Learning and the improving relationship between investment and q^*

Daniel Andrei
UCLA

William Mann
UCLA

Nathalie Moyen
Colorado

March 8, 2018

Abstract

We show that the relationship between aggregate investment and Tobin's q has become remarkably tight in recent years, contrasting with earlier times. We connect this change with the growing empirical dispersion in Tobin's q, which we document 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 greater stock of intangible assets. We confirm the model's predictions in the data, and we disentangle our learning mechanism from measurement error in q.

*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, as well as Tor-Erik Bakke (2018 MFA discussant) and presentation participants at the 2018 Midwest Finance Association Annual Meeting, 2018 Workshop on Asset Pricing in Zurich, Emory University, UCLA, University of Colorado Boulder, Vanderbilt University, and Washington University in St. Louis.

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 marginal q with a valuation ratio, average q or Tobin's q, so that a simple regression of investment on Tobin's q should have a strong fit. Researchers have found that this regression in fact performs quite poorly. While the Hayashi model assumptions may not hold exactly in the data, a stark disconnect between investments and valuations is deeply puzzling to financial economists. A large literature investigates the potential reasons why Tobin's q does not explain investment 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 1995–2015, comparable to the empirical performances of the bond price q regression proposed in Philippon (2009) and the total tangible and intangible investment-q regression in Peters and Taylor (2017). 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 (2017) focuses on the failure to measure intangible assets, which have grown substantially in the aggregate, and Philippon (2009) focuses on measuring q via bond markets to avoid relying on equity market valuations, which are increasingly volatile and may seem unreliable. We show that, counterintuitively, it is precisely the growing volatility in valuations, *especially* in intangible-intensive industries, that has contributed to the revived empirical performance of the classic regression.

To explain these recent developments, we propose a learning-based model of corporate investment. The model endogenously produces more variation in marginal

[1]For examples, see Fazzari, Hubbard, and Petersen (1988), Gilchrist and Himmelberg (1995), Kaplan and Zingales (1997), Erickson and Whited (2000), Gomes (2001), Cooper and Ejarque (2003), Moyen (2004), Philippon (2009), Eberly (2011), and Peters and Taylor (2017).

q. This learning-induced variation is informative about the firm's investment policy. The main result is that 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 affecting investment.

To motivate the intuition empirically, we establish several stylized facts. First, the volatility of aggregate q in the data is higher precisely during 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. 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-flow growth. Specifically, we assume that the expected cash-flow growth evolves over time, and the firm can never fully learn. We then allow the firm to acquire, at a cost, informative signals about the time-varying cash-flow growth. These features provide a theoretical foundation for the stochastic variation in marginal q under incomplete information. We show that learning endogenously amplifies the volatility of marginal q, thereby improving the fit of the investment-q regression.

Importantly, uncertainty and volatility in the firm's valuation are distinct concepts. When signals are more informative, there is lower uncertainty about future cash-flow growth at any given moment; but by the same token, beliefs become more volatile as they are more aggressively updated from one moment to the other. Furthermore, we show that firms endogenously choosing more informative signals are the ones facing greater fundamental uncertainty about their cash flows. Thus, more uncertain environments lead to more learning, which in turn endogenously lowers a firm's cash-flow uncertainty but amplifies volatility in Tobin's q. The learning-induced volatility is not noise, but rather is statistically informative about the firm's investment policy.

An 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 (2017). 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 have noticeably higher R^2 values in their investment-q panel regressions compared to those from the average industry. This stylized fact is documented in Peters and Taylor (2017) and 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. This learning mechanism works in the opposite direction as misspecification issues, which have been the focus in much of the prior research. When a theory calls for an important role for cash flow beyond the information captured by Tobin's q, excluding cash flow from the regression would create an omitted variable bias. With such a bias, the investment-q regression would 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.

To test the relative importance of the learning mechanism against this potential misspecification, we estimate two fixed-effects panel regressions within each industry: a regression of Tobin's q on cash flow, and the standard regression of investment on Tobin's q. We find that the R^2 values from these regressions are negatively correlated across industries. The investment-q regression fits best in industries where the q-cash flow regression fits the worst. This pattern supports the learning mechanism. While it does not rule out the possibility of misspecification, it does suggest that the empirical effects of the misspecification are outweighed by the learning mechanism.

In sum, we find that the classic q theory of investment works surprisingly well in

recent years, and counterintuitively it works best for firms with high volatilities in equity valuations, high levels of R&D investment, and low levels of tangibility. Our findings have several general implications. They suggest that learning-based models may be particularly well-suited to study corporate investment behavior. They also suggest that Tobin's q, "arguably the most common regressor in corporate finance" (Erickson and Whited, 2012), may be a better empirical proxy for the firm's investment opportunities than previously thought.

An empirical paper closely related to ours is Peters and Taylor (2017). They augment the simple investment-q regression by adjusting for intangible capital. While we report results using the classic definitions of investment and q in keeping with the previous literature, all results continue to hold when these quantities are adjusted for intangibles with the "total" investment and q series (results available on request).

Another related empirical paper is Gutiérrez and Philippon (2016). They 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 the correlations, which have improved in recent years. In the appendix, we extend our learning model to demonstrate how greater market power generates a lower Tobin's q slope and yet a higher investment-q regression R^2. Lindenberg and Ross (1981) and Cooper and Ejarque (2003) also examine competition-related implications on Tobin's q.

Our paper builds on a long theoretical literature investigating the q theory of investment.[2] The most closely related theory paper to ours is Abel (2017), to which we add a learning mechanism. In our model, as in Alti (2003), Moyen and Platikanov (2012), and Hennessy and Radnaev (2018), learning symmetrically and simultaneously occurs for the firm and the market alike. In contrast to our setting, many papers study settings in which managers are assumed to possess superior information compared to outsiders (e.g. Myers and Majluf, 1984), or the reverse channel in which managers extract information from their stock prices when making investment decisions.[3] We choose the simple, symmetric learning framework as it is powerful enough

[2]Foundational contributions to Tobin's q theory originate from Keynes (1936), Brainard and Tobin (1968), Tobin (1969), Mussa (1977), Lindenberg and Ross (1981), Abel (1983), and Salinger (1984), among many others.

[3]See Subrahmanyam and Titman (1999), Bresnahan, Milgrom, and Paul (1992), Dow and Gorton (1997), Goldstein and Guembel (2008), and Edmans, Goldstein, and Jiang (2015), among others.

to provide the empirical prediction of interest, namely that the fit of the investment-q regression improves when firms are more engaged in research and learning.

Finally, our paper complements the recent literature showing that financial markets have become more informative in recent years. Bai, Philippon, and Savov (2016) argue that the recent rise in price informativeness is due to greater information production in financial markets. Chen, Goldstein, and Jiang (2007) and Bakke and Whited (2010) document a stronger relationship between stock prices and investment for firms with more informative stock prices, whereas Dow, Goldstein, and Guembel (2017) demonstrate how the information production in financial markets can amplify business cycles. The theoretical models of Farboodi, Matray, and Veldkamp (2017) and Begenau, Farboodi, and Veldkamp (2017) describe how this rise in price informativeness affects capital allocation in the economy. In line with this growing body of evidence, we document a remarkable improvement in the relationship between investment and q in recent years, and point to learning as a plausible explanation for this trend.

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. Figure 1 plots and compares aggregate investment, and lagged aggregate Tobin's q, from 1975 to 2015. The figure is divided into two subperiods of 20 years each. At the bottom of each subperiod is the R^2 value that would be obtained from the standard regression of aggregate investment rate on lagged q using only the data from that subperiod. This regression is specified as

$$\frac{I_{t+1}}{K_t} = \alpha + \beta q_t + \epsilon_t, \tag{1}$$

where t indexes quarters, I is aggregate private nonresidential fixed investment, K is aggregate gross stock of private nonresidential fixed assets, and q is measured as the ratio of corporate financial value (value of outstanding equity and debt securities less inventories) to the aggregate stock of corporate nonresidential fixed assets. To construct these series, we use quarterly data from the Fed Flow of Funds and from NIPA tables, following the steps described in Hall (2001) and Philippon (2009). More details are provided in Appendix A.1.

[Figure 1 here]

The R^2 from regression (1) has been the primary focus of the empirical literature assessing the performance of the q theory of investment. During the first subperiod, 1975-1995, the relationship between aggregate investment and Tobin's q is disappointingly weak, and the standard regression achieves an R^2 of only 8%. This fact has been widely confirmed, e.g., Philippon (2009), Table III (top panel, second column). As a result, modern empirical research often describes the investment-q regression as an empirical failure. Many papers attempt to improve the classic regression in various ways, as discussed above. But in the second subperiod, 1995-2015, the investment-q regression performs much better. From 1995-2015, the R^2 is nearly 70%. Looking only at the more recent past, 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 solid blue and dashed red series are the year-over-year differences of the series from Figure 1. The R^2 values from the regression within the two 20-year subperiods are listed at the bottom of the figure, and they suggest the same conclusion as in Figure 1. The R^2 of the investment-q regression rose from less than 1% in 1975-1995 to greater than 48% in 1995-2015.

[Figure 2 here]

Also listed at the bottom of each subperiod in Figure 2 are the volatilities of the explanatory variable in the regression, differenced Tobin's q. These figures provide motivating evidence for our core mechanism. The volatility of Tobin's q is lower during the subperiod in which the investment-q regression performs worse, and it is higher during the subperiod in which the regression performs better.

Under the null hypothesis that the model is true, the investment-q regression should yield a higher R^2 when there is more dispersion in the key explanatory variable, Tobin's q. Thus, based on our results, 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 the next section, we use panel data from Compustat to examine more closely the empirical dispersion in Tobin's q and establish this pattern.

2.2 Increased dispersion in Tobin's q

Shifting our focus from the aggregate series discussed above, we next reconstruct the series of investment and valuation at the firm level. Using annual panel data on publicly-traded firms from Compustat, we confirm and explore the growing empirical dispersion in Tobin's q.

We decompose the volatility of Tobin's q along two dimensions: first between-firm, 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 Tobin's q in Compustat, then we smooth it by simple averaging over a rolling five-year lag in order to focus on the trend. The cross-sectional dispersion has trended upward over time. The standard deviation of Tobin's q has risen from less than 5 during the 1980s and 1990s, to over 10 during the 2000s.

[Figure 3 here]

In Figure 4, we investigate how Compustat firms have changed over time by examining the within-firm volatility. To create this figure, we proceed in two steps. First, we calculate for each Compustat firm the within-firm volatilities of Tobin's q during its entire lifetime in Compustat. This creates volatility measures of valuation 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 here]

Figure 4 reveals that within-firm volatilities of Tobin's q have greatly increased relative to their 1980 values. Tobin's q volatility has risen from less than 1 to over 5. The increase in Tobin's q volatility is especially noticeable in the late 1990s and early

2000s. The series thus show that the composition of Compustat has shifted towards firms with higher volatilities.

2.3 Better performance for firms with more volatile q

Motivated by the Compustat evidence above which shows that firms are exhibiting greater volatility of q, we next demonstrate that the investment-q regression works better where this volatility is greater.

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 0.25, while in the highest bin it is 12.53.[4]

The bin with the highest-volatility firms is where we find that the investment-q relationship is the tightest. To show this, we estimate standard panel regressions of investment on lagged Tobin's q:

$$\frac{I_{i,t+1}}{K_{it}} = \alpha_i + \beta q_{it} + \epsilon_{it}, \tag{2}$$

where i indexes firms, t indexes years, I is capital expenditures, K is gross property, plant, and equipment, q is defined as $\frac{V}{K}$, where V is the market value of equity plus book value of debt minus current assets. All of these definitions are taken from Peters and Taylor (2017).

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

[Table 1 here]

Figure 5 visually illustrates how the volatility in the data gives rise to these results. It samples 500 observations randomly from each of the lowest and the highest bins of volatility, 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 lowest-volatility bin shows no particular relationship between q and the investment rate, while the highest-volatility bin illustrates a fairly tight relationship.

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

[Figure 5 here]

In Table 1 and Figure 5, it may seem puzzling that the slope of the regression falls across bins of Tobin's q volatility, even as the R^2 increases. One explanation for a falling slope could be that the higher-volatility firms have market power. As argued in Cooper and Ejarque (2003), market power represented by decreasing returns to scale causes the slope of the regression to fall due to measurement error when using average q to proxy for marginal q.[5] R&D-intensive firms are a natural example of this pattern because they are often characterized by volatile valuations and market power gained from innovation. Mathematically, to reconcile the falling slope with the increase in R^2, one needs an offsetting source of higher volatility in Tobin's q. Our analysis shows that learning is a mechanism for generating the volatility. In the appendix, we simulate a model with both learning and decreasing returns to scale, and are able to reproduce the pattern in Table 1 and Figure 5. However, in the next section we present our benchmark learning model without decreasing returns to scale and therefore without measurement error.

If the large volatility in q is meaningless for investment, the improving fit of the investment-q regression should not obtain. 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 7.7, but the R^2 pattern across the regressions remains the same as in the previous table.

[Table 2 here]

In untabulated results, we observe the effect being even more dramatic when we winsorize investment as well. Under this approach, the R^2 of the fourth bin regression reaches over 20%. Further results show that the pattern is also robust to adding year fixed effects; to excluding all fixed effects; and to sorting on the stock price

[5] This measurement error is correlated with marginal q. It is therefore not suitable for the estimator of Erickson and Whited (2000), which assumes that the measurement error is independent of q.

volatility instead of Tobin's q volatility, confirming that higher volatility comes from the numerator of q, not its denominator.

Finally, we connect these cross-sectional patterns with the aggregate time-series patterns. In Figure 6, we separate the Compustat panel into two subperiods, 1975–1995 and 1995–2015, as we did in Figures 1 and 2. We redefine the four bins of Tobin's q volatility separately within these two subperiods, and report the R^2 from the panel regression within each bin and subperiod. The figure reproduces the stylized patterns documented so far. Within both subperiods, the fit of the regression increases steadily across the bins of volatility in Tobin's q. At the same time, the line for 1995-2015 is uniformly higher than the line for 1975-1995. These figures suggest that there have always been some firms for which the investment-q regression was tighter due to greater dispersion in Tobin's q, and that these firms have become more important in the aggregate in recent years.

[Figure 6 here]

In sum, the stylized facts discussed in this section demonstrate 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 below, we rationalize these facts with a learning model that explains why the types of firms appearing in the data in more recent years are likely to exhibit a tighter relationship between their investments and 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

$$dK_t = (I_t - \delta K_t)dt, \tag{3}$$

where I_t denotes the firm's investment decision.

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

$$\psi(I_t, K_t, \nu_t) = \frac{a}{2}\left(\frac{I_t}{K_t}\right)^2 K_t + \nu_t I_t, \tag{4}$$

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

$$d\nu_t = -\kappa\nu_t dt + \sigma_\nu dW_t^\nu. \tag{5}$$

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

$$\Pi(K_t, \theta_t) = \theta_t K_t, \tag{6}$$

where we use the output price as numéraire. Without loss of generality, we abstract from describing the flexible labor decision.[6]

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

$$d\theta_t = \lambda(\mu_t - \theta_t)dt + \sigma_\theta dW_t^\theta. \tag{7}$$

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 unobservable long-term mean μ_t evolves stochastically as described below.

[6]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 $\max_{L_t}\left[A_t L_t^\alpha K_t^{1-\alpha} - L_t\right] = \underbrace{(1-\alpha)\alpha^{\frac{\alpha}{1-\alpha}} A_t^{\frac{1}{1-\alpha}}}_{\equiv\theta_t} K_t \equiv \Pi(K_t, \theta_t)$.

3.2 Learning

The long-term mean around which θ_t evolves, μ_t, is not observable, and it follows a mean-reverting process

$$d\mu_t = \eta(\bar{\mu} - \mu_t)dt + \sigma_\mu dW_t^\mu. \tag{8}$$

For simplicity of exposition, we fix $\mu_0 = \bar{\mu}$. In the special case with $\sigma_\mu = 0$, the long-term mean would be observable, and the firm could choose how much to invest at each point in time knowing all the necessary information. However, as soon as $\sigma_\mu > 0$, the firm needs to update continuously its beliefs about μ_t.

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 (7). 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

$$ds_t = dW_t^\mu + \frac{1}{\sqrt{\Phi}} dW_t^s, \tag{9}$$

where 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 by the firm.

The following proposition and its corollary obtain from filtering theory (Liptser and Shiryaev, 1977), with the proof provided in Appendix A.2.

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

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \frac{\sigma_\theta}{\lambda}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right) d\widehat{W}_t^\theta + \sigma_\mu\sqrt{\frac{\Phi}{1+\Phi}}d\widehat{W}_t^s, \tag{10}$$

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 in equation (9), such that $\widehat{W}_t^s$ is a standard Brownian motion.

The Bayesian uncertainty, defined as $\zeta_t \equiv \mathbb{E}[(\mu_t - \widehat{\mu}_t)^2 \mid \mathcal{F}_t]$ where $\mathcal{F}_t$ is the

information set of the firm at time t, follows the deterministic process

$$\frac{d\zeta_t}{dt} = \frac{\sigma_\mu^2}{1+\Phi} - 2\eta\zeta_t - \frac{\lambda^2\zeta_t^2}{\sigma_\theta^2}. \tag{11}$$

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 A.2).

We assume that enough time has passed 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 for ζ, $\bar{\zeta}$, is obtained by setting the right-hand side of equation (11) to zero. This yields a quadratic equation with only one positive root:

$$\bar{\zeta} = \frac{\sigma_\theta^2}{\lambda^2}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{12}$$

Because learning in the model is constantly regenerated, the steady-state uncertainty is positive. It is increasing in σ_μ, decreasing in Φ, and goes to zero only in the limiting case as $\Phi \to \infty$ (when μ_t becomes perfectly observable).

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

$$\mathrm{Var}_t[\widehat{\mu}_t] = \sigma_\mu^2 - 2\eta\bar{\zeta}, \tag{13}$$

is strictly increasing in both σ_μ and Φ.

According to Corollary 1.1, the conditional variance of the filter increases when there is more uncertainty about the long-term mean μ_t or when the firm acquires information through a more informative signal Φ. Although the filtered long-term mean $\widehat{\mu}_t$ is less volatile than the truth μ_t (because the filter is a projection of μ_t on the observation filtration of the firm), Corollary 1.1 shows that learning with a more

informative signal Φ strictly increases the volatility of the filter itself. In the limit when $\Phi \to \infty$, the firm perfectly observes μ_t and the conditional variance of the filter reaches the conditional variance of the unobserved process, σ_μ^2.

Furthermore, Corollary 1.1 shows that learning affects uncertainty and volatility in opposite ways: learning with more informative signals (higher Φ) decreases the uncertainty $\bar{\zeta}$ in equation (12) but it increases the conditional volatility of the filter $\text{Var}_t[\widehat{\mu}_t]$ in equation (13). In other words, although the firm decreases uncertainty through learning at any moment in time, its beliefs become more volatile as they are updated from one moment to the other.

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 taking place does not change the conditional volatility of the cash-flow process (7) itself, which remains constant at σ_θ for any level of σ_μ. Learning, however, does increase 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,

$$V(K_t, \theta_t, \widehat{\mu}_t, \nu_t) = \max_I \mathbb{E}_t \left[\int_t^\infty e^{-r(s-t)} \left\{ \theta_s K_s - I_s - \psi(I_s, K_s, \nu_s) \right\} ds \right], \tag{14}$$

subject to equations (3) and (4), where r is the interest rate. The information set of the firm at time t is summarized by the capital stock K_t, the cash flow θ_t, the conditional expectation of cash-flow growth $\widehat{\mu}_t$, and the shock to the purchase price of capital ν_t.

The Hamilton-Jacobi-Bellman equation associated with problem (14) is

$$rV = \max_{I} \left\{ \theta K - I - \psi(I, K, \nu) + \mathcal{D}V(K, \theta, \widehat{\mu}, \nu) \right\}, \tag{15}$$

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

$$0 = V_K(K, \theta, \widehat{\mu}, \nu) - 1 - \psi_I(I, K, \nu). \tag{16}$$

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

$$V(K, \theta, \widehat{\mu}, \nu) = q(\theta, \widehat{\mu}, \nu)K. \tag{17}$$

Replacing the adjustment cost function (4) yields the following relationship between the rate of investment and q:

$$\frac{I_t}{K_t} = -\frac{1}{a} + \frac{1}{a} q(\theta_t, \widehat{\mu}_t, \nu_t) - \frac{1}{a}\nu_t. \tag{18}$$

Using equation (17) and solving for the optimal investment, we obtain the following partial differential equation for q (where q_x denotes the partial derivative of q with respect to the state variable x):

$$\begin{aligned} 0 = \theta_t + \frac{(1+\nu_t)^2}{2a} - \frac{1 + a(r+\delta) + \nu_t}{a} q + \lambda(\widehat{\mu}_t - \theta_t) q_\theta + \eta(\bar{\mu} - \widehat{\mu}_t) q_{\widehat{\mu}} - \kappa \nu_t q_\nu \\ + \frac{\sigma_\theta^2}{2} q_{\theta\theta} + \left(\frac{\sigma_\mu^2}{2} - \eta\bar{\zeta} \right) q_{\widehat{\mu}\widehat{\mu}} + \frac{\sigma_\nu^2}{2} q_{\nu\nu} + \lambda\bar{\zeta} q_{\theta\widehat{\mu}} + \frac{1}{2a} q^2. \end{aligned} \tag{19}$$

We solve this equation numerically by approximating $q(\theta, \widehat{\mu}, \nu)$ with Chebyshev polynomials.[7]

[7] 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 (18) and, counterfactually, this relationship would always have an R^2 of one. The shock to the capital purchase price causes the R^2 to be below one:

$$R^2 = \frac{\text{Var}[q(\theta_t, \widehat{\mu}_t, \nu_t)] \left(1 - \frac{\text{Cov}[q(\theta_t, \widehat{\mu}_t, \nu_t), \nu_t]}{\text{Var}[q(\theta_t, \widehat{\mu}_t, \nu_t)]}\right)^2}{\text{Var}[q(\theta_t, \widehat{\mu}_t, \nu_t)] + \text{Var}[\nu_t] - 2\,\text{Cov}[q(\theta_t, \widehat{\mu}_t, \nu_t), \nu_t]}. \tag{20}$$

The R^2 coefficient increases with the variance of q as long as the covariance between q and ν is negligible.[8] Notice also that a stronger regression coefficient for q in equation (18) does not mechanically affect the R^2, since the adjustment cost parameter a simplifies away from (20).

The firm's learning affects the R^2. This can be seen from an application of Itô's lemma on $q(\theta_t, \widehat{\mu}_t, \nu_t)$:

$$dq = \xi_t dt + \left(q_\theta \sigma_\theta + q_{\widehat{\mu}} \frac{\lambda}{\sigma_\theta} \bar{\zeta}\right) d\widehat{W}_t^\theta + q_{\widehat{\mu}} \sigma_\mu \sqrt{\frac{\Phi}{1+\Phi}} d\widehat{W}_t^s + q_\nu \sigma_\nu dW_t^\nu, \tag{21}$$

where ξ_t denotes the drift (its specific form does not matter for our analysis). 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 brackets above. Tobin's q also becomes sensitive to information shocks $d\widehat{W}_t^s$ through the third term above. Both these effects increase the volatility of $q(\theta_t, \widehat{\mu}_t, \nu_t)$ and, according to equation (20), the R^2 of the investment-q regression.

We illustrate the impact of learning on the R^2 by means of simulations. To this end, we implement a discretization of the continuous-time processes at a yearly frequency (see Appendix A.3). We then solve for the partial differential equation (19) and compute q_t for each simulated point $\{\theta_t, \widehat{\mu}_t, \nu_t\}$. The resulting value for q_t can then be replaced in equation (18), yielding the investment rate I_t/K_t. This completes the dataset necessary for implementing investment-q regressions.

[8] The R^2 depends on the relationship between q_t and ν_t. In our numerical calibration, we ensure that the covariance between q_t and ν_t is virtually zero, i.e., $q_\nu \approx 0$. This occurs for large values of κ, i.e., when the persistence of ν_t is close to zero. A non-negligible persistence of ν_t creates temporal dependence through which q_t depends on ν_t. Even in this case, the covariance term in equation (20) is of small magnitude, and does not impact our main intuition.

Figure 7 plots as an example one simulation of 100 yearly observations. The horizontal axis in each panel is the marginal q. The vertical axis represents the optimal investment rate I_t/K_t. The calibration used is as follows: $a = 16$, $r = 3\%$, $\delta = 10\%$, $\lambda = 0.5$, $\sigma_\theta = 0.1$, $\kappa = 5000$, $\sigma_\nu = 40$, $\bar{\mu} = 0.25$, and $\eta = 0.5$.

[Figure 7 here]

The left panel corresponds to the case of an unobservable μ_t 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 in equation (9), with $\Phi = 20$. Changes in μ_t are not yet perfectly observed, but with $\Phi = 20$ the signal in (9) is more informative relative to the cash-flow signal in equation (7). 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 average R^2 coefficients obtained from 5,000 such simulations are 18% for the left panel, 48% for the middle panel, and 56% for the right panel. We also notice that the volatility of Tobin's q increases with learning: it averages 0.18 in the first panel, 0.39 in the second panel, and 0.45 in the third panel.

Although learning improves the R^2 of the investment-q regression, it does not influence its slope, which remains equal to $1/a$ across all models. This can be seen in Figure 7, where the fitted line remains the same in the three simulated panels. In contrast, in the data of Figure 5 the slope decreases across the bins. As we discuss in Appendix A.4, the decreasing slope can be attributed to decreasing returns to scale, consistent with market power from R&D innovations.

3.5 Endogenous learning

In this section, we endogenize the information acquisition problem. Specifically, the firm can purchase a more informative signal (higher Φ) to 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 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 $\tilde{V}(K_0, \theta_0, \mu_0, \nu_0)$, and its associated cost, $c(\Phi)$, is a strictly increasing and convex

function with $c'(0) = 0$. With Φ as a parameter in $\tilde{V}(K_0, \theta_0, \mu_0, \nu_0)$, the problem is equivalent to the earlier model without an endogenous Φ. The optimal Φ^* is defined by the first-order condition $\tilde{V}_\Phi(K_0, \theta_0, \mu_0, \nu_0) = c'(\Phi^*)$, and there is an interior solution if and only if $\tilde{V}_{\Phi\Phi}(K_0, \theta_0, \mu_0, \nu_0) - c''(\Phi^*) < 0$.

We are interested in investigating whether the optimal purchase of information increases with the uncertainty about the long-term mean. Differentiating the first-order condition with respect to σ_μ and rearranging, we get

$$\frac{d\Phi^*}{d\sigma_\mu} = \frac{\tilde{V}_{\Phi\sigma_\mu}(K_0, \theta_0, \mu_0, \nu_0)}{c''(\Phi^*) - \tilde{V}_{\Phi\Phi}(K_0, \theta_0, \mu_0, \nu_0)}. \tag{22}$$

The denominator is positive if the problem has an interior solution. The optimal amount of information acquisition Φ^* increases in σ_μ if and only if an increase in the uncertainty about μ_t increases the marginal benefit of purchasing information ($\tilde{V}_{\Phi\sigma_\mu}(K_0, \theta_0, \mu_0, \nu_0) > 0$).

The problem therefore reduces to showing that $\tilde{V}_{\Phi\sigma_\mu}(K_0, \theta_0, \mu_0, \nu_0) > 0$. While there is no closed-form proof of this, it can be checked numerically as $\tilde{V}(K_0, \theta_0, \mu_0, \nu_0)$ is just the value function from the problem without an endogenous signal informativeness choice Φ. Figure 8 shows that this is indeed the case. The left panel plots the function $q(\theta_t, \widehat{\mu}_t, \nu_t)$, where Φ varies from 0 to 25 on the x-axis. Each line in the plot corresponds to a different value of $\sigma_\mu \in \{0.02, 0.08, 0.14\}$. In these plots, the state variables are fixed at $\theta_t = \widehat{\mu}_t = \bar{\mu}$ and $\nu_t = 0$, but the results remain the same with different values for the state variables. The right panel of Figure 8 presents the second derivative $q_{\Phi\sigma_\mu}$, which is approximated using finite difference.[9] This derivative is positive at all times, consistent with the optimal information acquisition level Φ^* increasing in σ_μ.

[Figure 8 here]

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

[9] In our model, the second derivative $q_{\Phi\sigma_\mu}$ has the same sign as $\tilde{V}_{\Phi\sigma_\mu}$.

and development.[10]

4 Empirical analysis of the model predictions

The model in the previous section demonstrates that the investment-q regression performs better in settings with learning, and that this effect is stronger among firms that endogenously acquire more information. In this section, we dig deeper into the empirical predictions of the model.

4.1 Better performance in high-tech industries

Section 3.5 contains the main prediction of the model, where the investment-q regression performs better among firms that endogenously choose to expend greater resources on information acquisition. Empirically, we are interested in identifying groups of firms where learning is most likely to take place. We focus on firms that decide to spend more on R&D. Our proposed learning mechanism should cause the investment-q regression to work better in industries featuring high investment in research. This insight provides testable cross-sectional implications of the model.

For an operational definition of a research-intensive industry, we use 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 refer to these as "research-intensive" or "high-tech" industries for the remainder of this paper. The industry classification follows Brown, Fazzari, and Petersen (2009), which shows that the seven industries account for nearly all the growth in aggregate R&D during the 1990s.

We build up our analysis of research-intensive industries in several layers. First, we examine the empirical distribution of Tobin's q in these industries compared to the average Compustat firm. Figure 9 displays the empirical density of Tobin's q for firm-years in the high-tech and other industries in the sample. For firms that spend more on research, the empirical distribution of Tobin's q is more skewed and more

[10]For simplicity, our model considers a purely cross-sectional learning decision, made one time in the firm's lifetime, rather than a dynamic decision. This allows us to connect our model to the cross-sectional distribution of firm-level R&D intensity, which is relatively stable, without introducing a second capital accumulation decision into the firm's problem.

dispersed.[11] In our model, this pattern arises because these firms' values are much more sensitive to the arrival of signals about future profitability.

[Figure 9 here]

Figure 10 calculates the within-firm volatility of Tobin's q for each firm, and plots the empirical density of this volatility, again separating out high-tech firms from the other industries. This volatility also follows a skewed distribution. It is higher on average for the high-tech firms than for the others.

[Figure 10 here]

It is not surprising that market valuations of high-tech companies are particularly volatile. What is less clear is that these fluctuations are highly predictive of investment, as expected under the q theory of investment. This contrasts with the alternative view that market value fluctuations arise from problems in measuring the firm's capital stock or from difficulties outsiders face in valuing the firm, which would make these fluctuations simply exogenous noise with respect to the firm's investment policy.

Tables 3 and 4 repeat the panel regressions of investment on lagged q with fixed effects, as specified earlier in equation (2) and implemented in Tables 1 and 2, where various columns separate out high-tech from other industries.

[Tables 3 and 4 here]

Columns 1 and 2 of Table 3 show that the standard investment-q panel regression fares better among high-tech firms: The R^2 value from the regression doubles from 11% to 22% when we move from the non-tech 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 the sample to years prior to 1995. The same discrepancy holds for

[11] The figure includes a mass of negative values of Tobin's q. Negative values are possible when subtracting out current assets from the numerator, which we do for consistency with Peters and Taylor (2017). Our analysis is qualitatively unaffected when not subtracting current assets, in which case there is no negative value for q.

these early years: the R^2 of the panel regression increases from under 7% for non-tech industries to over 21% for high-tech, in accord with the investment-q regression working better for high-tech industries.

Table 4 checks robustness to some alternative approaches. Column 1 shows that firm fixed effects are not driving the performance of the regression, as the (overall) R^2 from the pooled regression is similar to the (within) R^2 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 of the other papers in the Tobin's q literature. In untabulated results (available on request) we document further robustness to various combinations of fixed effects and approaches to winsorizing.

Columns 3 and 4 return to our main panel specification with firm fixed effects but no time fixed effect, and adds in annual R&D expense plus 30% of annual SG&A expense as a measure of intangible investment, following Peters and Taylor (2017). The conclusion remains the same as before: the regression works better in high-tech industries ($R^2 = 20\%$) than in other industries ($R^2 = 8\%$).

To show that the results extend beyond the coarse high-tech proxy, we examine R^2 values using more general measures of R&D intensity. In Figure 11, we examine evidence at the firm level. We sort firms in the Compustat panel into six bins based on their average R&D intensity (defined as the ratio of annual R&D to total assets) within their lifetime in Compustat. Within each bin, we estimate the panel investment-q regression. The figure plots the R^2 values obtained from the regression in each bin. These values show an increasing pattern, from about 10% in the lowest bin to over 25% in the highest bin.

[Figure 11 here]

In Figure 12, we perform a similar exercise at the industry level. We perform the investment-q regression separately within each 3-digit SIC industry in Compustat, and plot the resulting R^2 values against average R&D intensity calculated across firm-years in that industry. The figure shows a clear positive association. The high-tech industries from the previous analysis are marked with an "$\times$" in the figure. They are clustered near each other at high values of R&D intensity, and relatively high R^2 values, although not the highest observed R^2 across all industries.

[Figure 12 here]

The finding that the investment-q regression works better in high-tech industries was previously established in Peters and Taylor (2017). In their Section 5.1, Peters and Taylor use several cross-sectional proxies beyond the simple industry classification to explore a number of explanations for this fact, but ultimately reject all of them. They conclude: *"Why the classic q-theory fits the data better in high-intangible settings is also an interesting open question."* Our learning-based model of corporate investment provides a plausible explanation for this finding.

The growth of high-tech industries is key 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 13 shows that the firms in the high-tech industry classification represent a growing fraction of the number of firms and of book assets in Compustat. Similarly, Peters and Taylor (2017) show that their measure of intangible capital, which capitalizes past intangible investments such as R&D and SG&A, also increases over time in both Compustat and the aggregate data from the Fed Flow of Funds.

[Figure 13 here]

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 endogenous swings in valuations and investments.

4.2 Better performance with low cash flow-q correlation

We next explore a subtler implication of the model relating to the role of cash flows. The model predicts that the investment-q regression works better in settings where Tobin's q is less correlated with cash flow. This is because, in the model, q is less responsive to cash flow when the firm chooses to learn from other signals as well. The combination of both signals causes Tobin's q to be more informative for the firm's investment decisions, improving the fit of the investment-q regression.

This learning mechanism works in the opposite direction as misspecification issues. Consider the misspecification effect of omitting cash flow from the regression, when an

alternative theory (e.g., based on financial constraints) would predict that cash flow is an important variable. The resulting omitted variable bias would be smaller, and therefore the investment-q regression should work better, in settings where Tobin's q is *more* correlated with cash flow, because there is less empirical content in cash flow separate from q.

Building on this intuition, we test the relative importance of the learning mechanism against potential misspecification. We separate Compustat firms by industry at the 3-digit SIC code, and we investigate how the tightness of the fit between cash flow and Tobin's q is related to the tightness of the fit between investment and q. Within each industry, we estimate fixed-effects regressions of cash flow on lagged q, then of investment on lagged q. We save the R^2 values from both of these regressions for each industry, and plot them in Figure 14.

[Figure 14 here]

The pattern in the figure lends support to the learning mechanism. The industries with the tightest connection between q and investment (the highest values on the y-axis) are also the industries with the weakest connection between q and cash flow (the left-most values on the x-axis). Conversely, the industries with the tightest connection between q and cash flow are also the industries with the weakest connection between q and investment.

The overall pattern is contrary to what we would expect with a misspecification problem. If a variable (e.g., cash flow) is omitted from the regression but it is actually an important predictor of investment, the R^2 from the incorrect specification should *increase*, not *decrease*, when the included (q) and omitted (cash flow) variables are more highly correlated, because the omitted variable does not contain as much independent information. While our findings do not indicate that there is no misspecification in our model, i.e., cash flow may well be an important regressor in the true model, our findings do suggest that the empirical effects of the misspecification are outweighed by the learning mechanism.

4.3 Disentangling learning from measurement error in q

Our benchmark learning model does not include measurement error. In our setting, marginal q is always equal to average q. Empirically, however, Tobin's average q may

be a poor proxy for a number of reasons, such as measurement error in the firm's capital stock. In this section, we empirically disentangle our learning mechanism from measurement error.[12]

We draw on the large literature on measurement error in Tobin's q. Two contributions are especially relevant to our work. First, Erickson and Whited (2000) develop an estimator that is robust to measurement error by exploiting identifying information in the third- and higher-order moments of the empirical distribution of Tobin's q. Erickson, Jiang, and Whited (2014) improve on this approach by focusing on cumulants rather than moments. These approaches yield, among other things, estimates of two population R^2 values: First, the R^2 from the measurement regression of Tobin's average q on "true" marginal q, labeled τ^2; and second, the R^2 from the investment regression of investment rate on "true" marginal q, labeled ρ^2. We use these parameter estimates to quantify the cross-sectional importance of measurement error and the "true" performance of the q theory.

Second, Peters and Taylor (2017) focus on the role of intangibles, which are missing from the standard measurement of investment and average q. They propose to capitalize R&D and SG&A expenditures as intangible investments, and show that this approach improves the performance of the regression. The adjustment is largest for high-tech firms, for whom intangibles are relatively more important. We examine how the adjustment interacts with our learning mechanism in the cross-section.

As motivating evidence, we first examine the evolution of τ^2 and ρ^2 through time with and without the adjustment for intangibles. Figure 15 displays the time-series of τ^2 and ρ^2. The estimators of Erickson et al. (2014) are applied to rolling ten-year windows of Compustat data, using three cumulants to exactly identify the system.[13] The figure separately plots the series with (dashed lines) and without (solid lines) intangibles in the measures of investment and q.

[Figure 15 here]

First, consider the two series for τ^2, which are displayed in the left panel of Figure 15. These capture the degree of measurement error driving a wedge between average q and marginal q. A higher value of τ^2 corresponds to a greater R^2 in the measurement

[12] Throughout this section, we refer to measurement error that satisfies the identifying assumptions in Erickson and Whited (2000).

[13] Results are essentially unchanged if we use a greater number of cumulants to overidentify the system.

regression, and thus a lower degree of measurement error. In the early years, Figure 15 shows that the τ^2 values for total q and standard q are close together. This suggests that intangibles did not create a large amount of measurement error, consistent with the plots of aggregate intangible investment presented in Peters and Taylor (2017). In the later years, however, the two lines diverge, with the total-q intangible adjustment yielding a consistently better proxy for marginal q. Since the late 1990s, the quality of the standard q proxy has worsened, while the quality of total q has improved. As is well-known, intangibles are an increasingly important feature of the economy and accounting for them improves the measurement.

Second, consider the two series for ρ^2, which are displayed in the right panel of Figure 15. These capture the performance of the "true" investment regression, i.e., the regression of investment on the "true" marginal q. Using total investment, which includes R&D expenses and 30% of SG&A expenses as intangible investment, produces a consistently better fit than the standard investment with physical capital expenditures only. Assuming that the cumulant-estimator approach has addressed measurement error in Tobin's q, the discrepancy between total investment and standard investment is not driven by measurement error. Rather, it suggests that q theory also applies to intangible investment, and accounting for intangibles improves the empirical performance of the q theory.

Most importantly for our purpose, the ρ^2 values for both investment measures have trended upwards over time. Again, under the identifying assumptions of Erickson et al. (2014), this is not due to measurement error. Instead, it reflects the fact that the explanatory power of "true" marginal q on investment has improved, consistent with the motivating evidence from Figures 1 and 2. The improved fit of the regression is the main prediction of our learning model. Figure 15 empirically summarizes the importance of measurement error vis-à-vis the improving investment-q relationship.

Figure 16 explores the ρ^2 evidence cross-sectionally. The figure distinguishes between two different groups of firms. The gray bars belong to firms with below-median volatility of Tobin's q; and the black bars, to firms with above-median volatility. In the left pair of bars, this sorting is performed using standard measures of investment and Tobin's q, while in the right pair of bars, the sorting is performed using the total-investment and total-q adjustments.

[Figure 16 here]

The figure demonstrates that the R^2 fit remains much higher among firms with higher volatility of Tobin's q, even after correcting for the measurement error using Erickson et al. (2014). This pattern holds for both standard and total q.

Figure 17 presents a similar set of ρ^2 values, where firms are separated into high-tech and other industries as in Table 3, rather than sorted on the volatility of q. Regardless of whether standard or total q is used to proxy for marginal q, the fit of the q theory is higher among high-tech firms than in other industries.

[Figure 17 here]

For completeness, Figures 18 and 19 plot the τ^2 values for the cross-sections of the prior two figures. In the figures, we see that the total-q adjustment is particularly effective at addressing measurement error for high-tech firms and firms with high volatility of Tobin's q.

[Figure 18 here]

[Figure 19 here]

In sum, the investment-q regression fits better among high-tech industries and those with high volatility in Tobin's q, even after adjusting for measurement error in Tobin's q as in Erickson et al. (2014), and after accounting for intangibles as in Peters and Taylor (2017). This suggests that the empirical support for our model operates through a better fit of the true regression of investment on marginal q, not through differences in measurement error.

5 Conclusion

This paper is motivated by the empirical finding 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 these patterns with a learning-based model of corporate investment. Learning by firms endogenously produces more variation in marginal q, improving

the fit of the regression. The learning mechanism is relevant especially in research-intensive industries. Thus, the improved fit of the investment-q relationship is related to the substantial growth in expenditures on research and other intangibles in the aggregate. We investigate the model's predictions in the cross-section of firms in Compustat, and find empirical support for our learning mechanism.

In conclusion, even a simple version of the q theory of investment can describe the data quite well, when given sufficient variation in the key regression variable. Counterintuitively, this variation arises in firms far different from the canonical capital-intensive firms for which the theory was initially developed. Our findings suggest that corporate learning may be an important feature to capture in investment models, and that Tobin's q may be a particularly effective proxy for investment opportunities in R&D industries. Most importantly, as research-intensive firms are a growing segment of the economy, the future of the investment-q relationship looks increasingly bright.

A Appendix

A.1 Aggregate data from NIPA tables and Flow of Funds

This section details how we construct the aggregate quarterly series of investment and Tobin's Q, following Hall (2001) and Philippon (2009). The series from those papers are publicly available, but they end in 1999 and 2007 respectively, so we must reconstruct them with more recent data in order to extend the time-series.

Tobin's q The numerator of Tobin's q is the aggregate market value of corporate equity and corporate debt, minus corporate inventories. Aggregate market equity is series FL103164103 from the Fed's Flow of Funds website (note that this series was previously labeled FL103164003 until mid-2010).[14] Aggregate corporate debt is measured as financial liabilities (series FL104190005Q), minus financial assets (series FL104090005Q), plus the market value of outstanding bonds, minus the book value of outstanding bonds. The book value of outstanding bonds is the sum of the outstanding amounts of taxable corporate bonds (series FL103163003Q) and tax-exempt corporate bonds (series FL103162000Q).

The market value of bonds is calculated according to an algorithm employed in Hall (2001). Corporate bonds are assumed to be issued with ten-year maturities at a yield taken from a broad index (for taxable bonds, the BAA yield reported by Moody's; for tax-exempt bonds, the muni bond yields reported by the Federal Reserve's Table H.15). Market values are then recalculated for each vintage of bonds in each year by discounting their remaining scheduled payments at the then-prevailing yield, so that the market and book values of any vintage of bonds diverge after the issuance date.

The denominator of Tobin's q is the replacement cost of the firm's capital stock. This is measured by capitalizing gross corporate fixed investment (series FU105013005) at an annual depreciation rate of 10%, and initializing the stock series at $569 billion (taken from Hall (2001)). The investment series are deflated, and then the capital stock reflated, using the NIPA implicit deflator for fixed non-residential investment (from NIPA Table 7.1, line 32).

[14] Flow of Funds data can be downloaded at http://www.federalreserve.gov/datadownload/Choose.aspx?rel=Z.1

Investment The numerator of corporate investment is the seasonally-adjusted series of private nonresidential fixed investment (PNFI), available from the FRED website of the St. Louis Fed. The denominator is the gross capital stock, measured as the net stock of fixed assets (series K1NTOTL1ES000 on FRED) plus their aggregate depreciation (series M1NTOTL1ES000). Both investment and capital stock are deflated using the same NIPA deflator as above. The capital stock series is recorded only in the last quarter of each year, so we interpolate these year-end values to the other quarters of the year.

A.2 Proof of Proposition 1

The observable variables are the cash-flow process (7) and the signal (9). The unobservable variable is μ_t. Write the dynamics of the observable variables θ_t and s_t:

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

and of the unobservable variable μ_t:

$$d\mu_t = (\underbrace{\eta\bar{\mu}}_{a_0} + \underbrace{(-\eta)}_{a_1}\mu_t)dt + \underbrace{\sigma_\mu}_{b_1} dW_t^{\mu} + \underbrace{\begin{bmatrix} 0 & 0 \end{bmatrix}}_{b_2} \begin{bmatrix} dW_t^{\theta} \\ dW_t^{s} \end{bmatrix}. \tag{A.2}$$

We will apply the following standard theorem.

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

$$du_t = [a_0(t, s_t) + a_1(t, s_t)u_t]\, dt + b_1(t, s_t)dZ_t^u + b_2(t, s_t)dZ_t^s \tag{A.3}$$

$$ds_t = [A_0(t, s_t) + A_1(t, s_t)u_t]\, dt + B_1(t, s_t)dZ_t^u + B_2(t, s_t)dZ_t^s. \tag{A.4}$$

All the parameters can be functions of time and of the observable process. Liptser and Shiryaev (1977) show that the filter evolves according to (we drop the dependence of

coefficients on t and s_t for notational convenience):

$$d\widehat{u}_t = (a_0 + a_1\widehat{u}_t)dt + [(b \circ B) + \zeta_t A_1^\top](B \circ B)^{-1}[ds_t - (A_0 + A_1\widehat{u}_t)dt] \tag{A.5}$$

$$\frac{d\zeta_t}{dt} = a_1\zeta_t + \zeta_t a_1^\top + (b \circ b) - [(b \circ B) + \zeta_t A_1^\top](B \circ B)^{-1}[(b \circ B) + \zeta_t A_1^\top]^\top, \tag{A.6}$$

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

$$b \circ b = b_1 b_1^\top + b_2 b_2^\top \tag{A.7}$$

$$B \circ B = B_1 B_1^\top + B_2 B_2^\top \tag{A.8}$$

$$b \circ B = b_1 B_1^\top + b_2 B_2^\top. \tag{A.9}$$

In our setup, we obtain

$$b \circ b = \sigma_\mu^2 \tag{A.10}$$

$$B \circ B = \begin{bmatrix} \sigma_\theta^2 & 0 \\ 0 & \frac{\Phi+1}{\Phi} \end{bmatrix} \tag{A.11}$$

$$b \circ B = \begin{bmatrix} 0 & \sigma_\mu \end{bmatrix}, \tag{A.12}$$

and

$$[(b \circ B) + \zeta_t A_1^\top](B \circ B)^{-1} = \begin{bmatrix} \frac{\lambda\zeta_t}{\sigma_\theta^2} & \frac{\sigma_\mu \Phi}{1+\Phi} \end{bmatrix}. \tag{A.13}$$

Furthermore, the Bayesian uncertainty ζ_t follows the deterministic process

$$\frac{d\zeta_t}{dt} = \frac{\sigma_\mu^2}{1+\Phi} - 2\eta\zeta_t - \frac{\lambda^2\zeta_t^2}{\sigma_\theta^2}, \tag{A.14}$$

which has the following steady-state solution

$$\bar{\zeta} = \frac{\sigma_\theta^2}{\lambda^2}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{A.15}$$

Using (A.5) and (A.13) and replacing $\zeta_t = \bar{\zeta}$, we can write

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \begin{bmatrix} \frac{\lambda}{\sigma_\theta^2}\bar{\zeta} & \sigma_\mu \frac{\Phi}{1+\Phi} \end{bmatrix} \begin{bmatrix} d\theta_t - \lambda(\widehat{\mu}_t - \theta_t)dt \\ ds_t \end{bmatrix}. \tag{A.16}$$

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

$$d\theta_t = \lambda(\mu_t - \theta_t)dt + \sigma_\theta dW_t^\theta \tag{A.17}$$

$$d\theta_t = \lambda(\widehat{\mu}_t - \theta_t)dt + \sigma_\theta d\widehat{W}_t^\theta. \tag{A.18}$$

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:

$$d\theta_t - \lambda(\widehat{\mu}_t - \theta_t)dt = \sigma_\theta d\widehat{W}_t^\theta. \tag{A.19}$$

Furthermore, we can write the signal as

$$ds_t = dW_t^\mu + \frac{1}{\sqrt{\Phi}} dW_t^s = \sqrt{\frac{\Phi+1}{\Phi}} d\widehat{W}_t^s, \tag{A.20}$$

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

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \begin{bmatrix} \frac{\lambda}{\sigma_\theta}\bar{\zeta} & \sigma_\mu \sqrt{\frac{\Phi}{1+\Phi}} \end{bmatrix} \begin{bmatrix} d\widehat{W}_t^\theta \\ d\widehat{W}_t^s \end{bmatrix}, \tag{A.21}$$

which, after replacement of (A.15), yields:

$$d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \frac{\sigma_\theta}{\lambda}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right) d\widehat{W}_t^\theta + \sigma_\mu \sqrt{\frac{\Phi}{1+\Phi}} d\widehat{W}_t^s. \tag{A.22}$$

Notice that from (A.17)-(A.18) we can write:

$$d\widehat{W}_t^\theta = dW_t^\theta + \frac{\lambda}{\sigma_\theta}(\mu_t - \widehat{\mu}_t)dt, \tag{A.23}$$

and we also have from (A.20)÷

$$d\widehat{W}_t^s = \sqrt{\frac{\Phi}{1+\Phi}} ds_t. \tag{A.24}$$

We can therefore write Proposition 1.

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

$$\mathrm{Var}_t[\widehat{\mu}_t] = \sigma_\mu^2 - \frac{2\eta\sigma_\theta^2}{\lambda^2}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{A.25}$$

We can then compute

$$\frac{\partial \mathrm{Var}_t[\widehat{\mu}_t]}{\partial \sigma_\mu} = 2\sigma_\mu\left(1 - \frac{\eta\sigma_\theta}{(1+\Phi)\sqrt{\eta^2\sigma_\theta^2 + \frac{\lambda^2\sigma_\mu^2}{1+\Phi}}}\right) > 0 \tag{A.26}$$

and

$$\frac{\partial \mathrm{Var}_t[\widehat{\mu}_t]}{\partial \Phi} = \frac{\eta\sigma_\theta\sigma_\mu^2}{(1+\Phi)^2\sqrt{\eta^2\sigma_\theta^2 + \frac{\lambda^2\sigma_\mu^2}{1+\Phi}}} > 0, \tag{A.27}$$

which leads to Corollary 1.1.

A.3 Discretization used for simulations

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

$$\text{cash flow:} \quad d\theta_t = \lambda(\widehat{\mu}_t - \theta_t)dt + \sigma_\theta d\widehat{W}_t^\theta \tag{A.28}$$

$$\text{(scaled) signal:} \quad d\widehat{W}_t^s \tag{A.29}$$

$$\text{filter:} \quad d\widehat{\mu}_t = \eta(\bar{\mu} - \widehat{\mu}_t)dt + \Omega d\widehat{W}_t^\theta + \sigma_\mu\sqrt{\frac{\Phi}{1+\Phi}} d\widehat{W}_t^s \tag{A.30}$$

$$\text{capital purchase price shocks:} \quad d\nu_t = -\kappa\nu_t dt + \sigma_\nu dW_t^\nu, \tag{A.31}$$

where we define

$$\Omega \equiv \frac{\sigma_\theta}{\lambda}\left(\sqrt{\eta^2 + \frac{1}{1+\Phi}\frac{\lambda^2\sigma_\mu^2}{\sigma_\theta^2}} - \eta\right). \tag{A.32}$$

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

$$\frac{I_t}{K_t} = -\frac{1}{a} + \frac{1}{a}q(\theta_t, \widehat{\mu}_t, \nu_t) - \frac{1}{a}\nu_t, \tag{A.33}$$

which provides all the data necessary for the regressions. We implement the following discretization of the continuous-time processes (A.28)-(A.31):

$$\nu_{t+\Delta} = \nu_t e^{-\kappa\Delta} + \sigma_\nu\sqrt{\frac{1-e^{-2\kappa\Delta}}{2\kappa}}dW_t^\nu, \tag{A.34}$$

$$\widehat{\mu}_{t+\Delta} = \widehat{\mu}_t e^{-\eta\Delta} + \bar{\mu}\left(1-e^{-\eta\Delta}\right) + \sqrt{\frac{1-e^{-2\eta\Delta}}{2\eta}}\left(\Omega d\widehat{W}_t^\theta + \sigma_\mu\sqrt{\frac{\Phi}{1+\Phi}}d\widehat{W}_t^s\right), \tag{A.35}$$

$$\theta_{t+\Delta} = \theta_t e^{-\lambda\Delta} + \widehat{\mu}_t\left(1-e^{-\lambda\Delta}\right) + \sigma_\theta\sqrt{\frac{1-e^{-2\lambda\Delta}}{2\lambda}}d\widehat{W}_t^\theta. \tag{A.36}$$

A.4 Combining learning with decreasing returns to scale

This appendix extends the model from Section 3 to allow for market power, which we represent through decreasing returns to scale in the profit function (Cooper and Ejarque, 2003). Suppose the profit function is as follows:

$$\Pi(K_t, \theta_t) = \theta_t K_t^\alpha. \tag{A.37}$$

If $\alpha < 1$, the Hayashi (1982) conditions are violated and marginal q no longer equals average q, so that the use of average q in the investment-q regression induces measurement error.

As in Section 3, equation (14), the firm's objective function leads to a linear

relationship between investment and marginal q,

$$\frac{I}{K} = -\frac{1}{a} + \frac{1}{a} V_K(K, \theta, \widehat{\mu}, \nu) + \frac{1}{a}\nu, \tag{A.38}$$

where $V_K(K, \theta, \widehat{\mu}, \nu)$ denotes marginal q, i.e., the shadow cost of capital. Replacing this relationship in the HJB equation yields a partial differential equation for V. We solve this equation numerically by approximating $V(K, \theta, \widehat{\mu}, \nu)$ with Chebyshev polynomials. We compare three model specifications:

(i) A model without learning and with constant returns to scale ($\alpha = 1$). This is a special case of the model analyzed in Section 3, in which marginal q equals average q. Consequently, the investment-q regression exhibits no measurement error.

(ii) A model with learning and with decreasing returns to scale ($\alpha = 0.95$). In this specification, $\sigma_\mu = 0.1$ and the firm learns from cash-flow realizations and from the additional signal s_t, with $\Phi = 5$. In this model, the investment-q regression exhibits measurement error.

(iii) A model with learning and with decreasing returns to scale ($\alpha = 0.9$). In this specification, $\sigma_\mu = 0.15$ and the firm learns from cash-flow realizations and from the additional signal s_t, which is more informative with $\Phi = 20$. In this model, the investment-q regression exhibits measurement error.

For all the above specifications, the calibration is the same as in our baseline model of Section 3. The values chosen for the parameter α are in line with calibrations used in the literature (e.g., Gomes, 2001).

Table 5 presents simulation results. Each of the three models above is simulated 5,000 times at yearly frequency for 100 years. For each simulation, we run the investment-q regression using average q (which equals V/K) as a proxy for marginal q. Each row of the table shows the R^2 coefficient, the slope of the regression, and the volatility of average q (where all reported statistics are averaged over the 5,000 simulations).

[Table 5 here]

Column (1) shows that the R^2 coefficient increases with learning, which is our main result. Column (2) shows that the slope of the investment-q regression decreases with

market power. It might seem that this decreasing slope should also lead to a decrease in R^2, but the offsetting impact of learning leads to a net increase in R^2. The reason for the increase in R^2 is a substantial increase in the volatility of average q, as shown in column (3). As elaborated above, this results from learning.

Figure 20 depicts the relationship between investment and average q for two simulated samples. The blue triangles are generated from a simulation of Model (i) without learning and without decreasing returns to scale. The red crosses are generated from a simulation of Model (iii) with learning and decreasing returns to scale, $\alpha = 0.9$. This figure reproduces qualitatively the pattern of Figure 5, where the R^2 coefficients increase while the slope coefficients decrease. Indeed, while decreasing returns to scale dampen the slope coefficient, learning induces a higher volatility of average q, which ultimately leads to a higher R^2. We also notice that Tobin's q is on average higher with decreasing returns to scale, in line with the intuition from Lindenberg and Ross (1981) that q should persist above one for firms with monopoly rents.

[Figure 20 here]

References

Abel, Andrew, 1983, Optimal Investment under Uncertainty, *American Economic Review* 73, 228–233.

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?, *Journal of Finance* 58, 707–722.

Bai, Jennie, Thomas Philippon, and Alexi Savov, 2016, Have Financial Markets Become More Informative?, *Journal of Financial Economics* 122, 625–654.

Bakke, Tor-Erik, and Toni M Whited, 2010, Which Firms Follow the Market? An Analysis of Corporate Investment Decisions, *Review of Financial Studies* 23, 1941–1980.

Begenau, Juliane, Maryam Farboodi, and Laura Veldkamp, 2017, Big Data in Finance and the Growth of Large Firms *Working Paper*.

Brainard, William C., and James Tobin, 1968, Pitfalls in Financial Model Building, *American Economic Review Papers and Proceedings* 58, 99–122.

Brennan, Michael J, 1998, The Role of Learning in Dynamic Portfolio Decisions, *European Finance Review* 1, 295–306.

Bresnahan, Timothy, Paul Milgrom, and Jonathan Paul, 1992, The Real Output of the Stock Exchange, in *Output Measurement in the Service Sectors*, 195–216 (National Bureau of Economic Research, Inc).

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.

Chen, Qi, Itay Goldstein, and Wei Jiang, 2007, Price Informativeness and Investment Sensitivity to Stock Price, *Review of Financial Studies* 20, 619–650.

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.

Dow, James, Itay Goldstein, and Alexander Guembel, 2017, Incentives for Information Production in Markets where Prices Affect Real Investment, *Journal of the European Economic Association* 15, 877–909.

Dow, James, and Gary Gorton, 1997, Stock Market Efficiency and Economic Efficiency: Is There a Connection?, *Journal of Finance* 52, 1087–1129.

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.

Eberly, Andrew B. Abel Janice C., 2011, How Q and Cash Flow Affect Investment without Frictions: An Analytic Explanation, *Review of Economic Studies* 78, 1179–1200.

Edmans, Alex, Itay Goldstein, and Wei Jiang, 2015, Feedback Effects, Asymmetric Trading, and the Limits to Arbitrage, *American Economic Review* 105, 3766–3797.

Erickson, Timothy, Colin Jiang, and Toni Whited, 2014, Minimum Distance Estimation of the Errors-in-Variables Model Using Linear Cumulant Equations, *Journal of Econometrics* 183, 211–221.

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

Erickson, Timothy, and Toni Whited, 2012, Treating Measurement Error in Tobin's q, *Review of Financial Studies* 1286–1329.

Farboodi, Maryam, Adrien Matray, and Laura Veldkamp, 2017, Where Has All the Big Data Gone? *Princeton and NYU Working Paper*.

Fazzari, Steven, R. Glenn Hubbard, and Bruce Petersen, 1988, Financing Constraints and Corporate Investment, *Brookings Papers on Economic Activity* 1, 141–206.

Gilchrist, Simon, and Charles P. Himmelberg, 1995, Evidence on the Role of Cash Flow for Investment, *Journal of Monetary Economics* 36, 541–572.

Goldstein, Itay, and Alexander Guembel, 2008, Manipulation and the Allocational Role of Prices, *Review of Economic Studies* 75, 133–164.

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.

Hennessy, Christopher A, and Boris Radnaev, 2018, Learning and Leverage Cycles in General Equilibrium: Theory and Evidence, *Review of Finance* 22, 311–335.

Kaplan, Steven N, and Luigi Zingales, 1997, Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?, *Quarterly Journal of Economics* 112, 169–215.

Keynes, John Maynard, 1936, *The General Theory of Employment, Interest and Money* (London: Macmillan).

Lindenberg, Eric B., and Stephen Ross, 1981, Tobin's q Ratio and Industrial Organization, *Journal of Business* 54, 1–32.

Liptser, R, and A Shiryaev, 1977, Statistics of Random Processes, (Berlin, New York: Springer-Verlag).

Moyen, Nathalie, 2004, Investment-Cash Flow Sensitivities: Constrained versus Unconstrained Firms, *Journal of Finance* 59, 2061–2092.

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

Mussa, Michael, 1977, External and Internal Adjustment Costs and the Theory of Aggregate and Firm Investment, *Economica* 44, 163–178.

Myers, Stewart C., and Nicholas S. Majluf, 1984, Corporate Financing and Investment Decisions When Firms Have Information That Investors Do Not Have, *Journal of Financial Economics* 13, 187–221.

Peters, Ryan H, and Lucian A Taylor, 2017, Intangible capital and the investment-q relation, *Journal of Financial Economics* 123, 251–272.

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

Salinger, Michael A., 1984, Tobin's q, Unionization, and the Concentration-Profits Relationship, *RAND Journal of Economics* 15, 159–170.

Subrahmanyam, Avanidhar, and Sheridan Titman, 1999, The Going-Public Decision and the Development of Financial Markets, *Journal of Finance* 54, 1045–1082.

Tobin, James, 1969, A General Equilibrium Approach to Monetary Theory, *Journal of Money, Credit and Banking* 1, 15–29.

B Tables and figures

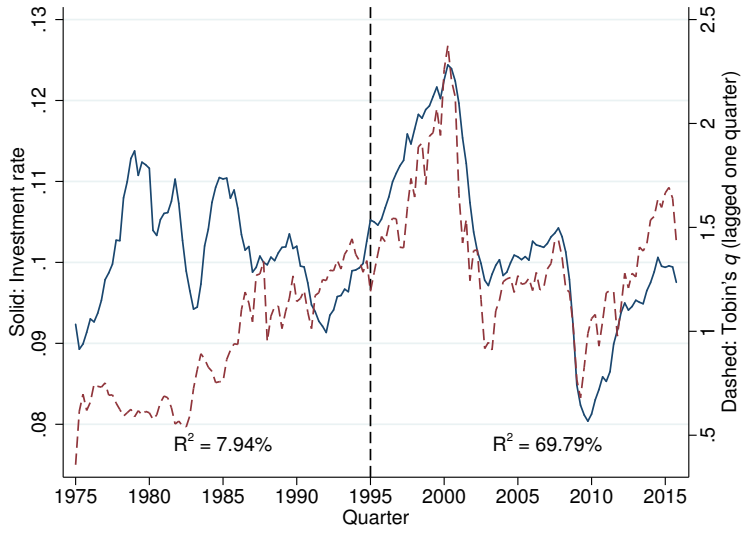

Figure 1: Aggregate quarterly investment rate and lagged Tobin's q.

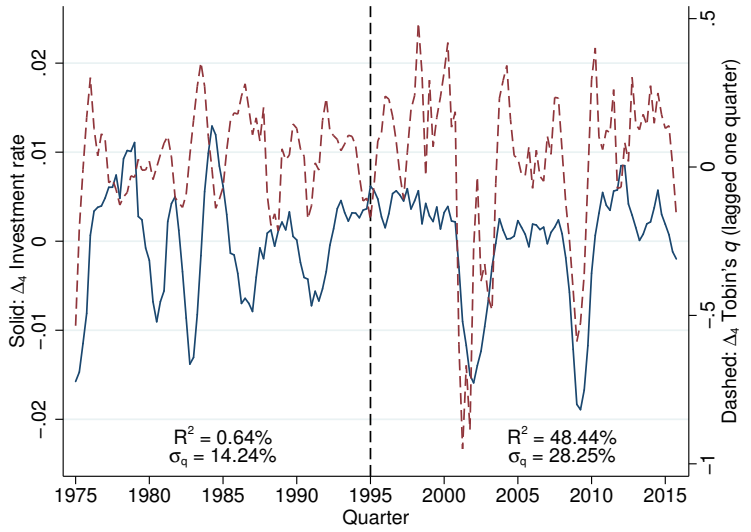

Figure 2: Year-over-year differences of aggregate quarterly investment rate and lagged Tobin's q.

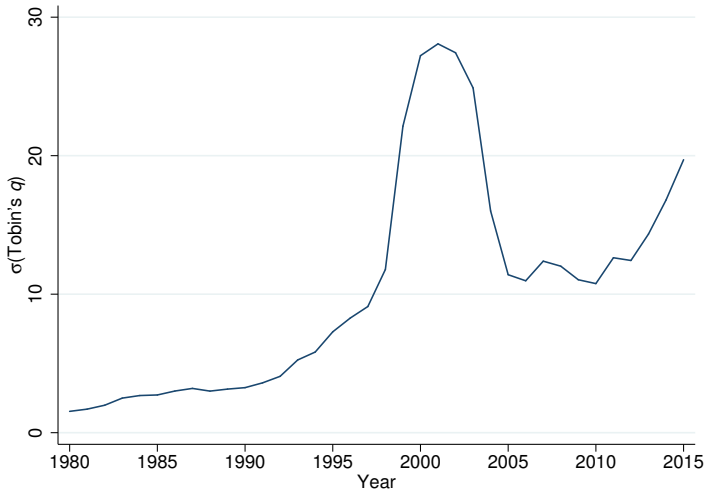

Figure 3: Between-firm dispersion in Tobin's q, 1980-2015. For each year, the figure plots the cross-sectional standard deviation of that year's Tobin's q across the firms in Compustat during that year. The series is smoothed over a five-year lag.

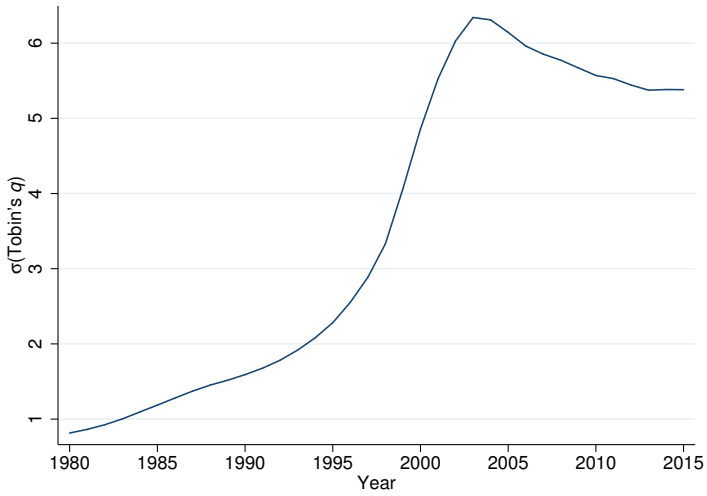

Figure 4: Within-firm dispersion in Tobin's q, 1980-2015. For each firm in Compustat, we calculate the within-firm volatility of Tobin's q during that firm's entire lifetime in Compustat. We then average that firm-level measure across all firms in Compustat for each year. The series is thus driven by changes in the composition of Compustat firms. Finally, the series is smoothed over a five-year lag.

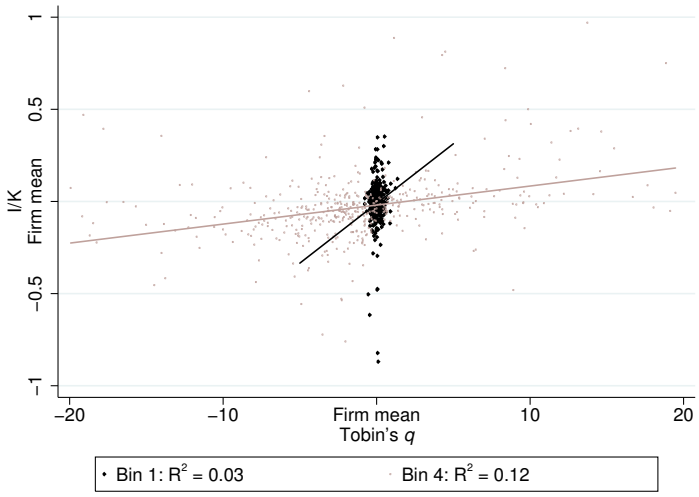

Figure 5: The scatter plot corresponds to Table 1. Each dot represents a firm-year observation, of which 500 are sampled from the lowest and highest bins of within-firm volatility in Tobin's q. The x-axis measures Tobin's q, and the y-axis measures the investment rate. Both numbers are demeaned within-firm to remove the firm fixed effect, so zeros on the axes correspond to the firm mean values. For each bin, the best-fit line of the same color reflects the regression in Table 1.

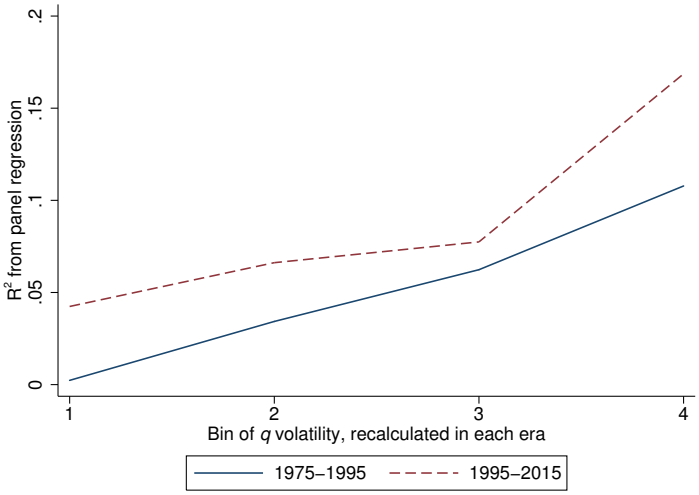

Figure 6: R^2 of the panel regression across four bins of volatility, where these bins are recalculated separately for the two subperiods 1975–1995 and 1995–2015.

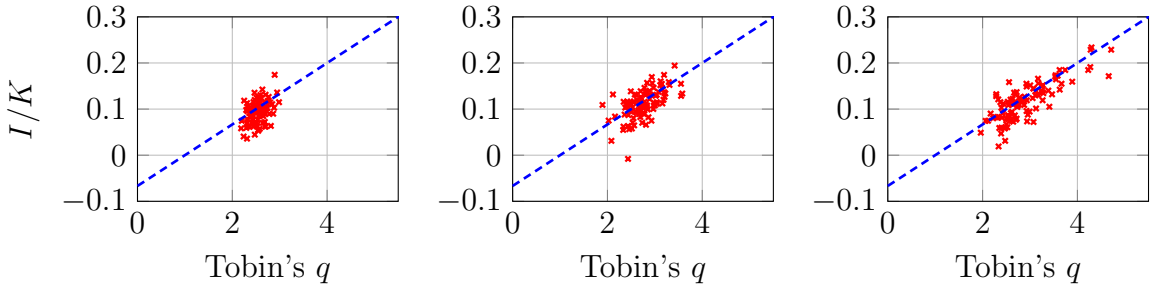

Figure 7: Relationship between investment and q for three different firms, for simulated samples of 100 yearly data points. In the left panel, the firm does not learn about μ_t, which is held constant at $\bar{\mu}$. In the middle panel, the firm learns about μ_t exclusively from the cash-flow process (7), i.e., $\Phi = 0$. In the right panel, the firm learns about μ_t from the cash-flow process (7) and from the signal (9) with $\Phi = 20$. The rest of the calibration used for these simulations is: $a = 16$, $r = 3\%$, $\delta = 10\%$, $\lambda = 0.5$, $\sigma_\theta = 0.1$, $\kappa = 5000$, $\sigma_\nu = 40$, $\bar{\mu} = 0.25$, and $\eta = 0.5$.

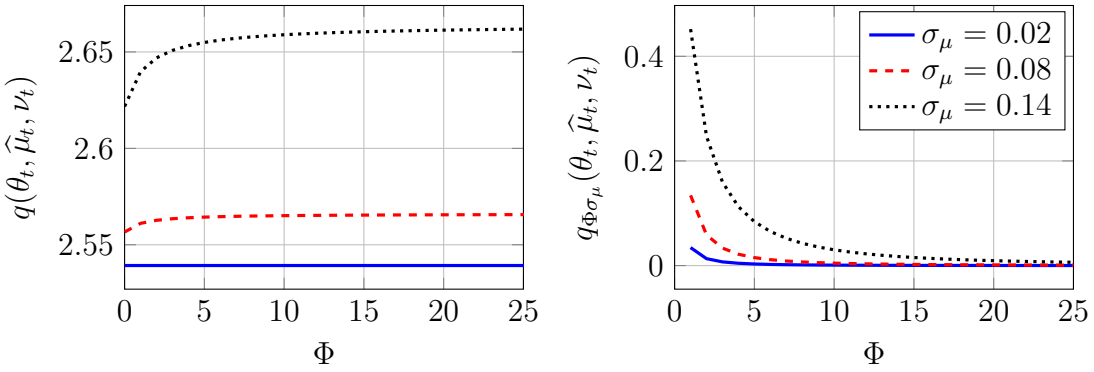

Figure 8: The left panel plots $q(\theta_t, \widehat{\mu}_t, \nu_t)$ when $\theta_t = \widehat{\mu}_t = \bar{\mu}$ and $\nu_t = 0$. Each line corresponds to a different value of $\sigma_\mu \in \{0.02, 0.08, 0.14\}$. The lines are plotted as functions of Φ, which goes from 0 to 25. The right panel uses the finite difference method to compute $q_{\Phi\sigma_\mu}$, which is positive in all cases.

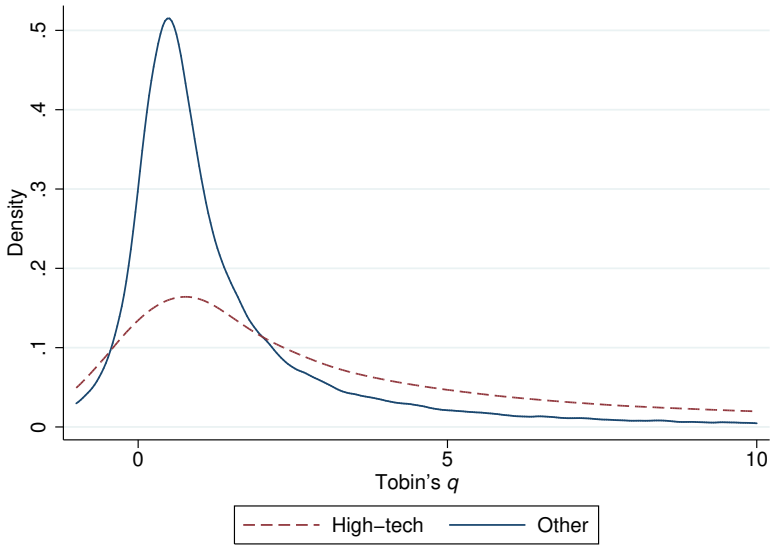

Figure 9: Empirical distribution of Tobin's q for firm-years in annual Compustat from 1975 to 2015, separating out high-tech industries from other industries. High-tech industries are defined as SIC codes 283, 357, 366, 367, 382, 384, and 737, following Brown et al. (2009).

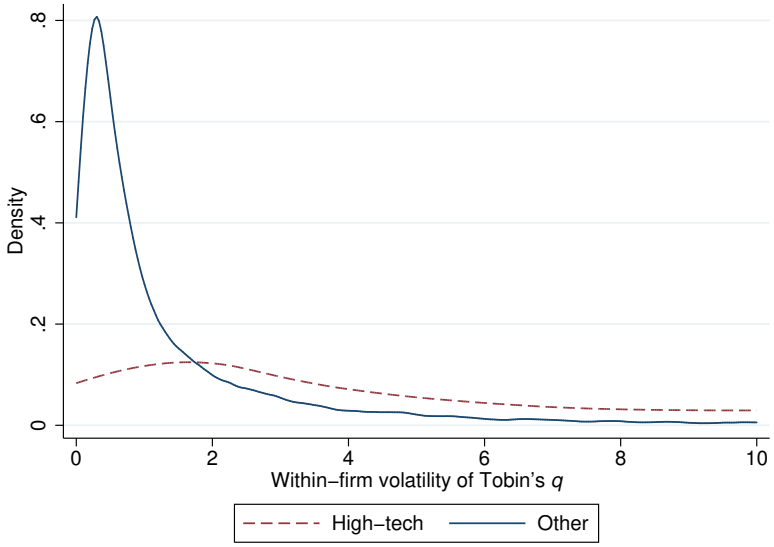

Figure 10: Empirical distribution of the within-firm volatility of Tobin's q for firm-years in annual Compustat from 1975 to 2015. The figure calculates the volatility for each firm, then averages high-tech industries separately from other industries. High-tech industries are defined as SIC codes 283, 357, 366, 367, 382, 384, and 737, following Brown et al. (2009).

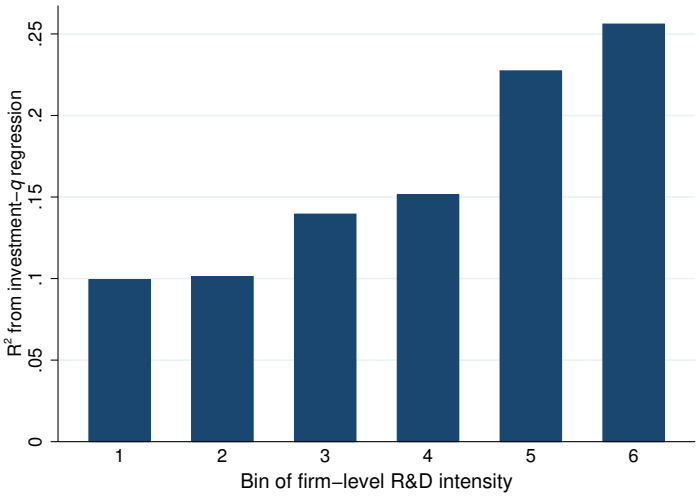

Figure 11: R^2 values from the panel investment-q regressions, performed separately for each bin of firm-level R&D intensity. R&D intensity is annual R&D expense divided by book assets, assigning zero for missing R&D data. It is calculated separately for each firm-year in Compustat from 1975-2015, then averaged within-firm.

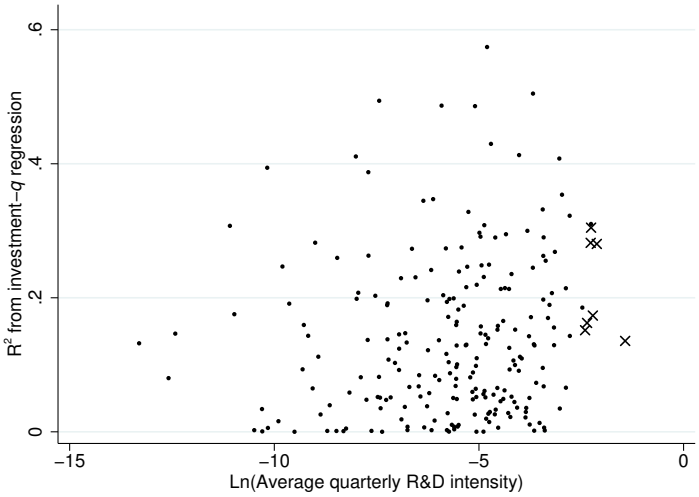

Figure 12: Each dot corresponds to a 3-digit SIC industry. The y-axis plots R^2 values from the panel investment-q regression performed separately in each industry. The x-axis plots the log of industry-average R&D intensity. R&D intensity is defined as annual R&D expense divided by book assets, assigning zero for missing R&D data. It is calculated separately for each firm-year in Compustat from 1975-2015, then averaged across firm-years in each 3-digit SIC industry in Compustat. Only industries with at least ten observations in Compustat are retained. The "×" markers denote the R&D-intensive industries identified in Brown et al. (2009).

Figure 13: 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).

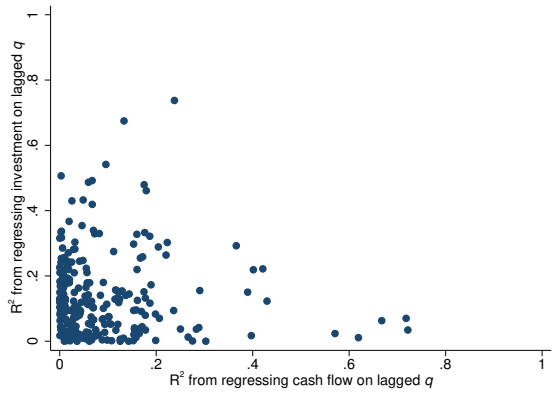

Figure 14: Each dot corresponds to a 3-digit SIC industry classification. The x-axis plots, for each industry, the R^2 value from a fixed-effects regression of cash flow on lagged Tobin's q. The y-axis plots, for the same industry, the R^2 from a fixed-effects regression of investment on lagged Tobin's q. The data are annual Compustat from 1975-2015. Only industries with at least ten observations in Compustat are retained. Cash flow is defined as income before extraordinary items plus depreciation expense. Cash flow, investment, and q are all winsorized at the 1st and 99th percentiles.

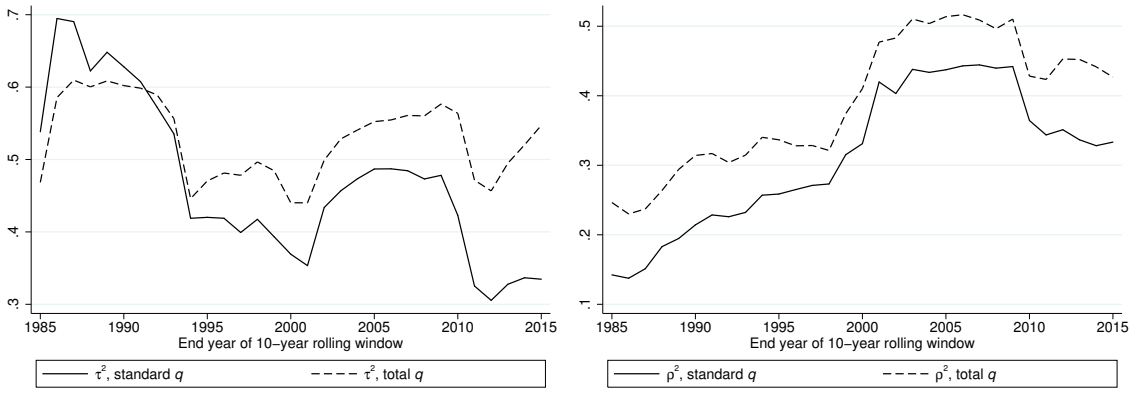

Figure 15: This figure plots the time-series of ρ^2 and τ^2 estimates recovered from the cumulant-estimator approach of Erickson, Jiang, and Whited (2014). Each estimate is calculated from a ten-year window ending in the year labeled on the axis. Solid lines use the standard definitions of investment and Tobin's q, while dashed lines use the total-q and total-investment measures defined in Peters and Taylor (2017).

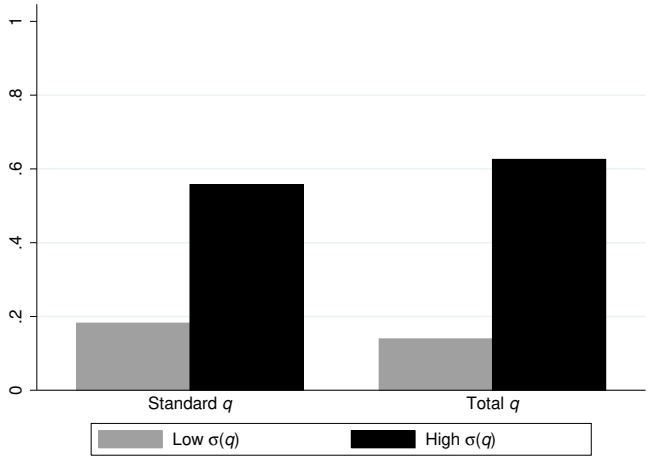

Figure 16: Estimates of ρ^2, the population R^2 of the investment equation on marginal q, following the cumulant-estimator approach of Erickson et al. (2014). The left pair of bars uses the standard measure of Tobin's q, and the right pair uses the intangible total-q adjustment of Peters and Taylor (2017). Within each pair, the left (gray) bar plots the estimate of ρ^2 for firms with below-median volatility of Tobin's q, and the right (black) bar plots the estimate for firms with above-median volatility.

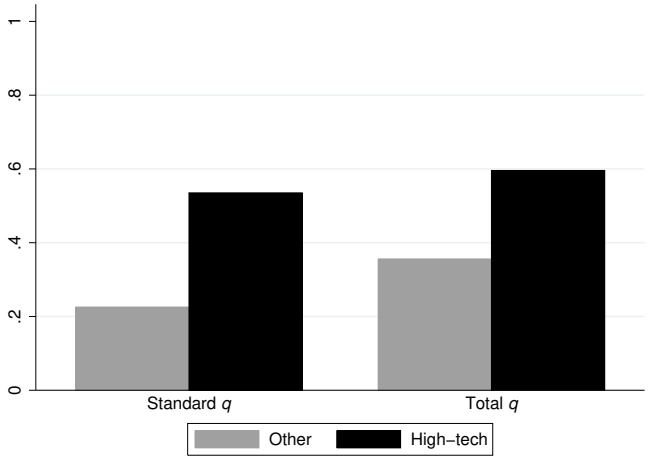

Figure 17: Estimates of ρ^2, the population R^2 of the investment equation on marginal q, following the cumulant-estimator approach of Erickson et al. (2014). The left pair of bars uses the standard measure of Tobin's q, and the right pair uses the intangible total-q adjustment of Peters and Taylor (2017). Within each pair, the right (black) bar plots the estimate for high-tech firms, and the left (gray) bar plots the estimate of ρ^2 for other firms.

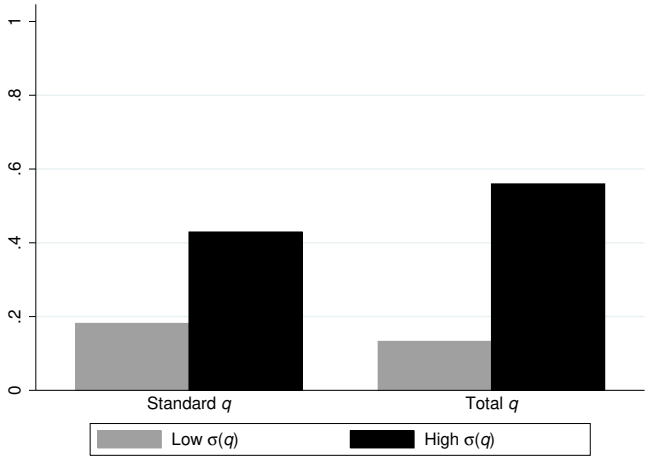

Figure 18: Estimates of τ^2, the population R^2 of the measurement equation of Tobin's q on marginal q, following the cumulant-estimator approach of Erickson et al. (2014). Within each pair, the left (gray) bar plots the estimate of ρ^2 for firms with below-median volatility of Tobin's q, and the right (black) bar plots the estimate for firms with above-median volatility.

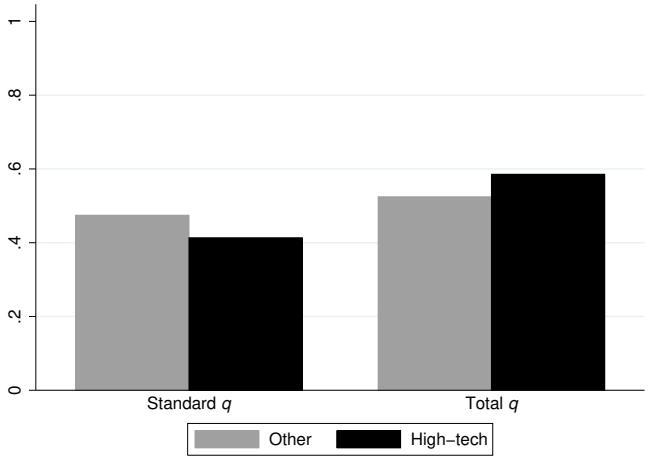

Figure 19: Estimates of τ^2, the population R^2 of the measurement equation of Tobin's q on marginal q, following the cumulant-estimator approach of Erickson et al. (2014). Within each pair, the right (black) bar plots the estimate for high-tech firms, and the left (gray) bar plots the estimate of ρ^2 for other firms.

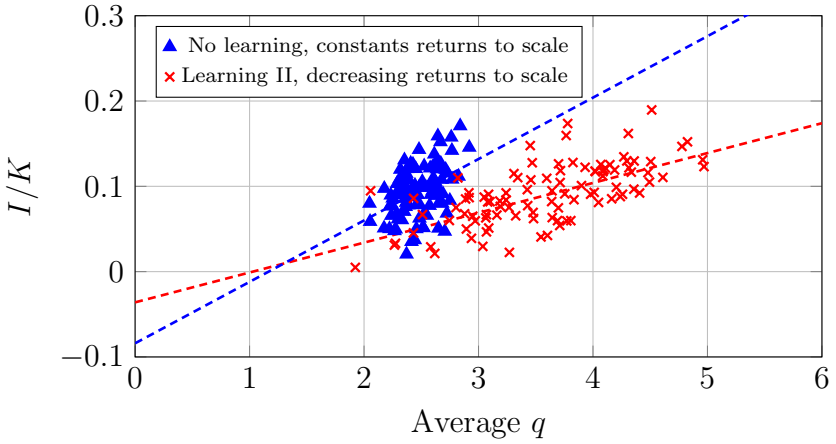

Figure 20: Relationship between investment and q without learning (Model (i) in Table 5, blue triangles) or with learning (Model (iii) in Table 5, red crosses). In Model (i), the profit function has constant returns to scale. In Model (iii), the profit function has decreasing returns to scale ($\alpha = 0.9$), which leads to measurement error in the regression and thus a lower slope. The simulations are performed over 100 years at an annual frequency.

	(1) I_t/K_{t-1}	(2) I_t/K_{t-1}	(3) I_t/K_{t-1}	(4) I_t/K_{t-1}
q_{t-1}	0.0777***	0.0658***	0.0377***	0.00470***
	(0.00498)	(0.00272)	(0.00163)	(0.000531)
Sample	Bin 1	Bin 2	Bin 3	Bin 4
Firm FE?	Yes	Yes	Yes	Yes
Obs.	34996	37752	38059	37332
R^2	0.0258	0.0603	0.0718	0.125

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. Firms are sorted into bins based on the within-firm volatility of Tobin's q, with bin 4 as the highest volatility. Standard errors are clustered by firm, and the table reports the within-firm R^2 of the regression.

	(1) I_t/K_{t-1}	(2) I_t/K_{t-1}	(3) I_t/K_{t-1}	(4) I_t/K_{t-1}
q_{t-1}	0.0765***	0.0653***	0.0397***	0.0166***
	(0.00499)	(0.00282)	(0.00166)	(0.000568)
Sample	Bin 1	Bin 2	Bin 3	Bin 4
Firm FE?	Yes	Yes	Yes	Yes
Obs.	34912	37704	38114	37409
R^2	0.0237	0.0559	0.0768	0.152

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_t/K_{t-1}	(2) I_t/K_{t-1}	(3) I_t/K_{t-1}	(4) I_t/K_{t-1}
q_{t-1}	0.0174***	0.0118***	0.0245***	0.0239***
	(0.000565)	(0.000303)	(0.00144)	(0.00136)
Sample	Non-high-tech	High-tech	Non-high-tech, pre-1995	High-tech, pre-1995
Firm FE?	Yes	Yes	Yes	Yes
Obs.	113297	34900	46464	9223
R^2	0.107	0.221	0.0678	0.211

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. "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 1975-2015. Columns 3 and 4 restrict to pre-1995 firm-years. Investment and q are winsorized at the 1st and 99th percentiles. Standard errors are clustered by firm, and the table reports the within-firm R^2 of the regression.

	(1) I_t/K_{t-1}	(2) I_t/K_{t-1}	(3) I_t^{tot}/K_{t-1}	(4) I_t^{tot}/K_{t-1}
q_{t-1}	0.00947***	0.0112***	0.0124***	0.00866***
	(0.000239)	(0.000281)	(0.000448)	(0.000259)
Sample	High-tech	High-tech	Non-high-tech	High-tech
Firm FE?	No	Yes	Yes	Yes
Year FE?	No	Yes	No	No
Obs.	34900	34900	113230	34899
R^2	0.218	0.305	0.0802	0.201

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 1 drops both firm and year fixed effects, and Column 2 includes both firm and year fixed effects as is done in Peters and Taylor (2017). 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.

Model	Market power	Calibration	(1) Avg. R^2	(2) Avg. slope	(3) Avg. $\sigma(q)$
(i) No learning	No ($\alpha = 1$)	$\sigma_\mu = 0$, $\Phi = 0$	0.179	0.063	0.184
(ii) Learning I	Yes ($\alpha = 0.95$)	$\sigma_\mu = 0.10$, $\Phi = 5$	0.349	0.053	0.345
(iii) Learning II	Yes ($\alpha = 0.90$)	$\sigma_\mu = 0.15$, $\Phi = 20$	0.402	0.038	0.614

Table 5: Simulations of three different models with varying degrees of market power. Row (i) considers a model without learning and with constant returns to scale, $\Pi(K_t, \theta_t) = \theta_t K_t$. Rows (ii) and (iii) consider models with learning and decreasing returns to scale, $\Pi(K_t, \theta_t) = \theta_t K_t^\alpha$, $\alpha < 1$. All other parameters are as in the baseline calibration (see Section 3). Each simulation contains 100 yearly data points. The average R^2 coefficients and the average slope coefficients from 5,000 regressions of I/K on average q (V/K) are reported in columns (a) and (b). Column (c) reports the mean volatility of average q over the 5,000 simulations.