction as misspecification issues, which have been the focus in much of the prior research. Consider a misspecified regression by omitting cash flow when an alternative theory (e.g., based on financial constraints) would predict that cash flow is an important variable. With such an omitted variable bias, the investment-q regression should work better in settings where Tobin's q is *more*, not less, correlated with cash flow, as the bias shrinks when q and cash flow are more highly correlated.

We test the relative importance of the learning mechanism against potential misspecification. To implement this test, we sort industries into quartiles based on their estimated correlations of Tobin's q with cash flow in Compustat data, and compare across these quartiles the R^2 from the investment-q regression. Figure 11 illustrates the results.

[Figure 11 here]

The pattern strongly supports the learning mechanism. The lowest quartile of correlation between q and cash flow features an R^2 of 18%, compared to 10% for the highest quartile. While this does not indicate that there is no misspecification in our model, it does suggest that the empirical effects of this misspecification are outweighed by the learning mechanism. Intuitively, when q is more volatile, it is highly correlated with investment but not with cash flow.

4.3 Better performance in low-interest-rate environments

Another natural implication of our model is that the performance of the investment-q regression should be sensitive to the level of interest rates. Fluctuations in the interest rate imply fluctuations in the firm's value function, relative to any component of the model residuals ν that does not scale with the interest rate. This increased volatility should again cause the standard regression to perform better.

Indeed, Figures 1 and 2 demonstrated exactly this result in the data: The middle era in the figure contained the highest nominal interest rates of the post-war era, and this coincides with the era of worst performance of the aggregate investment-q regression. In both the earlier and later eras, by constrast, interest rates were lower and the regression performed better. We take this stylized fact, which appears not to have been noticed in prior literature, as suggestive evidence that the performance of q theory can be understood by seeking out settings where Tobin's q varies more.

However, note that nominal interest rates in the last third of the figure were roughly the same as in the first third, yet the aggregate regression performs noticeably better in the last third than in the first third. This signals that interest rates alone cannot explain the time-series variation in the performance of the regression, leaving room for our model's mechanism based on variation in firms' learning.

5 Conclusion

This paper is motivated by our empirical observation that the relationship between aggregate investment and Tobin's q has become remarkably tight in recent years. This observation stands in contrast to a large literature showing that this regression performed quite poorly in the past.

We attribute the improvement in the empirical performance of the classic regression to an increase in the empirical variation in Tobin's q relative to residual factors affecting investment. We rationalize this increase in variation with a learning-based model of corporate investment. Better learning by firms endogenously produces more variation in marginal q, improving the fit of the regression. Our learning mechanism is relevant especially in intangible-intensive industries. Thus, the improved fit of the investment-q relationship fits well with the observation that expenditures on research and other intangible assets have grown substantially in the aggregate. Besides delivering the above empirical observation, our theoretical model has several testable predictions. We test these predictions in the cross-section of firms in Compustat, and find empirical support for our learning mechanism.

Our main conclusion is that even a simple version the q theory of investment can describe the data quite well, when given sufficient variation in the key regression variables. Counterintuitively, that variation occurs in firms that probably look the least like the classic capital-intensive firms for which the theory was initially developed. As these firms become a growing segment of the economy, the future of the investment-q relationship should look increasingly bright.

References

Abel, Andrew, 2017, The Effects of q and Cash Flow on Investment in the Presence of Measurement Error, *Journal of Financial Economics* (forthcoming).

Alti, Aydoḡan, 2003, How sensitive is investment to cash flow when financing is frictionless?, *The Journal of Finance* 58, 707–722.

Brennan, Michael J, 1998, The role of learning in dynamic portfolio decisions, *European Finance Review* 1, 295–306.

Brown, James R., Steven M. Fazzari, and Bruce C Petersen, 2009, Financing Innovation and Growth: Cash Flow, External Equity, and the 1990s R&D Boom, *Journal of Finance* 44, 151–185.

Cooper, Russell, and Joao Ejarque, 2003, Financial frictions and investment: requiem in q, *Review of Economic Dynamics* 6, 710–728.

Detemple, Jérôme, and Richard Kihlstrom, 1987, Acquisition d'information dans un modèle intertemporel en temps continu, *L'Actualité économique* 63, 118–137.

Dumas, Bernard, Alexander Kurshev, and Raman Uppal, 2009, Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility, *The Journal of Finance* 64, 579–629.

Erickson, Timothy, and Toni Whited, 2000, Measurement Error and the Relationship between Investment and q, *Journal of Political Economy* 108, 1027–1057.

Fazzari, Steven, R. Glenn Hubbard, and Bruce Petersen, 1988, Financing Constraints and Corporate Investment, *Brookings Papers on Economic Activity* 1988, 141–206.

Gomes, João, 2001, Financing Investment, *American Economic Review* 91, 1263–1285.

Gutiérrez, Germán, and Thomas Philippon, 2016, Investment-less Growth: An Empirical Investigation, *NYU working paper* .

Hall, Robert E., 2001, The Stock Market and Capital Accumulation, *American Economic Review* 91, 1185–1202.

Hayashi, Fumio, 1982, Tobin's Marginal q and Average q: A Neoclassical Interpretation, *Econometrica* 50, 213–224.

Kaplan, Steven N, and Luigi Zingales, 1997, Do investment-cash flow sensitivities provide useful measures of financing constraints?, *The quarterly journal of economics* 112, 169–215.

Liptser, R, and A Shiryayev, 1977, Statistics of random processes, Springer-Verlag, Berlin, New York.

Moyen, Nathalie, and Stefan Platikanov, 2012, Corporate investments and learning, *Review of Finance* 17, 1437–1488.

Peters, Ryan, and Lucian Taylor, 2016, Intangible Capital and the Investment-q Relation, *Journal of Financial Economics* (forthcoming).

Philippon, Thomas, 2009, The Bond Market's q, *Quarterly Journal of Economics* 124, 1011–1056.

A Tables and figures

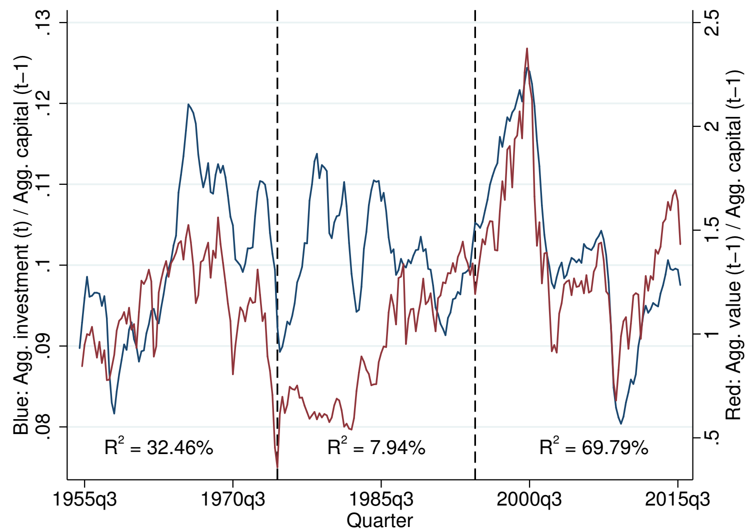

Figure 1: Aggregate quarterly investment rate and lagged Tobin's q. Both series are constructed as in Hall (2001). Compare R^2 values with Philippon (2009), Table III, top panel.

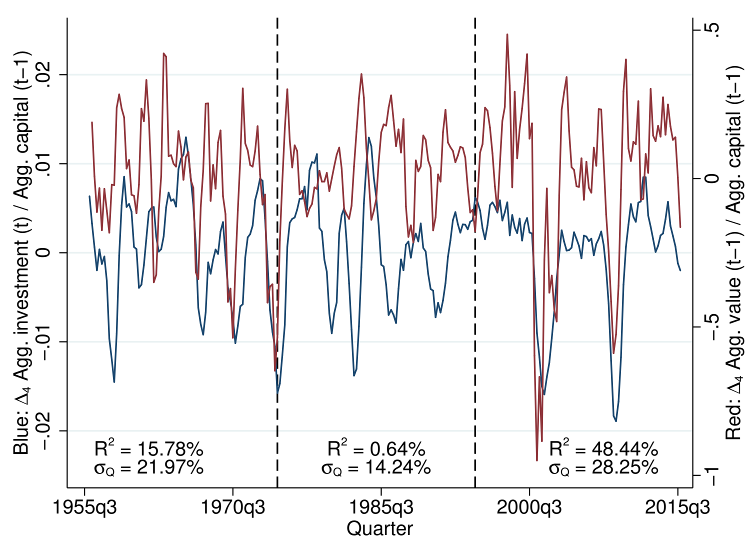

Figure 2: 4-quarter differences in investment and q. Both series are constructed as in Hall (2001). Compare R^2 values with Philippon (2009), Table III, bottom panel.

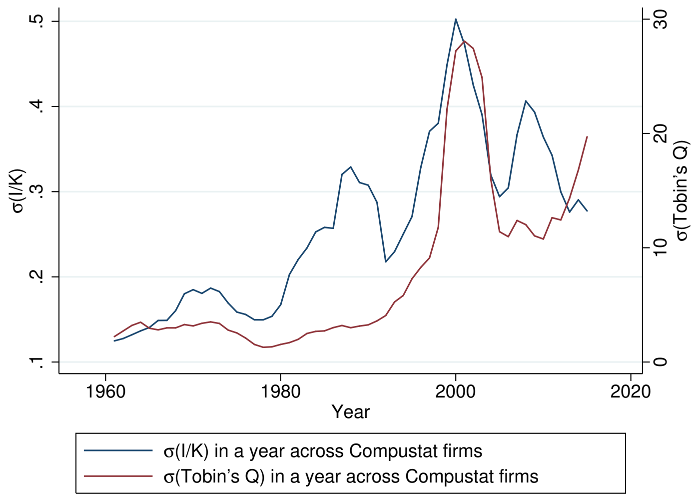

Figure 3: Between-firm dispersion in investment and Tobin's q, 1960-2015. For each year, the figure plots the cross-sectional standard deviation of that year's investment rate (left axis) and Tobin's q (right axis), across the firms in Compustat during that year. Both series are calculated as in Peters and Taylor (2016), and are then smoothed over a five-year lag.

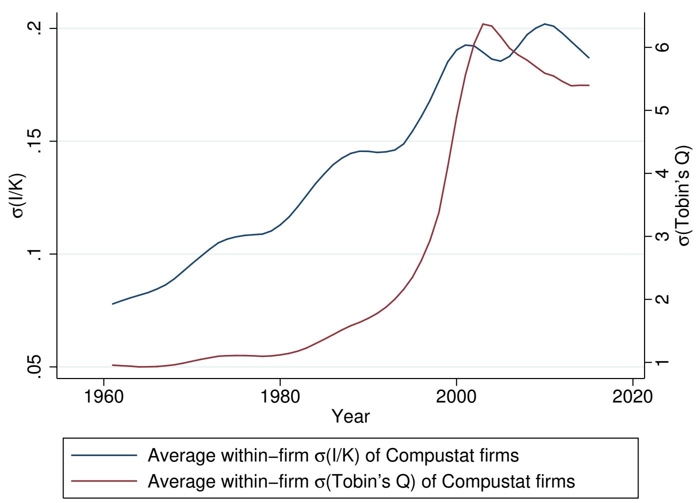

Figure 4: Within-firm dispersion in investment and Tobin's q, 1960-2015. For each firm in Compustat, we calculate the within-firm volatility of investment rate and Tobin's q during that firm's entire lifetime in Compustat. We then average that firm-level measure across all firms in Compustat for each year. The series are thus driven entirely by changes in the composition of Compustat firms.

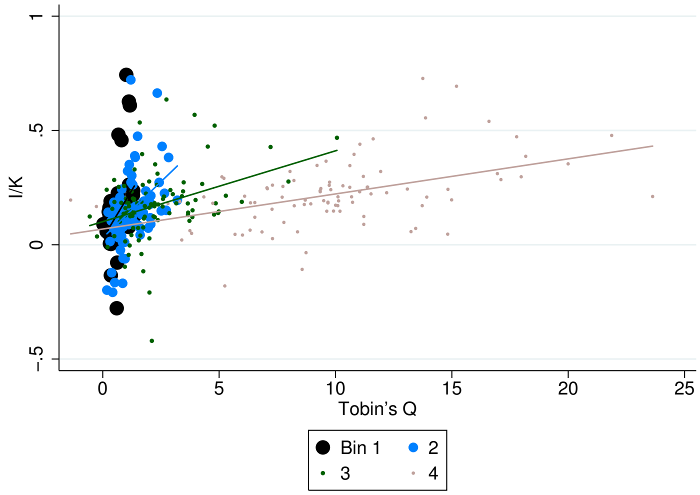

Figure 5: This scatter plot corresponds to Table 1. Each dot represents a firm-year observation. The x-axis measures Tobin's q, and the y-axis measures the investment rate, in both cases after demeaning within-firm. For each bin, the best-fit line of the same color reflects the regression in Table 1.

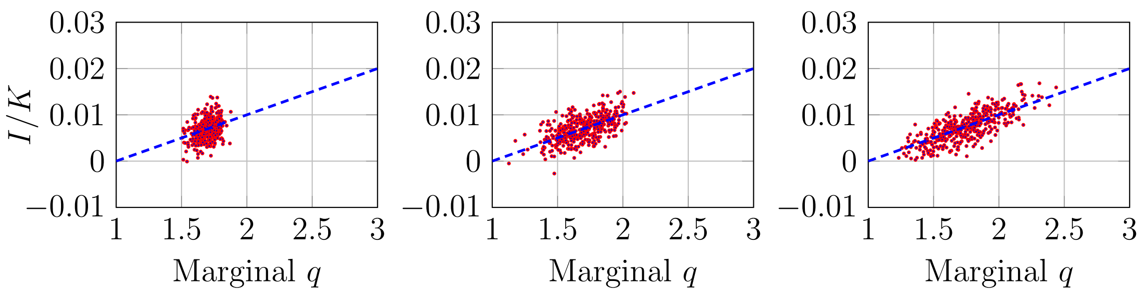

Figure 6: Relationship between investment and q for three different firms (simulated data). In the left panel, the firm does not learn about μ, which is fixed constant at $\bar{\mu}$. In the middle panel, the firm learns about μ exclusively from the cash-flow process (6) (i.e., $\Phi = 0$). In the right panel, the firm learns about μ from the cash-flow process (6) and from the signal (8) (i.e., $\Phi > 0$). The parameters used for these simulations are: $a = 100$, $r = 10\%$, $\delta = 2\%$, $\lambda = 1$, $\sigma_\theta = 0.1$, $\kappa = 100$, $\sigma_\nu = 3$, $\bar{\mu} = 0.2$, $\eta = 1$, $\sigma_\mu = 0.2$. For the right panel, we fix $\Phi = 100$.

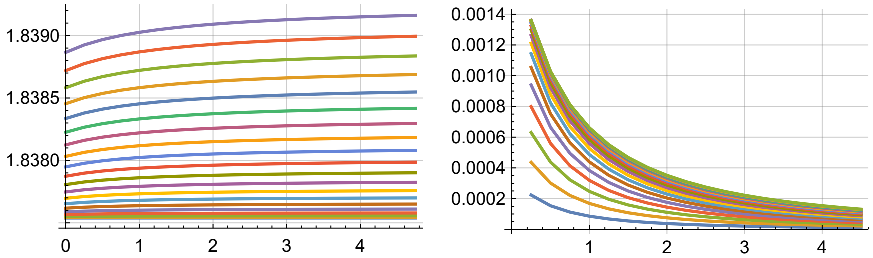

Figure 7: The left panel plots $q(\theta_t, \widehat{\mu}_t, \nu_t)$ when $\theta_t = \widehat{\mu}_t = \bar{\mu}$ and $\nu_t = 0$. Each line corresponds to a different value of σ_μ: for the bottom line $\sigma_\mu = 0.01$; for the top line $\sigma_\mu = 0.2$. The lines are plotted as functions of Φ, which goes from 0 to 4.75. The right panel uses the finite difference method to compute $q_{\Phi\sigma_\mu}$, which is positive in all cases.

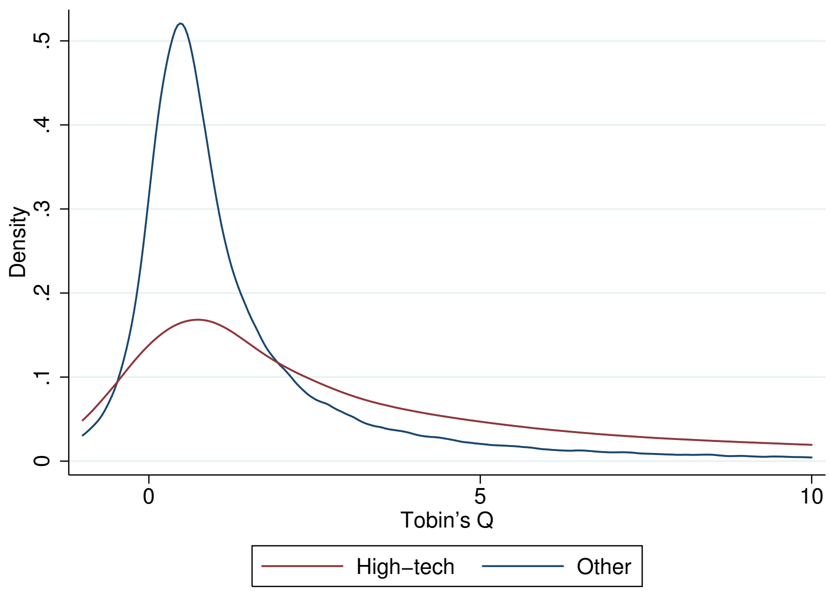

Figure 8: Empirical distribution of Tobin's q for firm-years in annual Compustat from 1956 to 2015, separating out high-tech industries from other industries. High-tech industries are defined as SIC codes 283, 357, 366, 367, 382, 384, and 737.

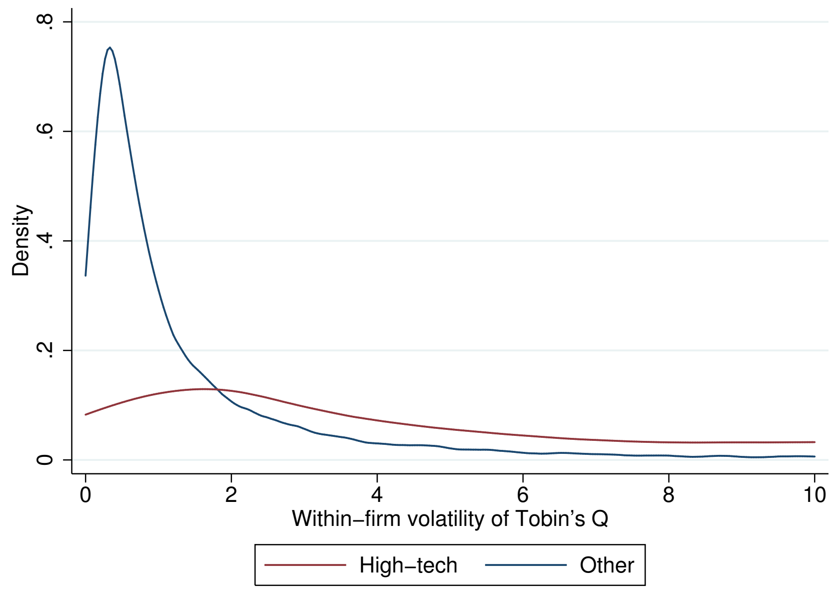

Figure 9: Empirical distribution of the within-firm volatility of Tobin's q for firm-years in annual Compustat from 1956 to 2015. The figure calculates the within-firm volatility, then retains one observation per firm, and separates out high-tech industries from other industries. High-tech industries are defined as SIC codes 283, 357, 366, 367, 382, 384, and 737, following Brown et al. (2009).

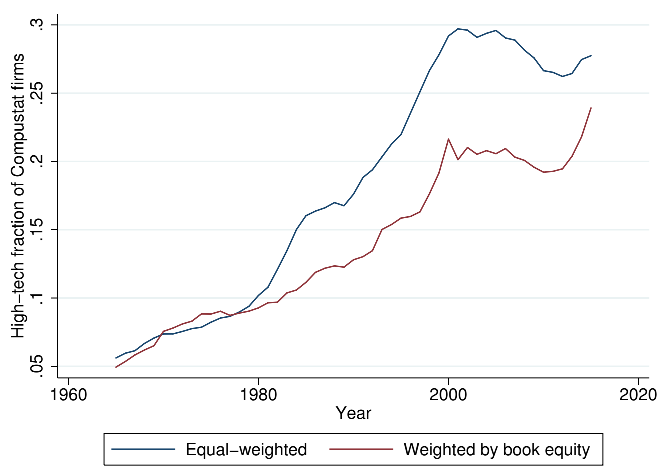

Figure 10: Fraction of firms in Compustat each year that fall into our classification of high-tech industries. The blue line is an equal-weighted average, while the red line weights firms by their shareholders' equity. High-tech industries are defined as SIC codes 283, 357, 366, 367, 382, 384, and 737, following Brown et al. (2009).

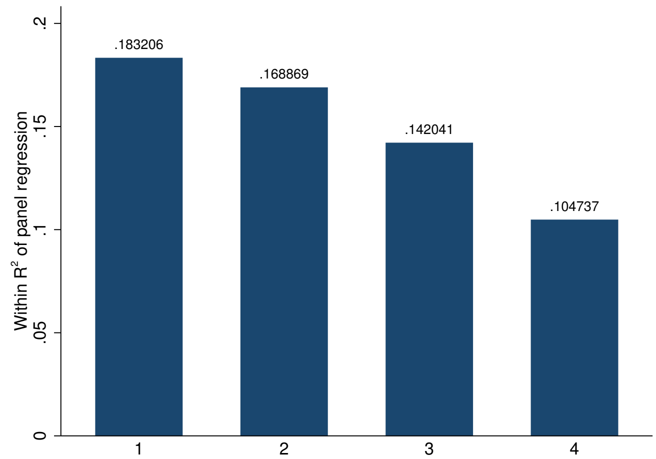

Figure 11: This figure plots within-R^2 values from panel regressions of investment on lagged Tobin's q. The data are annual Compustat from 1956-2015. The regressions are performed separately for quartiles of the industry-level correlation between q and cash flow, where industry is measured by 3-digit SIC code. We drop industries for which this correlation is estimated to be negative (about 20% of the sample), and firm-years with missing values of q or cash flow. Cash flow is defined as operating income before depreciation, scaled by gross PP&E. Cash flow, investment, and q are all winsorized at the first and 99th percentiles.

	(1)	(2)	(3)	(4)
	I_y/K_{y-1}	I_y/K_{y-1}	I_y/K_{y-1}	I_y/K_{y-1}
q_{y-1}	0.0801***	0.0638***	0.0375***	0.00474***
	(0.00447)	(0.00242)	(0.00157)	(0.000533)
Sample	Bin 1	Bin 2	Bin 3	Bin 4
Firm FE?	Yes	Yes	Yes	Yes
Obs.	37806	42365	43761	43723
R^2	0.0317	0.0619	0.0721	0.122

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1: This table performs panel regressions of investment on lagged Tobin's q, using annual data from Compustat, after sorting firms into bins based on the within-firm volatility of Tobin's q (with bin 4 as the highest volatility). Investment rate and Tobin's q are both defined as in Peters and Taylor (2016), and the same data filters are applied as in their analysis, except that the regression variables in this table are not winsorized. Standard errors are clustered by firm, and the table reports the within-firm R^2 of the regression.

	(1)	(2)	(3)	(4)
	I_y/K_{y-1}	I_y/K_{y-1}	I_y/K_{y-1}	I_y/K_{y-1}
q_{y-1}	0.0792***	0.0634***	0.0388***	0.0172***
	(0.00454)	(0.00250)	(0.00154)	(0.000562)
Sample	Bin 1	Bin 2	Bin 3	Bin 4
Firm FE?	Yes	Yes	Yes	Yes
Obs.	37692	42359	43814	43790
R^2	0.0291	0.0584	0.0795	0.146

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: This table repeats the analysis of Table 1, after winsorizing Tobin's q at the 1st and 99th percentiles.

	(1) $I_y/K_{y\text{-}1}$	(2) $I_y/K_{y\text{-}1}$	(3) $I_y/K_{y\text{-}1}$	(4) $I_y/K_{y\text{-}1}$
$q_{y\text{-}1}$	0.0172***	0.0119***	0.0207***	0.0161***
	(0.000524)	(0.000297)	(0.000855)	(0.000696)
Sample	Non-high-tech	High-tech	Non-high-tech, pre-2000	High-tech, pre-2000
Firm FE?	Yes	Yes	Yes	Yes
Obs.	133860	37641	81645	16662
R^2	0.104	0.217	0.0832	0.178

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3: This table performs panel regressions of investment on lagged Tobin's q using annual data from Compustat. Investment rate and Tobin's q are both defined as in Peters and Taylor (2016) (see discussion in the main text), and the same data filters are applied as in their analysis. "High-tech" refers to SIC codes 283, 357, 366, 367, 382, 384, and 737, following Brown et al. (2009). The data are annual Compustat from 1956-2015. Columns 3 and 4 restrict to pre-2000 firm-years. Investment and q are winsorized at the first and 99th percentiles. Standard errors are clustered by firm, and the table reports the within-firm R^2 of the regression.

	(1) $I_y/K_{y\text{-}1}$	(2) $I_y/K_{y\text{-}1}$	(3) $(I\text{+}R\&D)_y/K_{y\text{-}1}$	(4) $(I\text{+}R\&D)_y/K_{y\text{-}1}$
$q_{y\text{-}1}$	0.00934***	0.0113***	0.0245***	0.0278***
	(0.000229)	(0.000275)	(0.000832)	(0.000661)
Sample	High-tech	High-tech	Non-high-tech	High-tech
Firm FE?	No	Yes	Yes	Yes
Year FE?	No	Yes	No	No
Obs.	37641	37641	133860	37641
R^2	0.212	0.302	0.121	0.271

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: The regressions are as in Table 3, except as noted in each column: Column 2 adds year fixed effects (as done in Peters and Taylor (2016)), and Column 1 drops both firm and year fixed effects. In columns 3 and 4, R&D is added to capital expenditures as a measure of intangible "investment." Standard errors are clustered by firm, and the table reports the within-firm R^2 of the regression. (The year fixed effects in Column 2 contribute to the high R^2 value in that column.)

B Data details

The data from Hall (2001) and Philippon (2009) are publicly available, but they end in 1999 and 2007 respectively. For our analysis, it is critical to have the last few years of data. Therefore, we reconstructed the series of aggregate investment and Tobin's q following the discussion in those papers. As demonstrated in the figures and correlation matrices on the following page, our series are very close to the prior papers but not quite the same:

- Most notably, the q series reported in Philippon (2009) is much higher than the other two, reaching a maximum around 5. In contrast, the series from Hall (2001) peaks around 3, while our reconstruction peaks around 2.

- Our q series does not perfectly track Hall (2001) either, but this is completely explained by a restatement of the replacement-cost series in the Fed Flow of Funds (the denominator of Tobin's q) since his series were constructed.

- Hall (2001) does not perform the usual investment-q regression in his paper, so we do not reconstruct his investment series. Our reconstruction of the investment series in Philippon (2009) is tightly correlated with his but does not match perfectly, with the residual going in both directions. Further investigation should reveal whether this is due to a difference in the data analysis, or rather due to a restatement of the series, as was the case with the replacement-cost series above.

These discrepancies should not affect our linear regressions much, since all the series are correlated at 94% or more.

C Proofs

C.1 Proof of Proposition 1

The observable variables are the cash flow process (6) and the signal (8). The unobservable variable is μ_t. Write the dynamics of the observable variables θ_t and s_t:

$$\begin{bmatrix} d\theta_t \\ ds_t \end{bmatrix} = \left(\underbrace{\begin{bmatrix} -\lambda\theta_t \\ 0 \end{bmatrix}}_{A_0} + \underbrace{\begin{bmatrix} \lambda \\ 0 \end{bmatrix}}_{A_1} \mu_t \right) dt + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{B_1} dW_t^{\mu} + \underbrace{\begin{bmatrix} \sigma_\delta & 0 \\ 0 & \frac{1}{\sqrt{\Phi}} \end{bmatrix}}_{B_2} \begin{bmatrix} dW_t^{\theta} \\ dW_t^{s} \end{bmatrix}, \tag{20}$$

and of the unobservable variable μ_t:

$$d\mu_t = (\underbrace{\eta\bar{\mu}}_{a_0} + \underbrace{(-\eta)}_{a_1}\mu_t)dt + \underbrace{\sigma_\mu}_{b_1} dW_t^{\mu} + \underbrace{\begin{bmatrix} 0 & 0 \end{bmatrix}}_{b_2} \begin{bmatrix} dW_t^{\theta} \\ dW_t^{s} \end{bmatrix}. \tag{21}$$

We will apply the following standard theorem.

Theorem 1 *(Liptser and Shiryayev, 1977) Consider an unobservable process u_t and an observable process s_t with dynamics given by*

$$du_t = [a_0(t, s_t) + a_1(t, s_t)u_t]\, dt + b_1(t, s_t)dZ_t^u + b_2(t, s_t)dZ_t^s \tag{22}$$
$$ds_t = [A_0(t, s_t) + A_1(t, s_t)u_t]\, dt + B_1(t, s_t)dZ_t^u + B_2(t, s_t)dZ_t^s. \tag{23}$$

All the parameters can be functions of time and of the observable process. Liptser and Shiryayev (1977) show that the filter evolves according to (we drop the dependence of coefficients on t and s_t for notational convenience):

$$d\widehat{u}_t = (a_0 + a_1\widehat{u}_t)dt + [(b \circ B) + \zeta_t A_1^\top](B \circ B)^{-1}[ds_t - (A_0 + A_1\widehat{u}_t)dt] \tag{24}$$
$$\frac{d\zeta_t}{dt} = a_1\zeta_t + \zeta_t a_1^\top + (b \circ b) - [(b \circ B) + \zeta_t A_1^\top](B \circ B)^{-1}[(b \circ B) + \zeta_t A_1^\top]^\top, \tag{25}$$

where ζ_t is the posterior variance (or the Bayesian uncertainty) about u_t and

$$b \circ b = b_1 b_1^\top + b_2 b_2^\top \tag{26}$$
$$B \circ B = B_1 B_1^\top + B_2 B_2^\top \tag{27}$$
$$b \circ B = b_1 B_1^\top + b_2 B_2^\top. \tag{28}$$

In our setup, we obtain

$$b \circ b = \sigma_\mu^2 \tag{29}$$

$$B \circ B = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \frac{\Phi+1}{\Phi} \end{bmatrix} \tag{30}$$

$$b \circ B = \begin{bmatrix} 0 & \sigma_\mu \end{bmatrix}, \tag{31}$$

and

$$[(b \circ B) + \zeta_t A_1^\top](B \circ B)^{-1} = \begin{bmatrix} \frac{\lambda \zeta_t}{\sigma_\theta^2} & \frac{\sigma_\mu \Phi}{1+\Phi} \end{bmatrix}. \tag{32}$$

The Bayesian uncertainty ζ_t follows the deterministic process

$$\frac{d\zeta_t}{dt} = \frac{\sigma_\mu^2}{1+\Phi} - 2\eta\zeta_t - \frac{\lambda^2 \zeta_t^2}{\sigma_\theta^2}. \tag{33}$$

We assume that enough time has passed by such that the Bayesian uncertainty has reached a steady state. This is a common assumption in the literature on incomplete information (e.g., Dumas, Kurshev, and Uppal, 2009), and it fits well in our model with infinite horizon. The steady-state value solves $d\zeta_t = 0$. This is a quadratic equation with two solutions. Only one of the solutions is positive:

$$\bar{\zeta} = \frac{\sigma_\theta^2}{\lambda^2} \left(\sqrt{\eta^2 + \frac{1}{1+\Phi} \frac{\lambda^2 \sigma_\mu^2}{\sigma_\theta^2}} - \eta \right) \tag{34}$$

The steady-state uncertainty is a decreasing function of Φ and goes to zero as $\Phi \to \infty$ (in this case, μ_t is perfectly observable).

Using (24) and (32) and replacing $\zeta_t = \bar{\zeta}$, we can write

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \begin{bmatrix} \frac{\lambda}{\sigma_\theta^2}\bar{\zeta} & \sigma_\mu \frac{\Phi}{1+\Phi} \end{bmatrix} \begin{bmatrix} d\theta_t - \lambda(\widehat{\mu}_t - \theta_t)dt \\ ds_t \end{bmatrix} \tag{35}$$

The (observable) process θ can be written in two ways:

$$d\theta_t = \lambda(\mu_t - \theta_t)dt + \sigma_\theta dW_t^\theta \tag{36}$$

$$d\theta_t = \lambda(\widehat{\mu}_t - \theta_t)dt + \sigma_\theta d\widehat{W}_t^\theta \tag{37}$$

The first equation is written under the physical (true) probability measure. The second equation is written under the filtration of the firm, and $\widehat{W}_t^\theta$ is a standard Brownian motion under this filtration (intuitively, the second equation shows how the firm interprets the dynamics of the observable process θ). From these two equations, we obtain:

$$d\theta_t - \lambda(\widehat{\mu}_t - \theta_t)dt = \sigma_\theta d\widehat{W}_t^\theta \tag{38}$$

Furthermore, we can write the signal as

$$ds_t = dW_t^\mu + \frac{1}{\sqrt{\Phi}} dW_t^s = \sqrt{\frac{\Phi+1}{\Phi}} d\widehat{W}_t^s, \tag{39}$$

where $\widehat{W}_t^s$ is a standard Brownian motion independent of $\widehat{W}_t^\theta$. This leads to

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \begin{bmatrix} \frac{\lambda}{\sigma_\theta}\bar{\zeta} & \sigma_\mu \sqrt{\frac{\Phi}{1+\Phi}} \end{bmatrix} \begin{bmatrix} d\widehat{W}_t^\theta \\ d\widehat{W}_t^s \end{bmatrix}, \tag{40}$$

which, after replacement of (34), yields:

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \frac{\sigma_\theta}{\lambda}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right) d\widehat{W}_t^\theta + \sigma_\mu \sqrt{\frac{\Phi}{1+\Phi}} d\widehat{W}_t^s. \tag{41}$$

Notice that from (36)-(37) we can write:

$$d\widehat{W}_t^\theta = dW_t^\theta + \frac{\lambda}{\sigma_\theta}(\mu_t - \widehat{\mu}_t)dt, \tag{42}$$

and we also have from (39):

$$d\widehat{W}_t^s = \sqrt{\frac{\Phi}{1+\Phi}} ds_t. \tag{43}$$

We can therefore write Proposition 1.

From Proposition 1, the conditional variance of the filter $\widehat{\mu}_t$ is

$$\mathrm{Var}_t[\widehat{\mu}_t] = \sigma_\mu^2 - \frac{2\eta\sigma_\theta^2}{\lambda^2}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{44}$$

We can then compute

$$\frac{\partial \,\mathrm{Var}_t[\widehat{\mu}_t]}{\partial \sigma_\mu} = 2\sigma_\mu \left(1 - \frac{\eta\sigma_\theta}{(1+\Phi)\sqrt{\eta^2\sigma_\theta^2 + \frac{\lambda^2\sigma_\mu^2}{1+\Phi}}}\right) > 0 \tag{45}$$

and

$$\frac{\partial \operatorname{Var}_t[\widehat{\mu}_t]}{\partial \Phi} = \frac{\eta\sigma_\theta\sigma_\mu^2}{(1+\Phi)^2\sqrt{\eta^2\sigma_\theta^2 + \frac{\lambda^2\sigma_\mu^2}{1+\Phi}}} > 0, \tag{46}$$

which leads to Corollary 1.1.

C.2 Discretization used for simulations

The following processes are simulated under the filtration of the firm:

$$\begin{aligned}
\text{cash flow:} \quad & d\theta_t = \lambda(\widehat{\mu}_t - \theta_t)dt + \sigma_\theta d\widehat{W}_t^\theta & (47)\\
\text{(scaled) signal:} \quad & d\widehat{W}_t^s & (48)\\
\text{filter:} \quad & d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \Omega d\widehat{W}_t^\theta + \sigma_\mu\sqrt{\frac{\Phi}{1+\Phi}}d\widehat{W}_t^s & (49)\\
\text{adj. costs shocks:} \quad & d\nu_t = -\kappa\nu_t dt + \sigma_\nu dW_t^\nu, & (50)
\end{aligned}$$

where we define

$$\Omega \equiv \frac{\sigma_\theta}{\lambda}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{51}$$

Once we have simulated the four time-series above, we compute $q(\theta_t, \widehat{\mu}_t, \nu_t)$ for each simulated point. Then, we use the first order condition for investment to compute the investment-capital ratio for each simulated point:

$$\frac{I_t}{K_t} = 1\frac{1}{a} + \frac{1}{a}q(\theta_t, \widehat{\mu}_t, \nu_t) - \frac{1}{a}\nu_t, \tag{52}$$

which provides all the data necessary for the regressions. We implement the following discretization of the continuous-time processes (47)-(50):

$$\nu_{t+\Delta} = \nu_t e^{-\kappa\Delta} + \sigma_\nu\sqrt{\frac{1-e^{-2\kappa\Delta}}{2\kappa}}dW_t^\nu. \tag{53}$$

$$\widehat{\mu}_{t+\Delta} = \widehat{\mu}_t e^{-\eta\Delta} + \bar{\mu}\left(1 - e^{-\eta\Delta}\right) + \sqrt{\frac{1-e^{-2\eta\Delta}}{2\eta}}\left(\Omega d\widehat{W}_t^\theta + \sigma_\mu\sqrt{\frac{\Phi}{1+\Phi}}d\widehat{W}_t^s\right). \tag{54}$$

$$\theta_{t+\Delta} = \theta_t e^{-\lambda\Delta} + \widehat{\mu}_t\left(1 - e^{-\lambda\Delta}\right) + \sigma_\theta\sqrt{\frac{1-e^{-2\lambda\Delta}}{2\lambda}}d\widehat{W}_t^\theta. \tag{55}$$

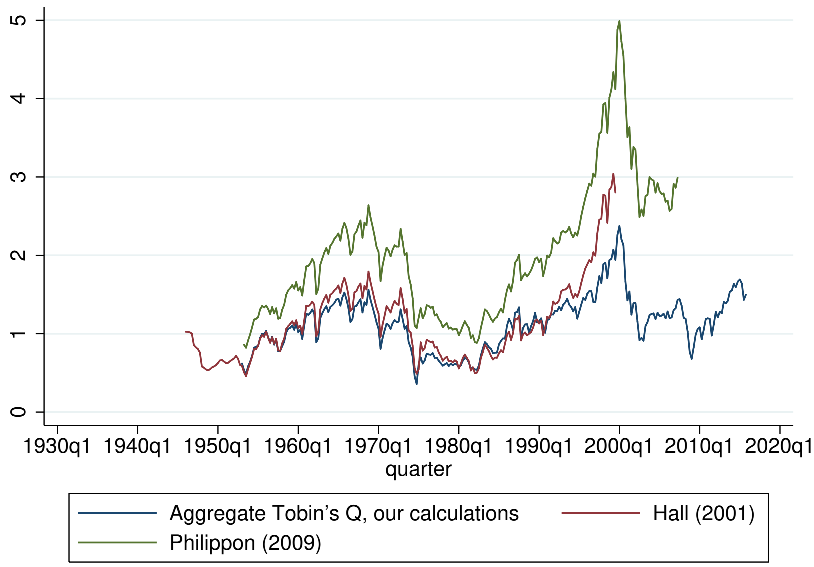

Figure 12: Plot of our series for aggregate Tobin's q against those from Hall (2001) and Philippon (2009).

	(1)		
	Tobin's Q	Hall '01	Philippon '09
Tobin's Q	1		
Hall '01	0.950***	1	
Philippon '09	0.943***	0.989***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Correlation matrix of our series for aggregate Tobin's q against those from Hall (2001) and Philippon (2009).

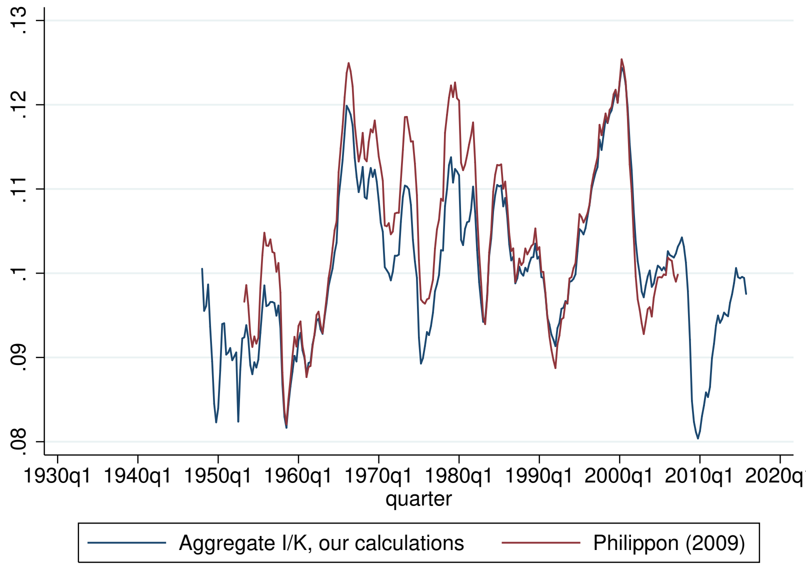

Figure 13: Plot of our series for investment against the series from Philippon (2009).

	(1)	
	Investment rate	Philippon '09
Investment rate	1	
Philippon '09	0.940***	1

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Correlation matrix of our series for investment against the series from Philippon (2009).