# The Effect of Tobin's $q$ Ratio on Corporate Investment: Evidence from the U.S. Firm Panel Data Set

Duc Thanh (Timothy) Nguyen
Department of Economics, Boston University

November 10, 2025

## Abstract

This paper examines the effect of Tobin's $q$ on corporate investment using a panel of 1,962 U.S. firms. Pooled OLS models show an investment elasticity of 0.24, which is robust to financial controls. Notably, non-NYSE/AMEX firms are more sensitive to $q$ than listed firms. Fixed-effects models yield a higher elasticity of 0.28, with Hausman tests favoring fixed over random effects. A fixed-effects IV model produces the largest elasticity (0.34), though tests fail to reject the exogeneity of $q$. Overall, Tobin's $q$ has a positive, meaningful causal effect on investment, though firm heterogeneity and financial variables remain significant factors.

*Keywords [proposed]:* Tobin's $q$; corporate investment; panel data; fixed effects; instrumental variables; financing constraints; firm heterogeneity.
*JEL classification [proposed]:* G31; E22; G32; C23; C26.

## 1 Introduction

Corporate investment plays a central role in shaping productivity, long-run growth, and firm-level competitiveness. When firms decide how much to invest, they compare the cost of installing new capital with the value that financial markets place on their existing assets. Tobin's $q$ ratio—the market value of the firm relative to the replacement cost of its assets—provides a natural summary of these expectations. If financial markets correctly anticipate future profitability, then a higher Tobin's $q$ should be associated with higher investment, while firms with low $q$ should cut back.

A large empirical literature has tested this prediction by estimating "$q$-models" of investment, but the evidence is mixed. Many studies find a positive and statistically significant relationship between investment and Tobin's $q$, yet the explanatory power of $q$ alone is often modest, and investment also appears to depend on variables such as cash flow, leverage, and R&D. At the same time, concerns have been raised that average Tobin's $q$ may be endogenous: it may reflect not only fundamentals, but also time-varying risk, mispricing, or feedback from investment itself. These issues make it difficult to interpret simple correlations between $q$ and investment as evidence of a strong causal effect.

This paper revisits the effect of Tobin's $q$ on corporate investment using a detailed panel dataset of U.S. corporations that includes information on investment, market value, assets, cash flow, debt, R&D, and advertising. Rather than modeling stock-market valuation as the outcome, I treat Tobin's $q$, constructed as a value-to-assets ratio, as the key regressor and study how firm-level investment responds to changes in $q$ once I control for financial variables and unobserved firm characteristics. The central research question is: to what extent does Tobin's $q$ have a causal effect on firms' investment-to-assets ratios, once we control for firm characteristics, financial variables, and time-invariant unobserved heterogeneity?

To answer this question, I estimate a sequence of models that progressively impose more structure and address more sources of bias. I begin with simple pooled regressions of investment on $q$, then add additional firm-level controls and interactions. I then exploit the panel nature of the data by estimating fixed-effects models that control for time-invariant firm heterogeneity and common year shocks. Finally, I use instrumental variables derived from present-value-of-dividends measures to address the potential endogeneity of Tobin's $q$. This approach allows for a systematic comparison between ordinary least squares, fixed-effects, and fixed-effects IV estimates of the $q$–investment relationship.

The rest of the paper is organized as follows. Section 2 reviews the theoretical and empirical literature on Tobin's $q$, investment, and financing constraints. Section 3 describes the U.S. firm panel dataset and presents summary statistics. Section 4 sets out the econometric framework and modelling strategy. Section 5 reports the empirical results, and Section 6 discusses their economic interpretation, limitations, and implications. Section 7 concludes.

## 2 Literature Review

Tobin's $q$-theory of investment predicts that firms invest until the market value of an extra unit of capital equals its replacement cost, so investment should rise with the ratio of market value to replacement cost. Hayashi (1982) shows that under restrictive conditions—constant returns to scale and a particular form of adjustment costs—marginal $q$ equals average $q$, which can be measured as a market-value-to-assets ratio. This result justifies using average Tobin's $q$ in empirical investment equations, but it also highlights that violations of these assumptions, as well as measurement error in capital and market value, can weaken the link between average $q$ and true investment opportunities.

A common empirical strategy in this literature is to regress an investment ratio on Tobin's $q$ and additional controls such as cash flow. For example, Blundell et al. (1992) estimate panel regressions of the form

$$\frac{I_{it}}{K_{i,t-1}} = \alpha_i + \lambda_t + \beta\, q_{it} + \gamma \frac{CF_{it}}{K_{i,t-1}} + u_{it},$$

using a large panel of manufacturing firms and finding the coefficient on $q_{it}$ is positive and statistically significant, but economically modest, and that including cash flow improves the fit of the model. Bond et al. (2003) use similar investment equations for firm-level panels from several European countries and also conclude that Tobin's $q$ has a limited ability to explain investment dynamics once other financial variables are included. Relative to these studies, my preferred fixed-effects estimates in Section 5 suggest a somewhat larger elasticity of investment with respect to Tobin's $q$—about 0.28 in the within-firm OLS model and about 0.34 when Tobin's $q$ is instrumented—but like the previous literature, my models also have a fairly low $R^2$ and indicate that $q$ is only one among several determinants of investment.

The dataset used in this paper comes from Hall and Hall (1993),[1] who study the "value and performance" of U.S. corporations by comparing stock-market valuations with the present discounted value of dividends. Their focus is on explaining valuation gaps using variables such as investment, R&D, and leverage, rather than on estimating an explicit investment equation. In contrast, this paper takes Tobin's $q$ (constructed as a value-to-assets ratio) as the key regressor and examines how investment responds to changes in $q$ in a firm-level panel. Following the financing-constraints literature, I include cash flow, debt, and R&D and advertising intensities as controls, and I address concerns about the endogeneity of $q$ using fixed effects and instruments derived from Hall and Hall's present-value-of-dividends measures. This design allows a reassessment of how large and robust the effect of Tobin's $q$ on corporate investment is in the U.S. firm panel dataset.

---

[1] Data and variable definitions are documented at `https://users.ssc.wisc.edu/~behansen/econometrics/invest_description.pdf`.

## 2.1 Theoretical Foundations

The theoretical basis for using Tobin's $q$ as a determinant of investment originates with Tobin (1969) and Brainard and Tobin (1968), who argued that firms should invest when the market valuation of capital exceeds its replacement cost. The formal link between marginal and average $q$ was established by Hayashi (1982), who demonstrated that under constant returns to scale and convex adjustment costs, the observable average $q$ is a sufficient statistic for investment. However, Abel and Blanchard (1986) showed that when adjustment costs are more general or when firms face irreversibility constraints, the gap between marginal and average $q$ can be substantial, introducing measurement error into empirical investment equations.

Subsequent theoretical work has refined these insights. Abel and Eberly (1994) develop a model with irreversible investment and show that the option value of waiting generates a wedge between $q$ and the investment threshold, implying that average $q$ may systematically overstate investment incentives when uncertainty is high. Caballero and Engel (1999) introduce heterogeneous firms with lumpy adjustment and demonstrate that aggregate investment dynamics can differ markedly from the smooth adjustment implied by standard $q$ models. These theoretical developments underscore the importance of controlling for firm heterogeneity and using panel methods, as I do in this paper.

## 2.2 Empirical Evidence on Tobin's $q$ and Investment

The empirical literature on $q$ and investment is extensive but yields mixed conclusions regarding the strength of the relationship. Early work by Summers (1981) found that $q$ models explain relatively little of the variation in aggregate investment, a finding that has persisted in many subsequent studies. Blundell et al. (1992) estimate panel regressions using a large sample of UK manufacturing firms and report a positive but economically modest coefficient on $q$, with cash flow providing additional explanatory power. Bond et al. (2003) extend this analysis to firm-level panels from Belgium, France, Germany, and the United Kingdom and also find that $q$ has limited ability to explain investment dynamics once financial variables are included.

A key methodological concern in this literature is measurement error in Tobin's $q$. Erickson and Whited (2000) demonstrate that standard proxies for $q$ are noisy measures of true marginal $q$ and that this measurement error can severely attenuate the estimated coefficient in investment regressions. They propose higher-order moment estimators to address this problem and find substantially larger $q$ coefficients after correction. Cummins et al. (2006) use analysts' earnings forecasts as an alternative measure of fundamentals and report a much stronger relationship between investment and $q$ when measurement error is reduced. These findings motivate my use of instrumental variables in Section 5, where I employ the present discounted value of dividends and asset-pricing terms from Hall and Hall (1993) to address potential attenuation bias.

## 2.3 Financial Constraints and Investment

A central debate in the investment literature concerns the role of financial constraints. Fazzari et al. (1988) argue that the sensitivity of investment to cash flow, after controlling for $q$, reflects financing frictions: firms that face higher costs of external finance rely more heavily on internal funds. This interpretation was challenged by Kaplan and Zingales (1997), who show that the relationship between financial constraints and investment–cash flow sensitivity is not monotonic, casting doubt on the use of cash-flow coefficients as indicators of financing constraints.

The financial-constraints literature has since developed in several directions. Whited (1992) demonstrates that debt capacity and credit-market imperfections can generate investment–cash flow sensitivity even in well-specified $q$ models. Gilchrist and Himmelberg (1995) use the fundamental component of $q$ estimated from a VAR framework and find that cash flow retains explanatory power after removing measurement error in $q$, supporting the financing-constraints interpretation. More recently, Hadlock and Pierce (2010) propose the SA index based on firm size and age as a more reliable measure of financial constraints than earlier classification schemes. In my analysis, I include cash flow, leverage, and an NYSE/AMEX listing dummy as controls to capture different dimensions of financial frictions and test whether the $q$–investment relationship varies with firms' likely access to external capital markets.

## 3 Data Description and Summary Statistics

### 3.1 Data Source and Sample

The data for this paper come from the firm-level panel compiled by Hall and Hall (1993), "The Value and Performance of U.S. Corporations," downloaded from Bronwyn Hall's data archive. The dataset contains accounting and stock-market information for U.S. corporations identified by the first six digits of their CUSIP code and followed annually from 1960 to 1991.

In the Stata implementation, I construct a firm identifier from the CUSIP code and declare the data as a panel with firm and year as the two dimensions. This yields an unbalanced panel with 1,962 firms and 27,566 firm-year observations in total. Firms enter and exit the sample at different dates and are observed for varying numbers of years, reflecting listing and delisting on U.S. stock exchanges. My estimation sample drops observations with missing values in the variables used in the baseline specification (investment, Tobin's $q$, cash flow, debt, R&D, advertising, and sales). All variables used in the paper come from the same web source (Hall and Hall's `pstar` dataset); any additional transformations, such as logs and interactions, are constructed by me in Stata. I also construct two time dummies to capture the major oil-price shock episodes in the sample. The variable $\text{oil7374}_t$ equals 1 for observations in 1973 and 1974 and 0 otherwise. The variable $\text{oil7981}_t$ equals 1 for observations in 1979, 1980, and 1981 and 0 otherwise. These indicators are designed to absorb the common shifts in investment associated with the 1973–74 and 1979–81 oil crises and the associated recessions. In the pooled regressions without full year fixed effects, I include these oil-crisis dummies as additional controls. Table 1 lists the core variables used in the analysis.

### 3.2 Summary Statistics

For detailed summary statistics and figures, see Appendices 1–4. The estimation sample contains 27,566 firm-year observations. Average investment intensity is about 0.10 of beginning-of-year assets (sd $\approx 0.08$, range 0–2.38). Average Tobin's $q$ is 1.62 but with a very wide range (0–407), and annual sales average about 730 million USD, again with substantial dispersion. The log-transformed variables $\ln(\text{inva})$, $\ln(\text{vala})$, and $\ln(\text{sales})$ have much tighter distributions and are used in the main regressions.

The extreme observations come from only a handful of firm-year pairs. The lowest investment-to-assets ratio in the estimation sample is 0.00032 for firm 636635 (CUSIP 636635) in 1977, while the highest value, 2.38, occurs for firm 250568 (CUSIP 250568) in 1990. Similarly, Tobin's $q$ ranges from essentially zero (0.00032 for firm 265720 in 1969) to about 407.09 (firm 911843 in 1990). These extreme observations motivate the use of log transformations and robustness checks, but they represent only a tiny fraction of the 27,566 firm-year observations in the panel.

Appendix 2 shows that log Tobin's $q$ has a roughly symmetric interquartile range around zero but a number of extreme low and high outliers. Appendix 4 plots $\ln(\text{inva})$ against $\ln(\text{vala})$ and reveals a clear but noisy positive slope, consistent with $q$-theory. The correlation matrix in Appendix 3 confirms a moderate correlation between $\ln(\text{inva})$ and $\ln(\text{vala})$ (about 0.33), strong comovement between cash flow and advertising intensity, and reasonably low correlations between $\ln(\text{inva})$ and the proposed instruments, while the instruments are correlated with Tobin's $q$, as desired. The moderate but clearly positive association between $\ln(\text{inva})$ and $\ln(\text{vala})$ supports using a log–log specification as one of the main functional forms in Section 4.

**Table 1.** Core variables used in the analysis

| Variable | Description | Type | Units | Role in analysis | Min. | Max. |
|---|---|---|---|---|---|---|
| ln_inva | Log investment / beginning-of-year assets | Continuous | Log of ratio | Dependent variable | –8.047 | 0.867 |
| ln_vala | Log Tobin's $q$ (total market value / assets) | Continuous | Log of ratio | Key regressor | –8.047 | 6.009 |
| cfa | Cash flow / assets | Continuous | Ratio (share of assets) | Control: internal finance | –5.726 | 17.785 |
| debta | Long-term debt / assets | Continuous | Ratio (share of assets) | Control: leverage | –0.009 | 9.545 |
| rnda | R&D / assets | Continuous | Ratio (share of assets) | Control: R&D intensity | 0.000 | 12.853 |
| adva | Advertising / assets | Continuous | Ratio (share of assets) | Control: advertising intensity | 0.000 | 6.496 |
| ln_sales | Log annual sales | Continuous | Log of million USD | Control: firm size | –6.908 | 11.179 |
| nyseamex | 1 if listed on NYSE or AMEX, 0 otherwise | Dummy (0/1) | Indicator | Control; interaction with ln_vala | 0.000 | 1.000 |
| inva | Investment / beginning-of-year assets | Continuous | Ratio (share of assets) | Underlying variable for ln(inva) | 0.000 | 2.380 |
| vala | Tobin's $q$: total market value / assets | Continuous | Ratio (no units) | Underlying variable for ln_vala, vala2 | 0.000 | 407.094 |
| sales | Annual sales | Continuous | Million USD | Underlying variable for ln(sales) | 0.001 | 71,643.375 |
| vala2 | Tobin's $q$ squared | Continuous | Ratio squared | Used in quadratic specification | 0.000 | 166,000.000 |
| ln_vala_nyse | Log Tobin's $q$ × NYSE/AMEX dummy | Continuous | Log of ratio | Interaction term | –8.047 | 6.009 |
| pstar | Present discounted value of dividends | Continuous | Million USD | Instrument for log Tobin's $q$ in IV | 0.026 | 539.690 |
| h0 | Asset-pricing term $h0$ | Continuous | Index | Instrument for log Tobin's $q$ in IV | 0.000 | 94.311 |
| h1 | Asset-pricing term $h1$ | Continuous | Index | Instrument for log Tobin's $q$ in IV | 0.000 | 84.524 |
| oil7374 | Oil-crisis dummy: 1973–1974 | Dummy (0/1) | Indicator | Time dummy, first oil-crisis period | 0.000 | 1.000 |
| oil7981 | Oil-crisis dummy: 1979–1981 | Dummy (0/1) | Indicator | Time dummy, second oil-crisis period | 0.000 | 1.000 |

*Notes:* All variables are taken from the Hall and Hall (1993) `pstar` dataset; logs, squares, and interactions are constructed by the author in Stata. Minima and maxima are reported over the full available sample.

# 4 Econometric Framework and Modelling

## 4.1 Goal and Main Variables

The goal of this paper is to estimate the causal effect of a firm's stock-market valuation, measured by Tobin's $q$, on its investment intensity. I use a panel of U.S. corporations indexed by $i = 1,\ldots,N$ and years indexed by $t = 1,\ldots,T$. The dependent variable is the investment-to-assets ratio, $\text{inva}_{it}$, defined as investment during fiscal year $t$ divided by beginning-of-year assets for firm $i$. The key explanatory variable is Tobin's $q$, $\text{vala}_{it}$, defined as the ratio of the firm's market value to the book value of its assets.

I also include a set of control variables that may affect investment and may be correlated with Tobin's $q$: cfa (cash flow/assets), debta (long-term debt/assets), rnda (R&D/assets), adva (advertising/assets), sales (later used in logs as a size proxy), and nyseamex (a dummy equal to 1 if the firm is listed on NYSE or AMEX and 0 otherwise). In addition, in the pooled regressions I include two time-dummy variables that capture the major oil-price shock episodes in the sample: $\text{oil7374}_t$ (equal to 1 for 1973–1974, 0 otherwise) and $\text{oil7981}_t$ (equal to 1 for 1979–1981, 0 otherwise). The coefficient on $\text{vala}_{it}$ is interpreted as the change in the firm's investment intensity when Tobin's $q$ increases, holding the other variables constant.

## 4.2 Simple Regression and Functional Form

I begin with a very simple pooled regression that relates the investment-to-assets ratio to Tobin's $q$ and ignores both the panel structure and the other firm-level controls:

$$\text{inva}_{it} = b_0 + b_1\,\text{vala}_{it} + u_{it}. \tag{M1}$$

Here $u_{it}$ is an error term that captures all remaining determinants of investment. Under the assumption $E(\text{vala}_{it}) = 0$, ordinary least squares (OLS) yields an unbiased estimate of $b_1$. I expect $b_1 > 0$: when the market values a firm highly relative to its replacement cost, investment opportunities are more attractive and firms should invest more. Model (M1) provides a useful benchmark but is likely too simple, because it ignores potential non-linearities, omitted firm characteristics, and the panel nature of the data.

To examine whether a linear specification is appropriate, I plot investment against Tobin's $q$ and add a fitted line. If the relationship appears curved, I estimate richer functional forms. One possibility is a quadratic in Tobin's $q$, which allows the marginal effect of $q$ on investment to be stronger at low or moderate values of $q$ and then flatten at higher values. Economically, this would be consistent with firms responding sharply when valuation first signals profitable opportunities, but less strongly once the most attractive projects are already being funded:

$$\text{inva}_{it} = c_0 + c_1\,\text{vala}_{it} + c_2\,\text{vala}_{it}^2 + v_{it}. \tag{M2}$$

A negative coefficient $c_2$ would indicate that the marginal effect of Tobin's $q$ on investment declines at high values of $q$. Because both $\text{inva}_{it}$ and $\text{vala}_{it}$ are positive for most firm-year observations, I also consider a log–log specification:

$$\ln(\text{inva}_{it}) = d_0 + d_1\,\ln(\text{vala}_{it}) + e_{it}. \tag{M3}$$

In this model, $d_1$ is an elasticity: it gives the percentage change in the investment-to-assets ratio associated with a 1 percent change in Tobin's $q$. I compare the goodness of fit of models (M1)–(M3), using measures such as $R^2$ and the root mean squared error, and I select the functional form that fits best and is economically sensible. That chosen form becomes the baseline specification for the rest of the analysis. In all simple regressions I compute heteroskedasticity-robust standard errors.

Figure 1 plots log(investment/assets) against log Tobin's $q$ for all firm-year observations, with a fitted regression line. The cloud of points slopes upward, indicating a clear positive association between $q$ and investment intensity. However, the relationship is quite noisy: most observations are concentrated in a dense

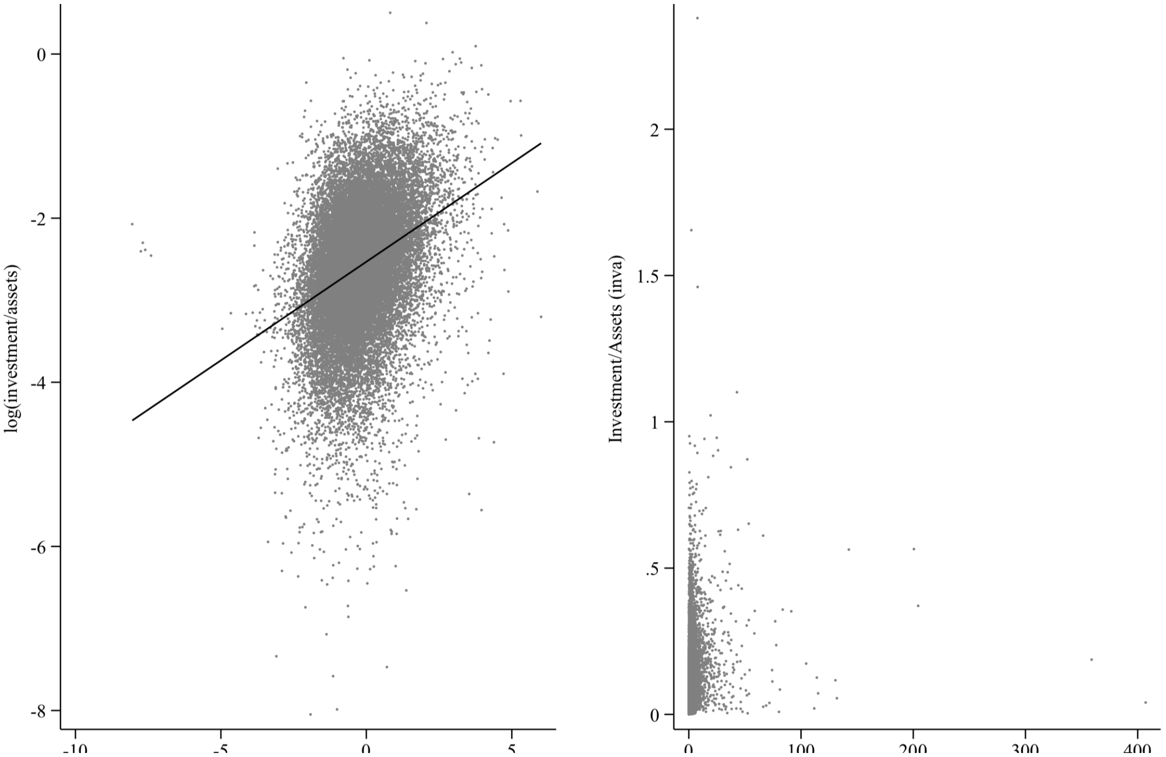


**Figure 1.** Investment vs. Tobin's $q$: raw versus log–log comparison.

oval around the center, with substantial vertical dispersion and a few extreme outliers at very low $q$. This pattern suggests that $q$ helps explain investment but that additional controls and a flexible (log–log) functional form are needed to capture the underlying relationship more accurately.

## 4.3 Multiple Regression with Controls and Interaction

The next step is to extend the preferred simple model by adding firm-level controls that may influence investment and are plausibly correlated with Tobin's $q$. Using the log–log form as the baseline, the pooled multiple-regression specification is

$$\begin{aligned}\ln(\text{inva}_{it}) = \beta_0 + \beta_1 \ln(\text{vala}_{it}) + \beta_2\,\text{cfa}_{it} + \beta_3\,\text{debta}_{it} + \beta_4\,\text{rnda}_{it} + \beta_5\,\text{adva}_{it}\\ + \beta_6 \ln(\text{sales}_{it}) + \beta_7\,\text{oil7374}_t + \beta_8\,\text{oil7981}_t + \varepsilon_{it}.\end{aligned} \tag{M4}$$

Comparing $\beta_1$ in (M4) with the coefficient on $\ln(\text{vala}_{it})$ in the simple log–log model (M3) reveals the extent of omitted-variable bias in the simple regression. If the coefficient on Tobin's $q$ changes substantially after adding cash flow, leverage, R&D, advertising, size, and the oil-crisis dummies, this indicates that these variables were confounding the relationship between valuation and investment. The coefficients $\beta_7$ and $\beta_8$ capture average shifts in investment intensity during the 1973–74 and 1979–81 oil-crisis periods relative to other years. In these pooled multiple regressions I continue to use heteroskedasticity-robust standard errors.

The responsiveness of investment to Tobin's $q$ may differ between firms that are listed on NYSE/AMEX and firms that are not. To allow for this possibility, I add an interaction between $\ln(\text{vala}_{it})$ and the NYSE/AMEX dummy:

$$\begin{aligned}\ln(\text{inva}_{it}) = \gamma_0 + \gamma_1 \ln(\text{vala}_{it}) + \gamma_2\,\text{nyseamex}_{it} + \gamma_3\,[\ln(\text{vala}_{it}) \times \text{nyseamex}_{it}]\\ + \gamma_4\,\text{cfa}_{it} + \gamma_5\,\text{debta}_{it} + \gamma_6\,\text{rnda}_{it} + \gamma_7\,\text{adva}_{it} + \gamma_8 \ln(\text{sales}_{it})\\ + \gamma_9\,\text{oil7374}_t + \gamma_{10}\,\text{oil7981}_t + \eta_{it}.\end{aligned} \tag{M5}$$

In specification (M5), the elasticity of investment with respect to Tobin's $q$ for firms that are not listed on NYSE/AMEX (that is, $\text{nyseamex}_{it} = 0$) equals $\gamma_1$. For listed firms ($\text{nyseamex}_{it} = 1$), the elasticity is $\gamma_1 + \gamma_3$. A test of the null hypothesis $\gamma_3 = 0$ shows whether the response to Tobin's $q$ differs systematically between listed and non-listed firms.

### 4.4 Panel Data and Fixed Effects

The dataset follows each firm over several years, which means that unobserved time-invariant characteristics, such as long-run management quality or corporate culture, may affect both Tobin's $q$ and investment. If these characteristics are not taken into account, the pooled regressions may suffer from bias. To address this concern, I estimate a fixed-effects model that includes a separate intercept for each firm and a separate intercept for each year:

$$\begin{aligned}\ln(\text{inva}_{it}) = a_i + \lambda_t + \delta_1 \ln(\text{vala}_{it}) + \delta_2\,\text{cfa}_{it} + \delta_3\,\text{debta}_{it} + \delta_4\,\text{rnda}_{it} \\ + \delta_5\,\text{adva}_{it} + \delta_6 \ln(\text{sales}_{it}) + v_{it}.\end{aligned} \tag{M6}$$

The term $a_i$ is a firm fixed effect that captures all time-invariant differences between firms, and $\lambda_t$ is a year fixed effect that captures shocks common to all firms in year $t$, such as macroeconomic conditions. In particular, the year dummies absorb the effects of aggregate events like the 1973–74 and 1979–81 oil crises. Because $\text{oil7374}_t$ and $\text{oil7981}_t$ are just specific combinations of the year dummies, including them in the fixed-effects models would be redundant; in practice, they are perfectly collinear with $\lambda_t$ and are dropped by the software. The fixed-effects estimator therefore uses within-firm changes over time, with year fixed effects capturing all common macro shocks, to identify $\delta_1$.

For comparison, I also estimate a random-effects version of (M6). A Hausman test is then used to decide whether the random-effects assumptions are plausible; if the test rejects random effects in favour of fixed effects, I treat the fixed-effects estimates as my main specification. In all panel models I cluster standard errors at the firm level to allow for arbitrary serial correlation within firms.

### 4.5 Endogeneity of Tobin's $q$ and Instrumental Variables

Even after controlling for fixed firm and year effects and adding observed covariates, Tobin's $q$ may remain endogenous in the investment equation. Investment shocks today may influence future stock prices and therefore future $q$, and time-varying investment opportunities or risk factors that are not fully captured by the controls may affect both $q$ and investment. Measurement error in Tobin's $q$ can also bias the estimated coefficient towards zero.

To address these issues, I use a fixed-effects instrumental-variables (IV) approach. I treat $\ln(\text{vala}_{it})$ as an endogenous regressor and use instruments that are strongly related to Tobin's $q$ but plausibly affect current investment only through their impact on $q$. In this dataset, I use the present discounted value of dividends, $\text{pstar}_{it}$, and two additional variables derived from the asset-pricing relation, $\text{h0}_{it}$ and $\text{h1}_{it}$, as instruments. In the first stage, I model the endogenous regressor as a function of the instruments, the observable firm characteristics, and firm and year fixed effects:

$$\begin{aligned}\ln(\text{vala}_{it}) = a_i + \lambda_t + \pi_1\,\text{pstar}_{it} + \pi_2\,\text{h0}_{it} + \pi_3\,\text{h1}_{it} + \pi_4\,\text{cfa}_{it} + \pi_5\,\text{debta}_{it} \\ + \pi_6\,\text{rnda}_{it} + \pi_7\,\text{adva}_{it} + \pi_8 \ln(\text{sales}_{it}) + v_{it}.\end{aligned}$$

Here $a_i$ and $\lambda_t$ are firm and year fixed effects, and $v_{it}$ is the first-stage error term. The coefficients on $\text{pstar}_{it}$, $\text{h0}_{it}$, and $\text{h1}_{it}$ capture how the present value of dividends and the pricing terms shift Tobin's $q$; I assess instrument relevance using the first-stage $F$-statistic on these three variables. The IV fixed-effects model has the same "second-stage" form as the fixed-effects regression in (M6), but $\ln(\text{vala}_{it})$ is now instrumented:

$$\begin{aligned}\ln(\text{inva}_{it}) = a_i + \lambda_t + \theta_1 \ln(\text{vala}_{it}) + \theta_2\,\text{cfa}_{it} + \theta_3\,\text{debta}_{it} + \theta_4\,\text{rnda}_{it} \\ + \theta_5\,\text{adva}_{it} + \theta_6 \ln(\text{sales}_{it}) + \varepsilon_{it}.\end{aligned} \tag{M7}$$

In practice, $\ln(\text{vala}_{it})$ is first regressed on the instruments, the controls, and the fixed effects, and the predicted component is then used in equation (M7). I check instrument strength using the first-stage $F$-statistic reported by the software and, when there are more instruments than endogenous variables, I perform an over-identification test to assess whether the instruments as a group are consistent with the model assumptions. The coefficient $\theta_1$ in model (M7) is my main estimate of the causal effect of Tobin's $q$ on corporate investment. In the results section, I compare this IV fixed-effects estimate with the OLS and fixed-effects estimates to show how accounting for endogeneity changes the conclusions.

# 5 Results

## 5.1 Baseline Specification and Functional Form

Appendix 5 reports three simple pooled regressions of investment intensity on Tobin's $q$. In column (1), the coefficient on Tobin's $q$ is 0.003 with a robust standard error of 0.001, so higher $q$ is significantly associated with higher investment intensity, but the fit is weak ($R^2 = 0.031$, RMSE = 0.079 in levels). In column (2), adding a quadratic term raises $R^2$ only to 0.062, so allowing for curvature in levels brings little improvement. The log–log model in Appendix 5, column (3), performs best: the elasticity of investment with respect to $q$ is 0.240, meaning a 1% increase in $q$ is associated with about a 0.24% increase in investment intensity. The $R^2$ rises to 0.108 (RMSE = 0.711 in logs), a noticeable gain relative to the level specifications. Residual plots for (M3) (Appendix 6) show residuals roughly symmetric around zero and without strong systematic patterns, which supports using the log–log form as the baseline.

I also ran diagnostic tests on these simple models. For the preferred log–log specification (M3), the RESET test reports $F(3, 27{,}561) = 21.37$ ($p < 0.001$), indicating that some non-linearities or omitted variables likely remain, while the Breusch–Pagan test gives $\chi^2(1) = 3.16$ ($p = 0.075$) and White's test $\chi^2(2) = 246.50$ ($p < 0.001$). For the level specifications (M1) and (M2), both RESET and Breusch–Pagan/White tests are even more extreme—for example, for (M1) the BP statistic is $\chi^2(1) = 27{,}510.83$ and White's $\chi^2(2) = 1{,}267.44$ (both $p < 0.001$)—confirming that those models fit the data poorly and exhibit strong heteroskedasticity. Given this evidence and the large cross-sectional variation in the data, I treat heteroskedasticity as likely and therefore report heteroskedasticity-robust standard errors for all pooled regressions and firm-clustered standard errors for panel models in the rest of the paper.

## 5.2 Multiple Regression with Controls

Table 2, column (3), reports the pooled log–log specification with controls, oil-crisis dummies, the NYSE/AMEX dummy, and the interaction between ln(vala) and NYSE/AMEX. In that specification, the coefficient on ln(vala) is 0.277 (s.e. 0.011), so for firms not listed on NYSE/AMEX a 10 percent increase in Tobin's $q$ is associated with about a 2.8 percent increase in the investment-to-assets ratio. The same column also reports positive coefficients on oil7374 (0.207) and oil7981 (0.088), implying higher investment intensity during those years conditional on $q$ and the other controls. Model fit improves relative to the simple log–log regression: Table 2 reports $R^2 = 0.162$ and RMSE = 0.689 in column (3), compared with 0.108 and 0.711 in column (2).

The control coefficients in Table 2, column (3), are economically sensible. Cash flow is 0.206 (s.e. 0.035), debt is 0.259 (0.022), R&D intensity is 0.220 (0.058), and firm size is 0.079 (0.003), while advertising

**Table 2.** Investment/assets and Tobin's $q$: pooled OLS, fixed effects, and FE–IV

| | (1) Linear | (2) Log–log | (3) Pooled OLS | (4) FE | (5) FE–IV |
|---|---|---|---|---|---|
| ln_vala | — | 0.240$^{***}$ | 0.277$^{***}$ | 0.280$^{***}$ | 0.339$^{***}$ |
| | | (0.004) | (0.011) | (0.014) | (0.059) |
| vala | 0.003$^{***}$ | — | — | — | — |
| | (0.001) | | | | |
| ln_vala × NYSE/AMEX | — | — | −0.040$^{***}$ | — | — |
| | | | (0.011) | | |
| cfa | — | — | 0.206$^{***}$ | 0.051 | −0.011 |
| | | | (0.035) | (0.045) | (0.077) |
| debta | — | — | 0.259$^{***}$ | 0.123$^{***}$ | 0.133$^{***}$ |
| | | | (0.022) | (0.026) | (0.029) |
| rnda | — | — | 0.220$^{***}$ | 0.133$^{***}$ | 0.116$^{*}$ |
| | | | (0.058) | (0.045) | (0.063) |
| adva | — | — | −0.950$^{***}$ | −0.063 | −0.011 |
| | | | (0.079) | (0.115) | (0.135) |
| ln_sales | — | — | 0.079$^{***}$ | 0.168$^{***}$ | 0.176$^{***}$ |
| | | | (0.003) | (0.020) | (0.021) |
| NYSE/AMEX dummy | — | — | −0.147$^{***}$ | — | — |
| | | | (0.011) | | |
| oil7374 | — | — | 0.207$^{***}$ | — | — |
| | | | (0.015) | | |
| oil7981 | — | — | 0.088$^{***}$ | — | — |
| | | | (0.012) | | |
| Constant | 0.094$^{***}$ | −2.531$^{***}$ | −2.956$^{***}$ | −3.044$^{***}$ | −3.089$^{***}$ |
| | (0.001) | (0.004) | (0.020) | (0.190) | (0.195) |
| $N$ | 27,566 | 27,566 | 27,566 | 27,566 | 27,566 |
| $R^2$ / within $R^2$ | 0.031 | 0.108 | 0.162 | 0.120 | 0.117 |
| RMSE | 0.079 | 0.711 | 0.689 | 0.545 | 0.546 |

*Notes:* Robust standard errors in parentheses; $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$. The dependent variable is investment/assets in column (1) and log investment/assets in columns (2)–(5). Column (3) is a pooled OLS specification with controls, oil-crisis dummies, the NYSE/AMEX dummy, and the ln(vala) × NYSE/AMEX interaction; column (4) adds firm and year fixed effects; column (5) instruments ln(vala) with pstar, h0, and h1.

intensity is negative at −0.950 (0.079). I therefore continue to report heteroskedasticity-robust standard errors for the pooled regressions. Multicollinearity also appears limited: Appendix 7 reports a mean VIF of 1.83 and a maximum VIF of 3.53 for the pooled specification with the interaction term.

## 5.3 Heterogeneous Effects and Interaction

Table 2, column (3), also shows a negative interaction coefficient on ln(vala) × NYSE/AMEX of −0.040 (s.e. 0.011). This means that being listed on NYSE/AMEX lowers the elasticity of investment with respect to Tobin's $q$ by about 0.04 relative to non-listed firms. The implied elasticity is 0.277 for non-listed firms and 0.237 for NYSE/AMEX firms, so investment remains positively related to $q$ in both groups but is less sensitive among listed firms.

**Table 3.** Panel-data regressions: fixed versus random effects (dependent variable: ln_inva)

| | (1) Fixed effects | (2) Random effects |
|---|---|---|
| ln_vala | 0.280 | 0.281 |
| | (0.0069) | (0.0063) |
| cfa | 0.051 | 0.112 |
| | (0.0200) | (0.0187) |
| debta | 0.123 | 0.149 |
| | (0.0152) | (0.0145) |
| rnda | 0.133 | 0.171 |
| | (0.0298) | (0.0281) |
| adva | −0.063 | −0.400 |
| | (0.0744) | (0.0633) |
| ln_sales | 0.168 | 0.090 |
| | (0.0086) | (0.0052) |
| Constant | −3.044 | −2.772 |
| | (0.232) | (0.231) |
| Year FE | Yes | Yes |
| Firm FE | Yes | No |
| Number of firms | 1,962 | 1,962 |
| Observations | 27,566 | 27,566 |
| Within $R^2$ | 0.120 | 0.116 |
| Overall $R^2$ | 0.107 | 0.151 |
| Hausman test (FE vs RE) | $\chi^2(37) = 310.02,\ p = 0.000$ | |

*Notes:* Standard errors in parentheses. Both specifications include year fixed effects; the fixed-effects model additionally includes firm fixed effects. The Hausman test compares the fixed- and random-effects estimators.

The other coefficients in Table 2, column (3), continue to look reasonable: cash flow, leverage, R&D intensity, and firm size are positively associated with investment, advertising is negative, and the NYSE/AMEX dummy itself is −0.147 (s.e. 0.011), suggesting lower average investment intensity for listed firms conditional on Tobin's $q$ and the other controls. Appendix 7 again reports modest VIFs for this specification (mean 1.83; max 3.53), so multicollinearity does not appear severe. As in earlier pooled models, I interpret these coefficients using heteroskedasticity-robust standard errors.

## 5.4 Panel Data: Fixed and Random Effects

To exploit the panel structure of the data, I next estimate fixed-effects (FE) and random-effects (RE) versions of the baseline log–log model with year dummies, as reported in Table 3. The FE regression uses within-firm variation over time and controls for any time-invariant firm characteristics through firm-specific intercepts. The within $R^2$ is 0.120, only slightly lower than the pooled multiple regression, indicating that a meaningful share of the variation in log investment intensity is explained by within-firm movements in Tobin's $q$ and the controls. The coefficient on log Tobin's $q$ is 0.280 (robust standard error 0.014, $t \approx 20.2$), so even after conditioning on firm fixed effects, year dummies, cash flow, leverage, R&D, advertising, and log sales, higher market valuation is strongly associated with higher investment intensity. The coefficients on debt, R&D, and firm size remain positive and statistically significant, while the cash-flow effect becomes small and imprecise once firm fixed effects are included. An $F$-test of the joint significance of the firm effects,

**Table 4.** Fixed-effects IV regression of log investment/assets on Tobin's $q$ and controls

| Variable | Coefficient | Robust s.e. |
|---|---|---|
| ln_vala | 0.339 | 0.059 |
| cfa | −0.011 | 0.077 |
| debta | 0.133 | 0.029 |
| rnda | 0.116 | 0.063 |
| adva | −0.011 | 0.135 |
| ln_sales | 0.176 | 0.021 |
| Constant | −3.089 | 0.195 |
| Observations | 27,566 | |
| Firms | 1,962 | |
| Within $R^2$ | 0.117 | |
| Overall $R^2$ | 0.115 | |
| Wald $\chi^2$ | 4,549.75 | |

*Notes:* Firm and year fixed effects included; robust standard errors clustered by firm in parentheses. ln(vala) is instrumented with pstar, h0, and h1.

$F(1{,}961, 25{,}567) = 9.40$ with $p < 0.001$, strongly rejects the hypothesis that all firm intercepts are equal, providing clear evidence that unobserved, time-invariant firm heterogeneity matters.

The RE model produces a very similar elasticity of investment with respect to Tobin's $q$ (0.281 with standard error 0.006), but the Hausman test comparing FE and RE, $\chi^2(37) = 310.02$ with $p < 0.001$, rejects the null hypothesis that the RE estimator is consistent. This indicates that the unobserved firm effects are correlated with Tobin's $q$ and other regressors, so the RE assumptions are not credible in this setting. I therefore treat the FE specification with firm-clustered standard errors as my preferred panel model and use it as the benchmark for interpreting the causal effect of Tobin's $q$ on corporate investment in the remainder of the paper.

## 5.5 Instrumental Variables and the Endogeneity of Tobin's $q$

In this subsection I treat log Tobin's $q$ as endogenous and estimate a fixed-effects IV model using the present discounted value of dividends (pstar) and the asset-pricing terms h0 and h1 as instruments. The FE–IV regression (Table 4) is estimated with firm and year fixed effects and firm-clustered standard errors. The IV estimate of the elasticity of investment with respect to Tobin's $q$ is 0.339 (robust s.e. 0.059, $z \approx 5.7$, $p < 0.001$, 95% CI $\approx [0.223, 0.455]$). This is noticeably larger than the corresponding FE OLS estimate (about 0.28), suggesting that ignoring endogeneity—through measurement error or feedback from investment to market value—tends to understate the responsiveness of investment to valuation.

Most control coefficients remain economically sensible. Leverage (debta) and firm size (ln_sales) are still positive and precisely estimated: a one-unit increase in debta is associated with about a 0.13 increase in log investment/assets (s.e. 0.029), and a 1% increase in sales raises investment intensity by roughly 0.18% (coefficient 0.176, s.e. 0.021). By contrast, cash flow, R&D, and advertising lose significance once Tobin's $q$ is instrumented, which is consistent with part of their apparent effect in the OLS and FE models being absorbed by the corrected valuation term. The within $R^2$ of the FE–IV model ($\approx 0.117$) is very close to the FE OLS value, and the regressors are strongly jointly significant (Wald $\chi^2(37) \approx 4{,}550$, $p < 0.001$), indicating that the IV specification still explains a substantial share of within-firm investment variation while correcting for endogeneity in Tobin's $q$.

As a check on the IV strategy, I estimated a pooled 2SLS version of the model and ran standard instrument-validity tests (Appendix 8). The first-stage regression shows that the instruments are strongly relevant: the joint $F$-statistic for pstar and h1 is 187.1 (partial $R^2 \approx 0.034$, $p < 0.001$). The over-identification test is less comforting: the score $\chi^2(1) = 4.79$ ($p = 0.029$) rejects the null that both instruments are perfectly exogenous, suggesting that at least one may be weakly correlated with the investment disturbance. At the same time, the endogeneity tests for $\ln(\text{vala})$ give $\chi^2(1) = 1.83$ ($p = 0.18$) and $F(1, 27{,}527) = 1.72$ ($p = 0.19$), so I cannot reject exogeneity of Tobin's $q$. I therefore view the FE–IV estimates as a robustness check that yields a somewhat larger elasticity of investment with respect to $q$, rather than a clearly preferred baseline relative to the FE-OLS specification.

### 5.6 Goodness of Fit and $F$-Tests

Across specifications, the models have moderate explanatory power. In the pooled regressions shown in Table 2, the $R^2$ rises from 0.108 in the simple log–log model (column (2)) to 0.162 in the pooled specification with controls and interaction (column (3)). In the panel models, Table 2 reports within $R^2$ values of 0.120 for the fixed-effects model (column (4)) and 0.117 for the FE–IV model (column (5)), indicating that a meaningful share of within-firm variation in investment intensity is explained by Tobin's $q$ and the controls. Table 5 also shows that the regressors are jointly highly significant across specifications. These results imply that the added controls, interaction term, and fixed effects improve fit while leaving the core positive relationship between $q$ and investment intact.

### 5.7 Oil-Crisis Dummies and Robustness

The sample includes two major oil-price shocks (1973–74 and 1979–81). To check whether these episodes drive the results, I re-estimate the pooled log–log models including dummies for each crisis period.

Adding these dummies to the pooled specification does not materially change the estimated elasticity of investment with respect to Tobin's $q$. In the pooled specification reported in Table 2, column (3), both oil dummies are positive, indicating higher average investment intensity in 1973–74 and 1979–81 conditional on $q$ and the other controls, but the main $\ln(\text{vala})$ coefficient remains positive and precisely estimated.

The same conclusion holds for the interaction model (M5): the $q$ elasticities for NYSE/AMEX and non-listed firms remain similar to the baseline, and the difference between them stays large and significant.

In the fixed-effects and fixed-effects IV models, year fixed effects already absorb these shocks, so adding oil-crisis dummies is redundant and does not change the results. Overall, the findings are robust to controlling for major oil-price shocks and are not driven by a few extreme macroeconomic episodes.

**Table 5.** Summary of $F$-tests

| Model | Description | Test | Statistic | d.f. | $p$-value | Interpretation |
|---|---|---|---|---|---|---|
| M3 | Simple log–log pooled OLS: ln(inva) on ln(vala) | $F$-test of all slopes $= 0$ | $F = 3345.63$ | (1, 27,564) | $< 0.001$ | ln(vala) is highly jointly significant; even the simple model explains non-trivial variation in ln(inva). |
| M4 | Pooled OLS with controls | Robust $F$-test of all slopes $= 0$ | $F = 575.09$ | (6, 27,559) | $< 0.001$ | All regressors ($q$, cash flow, debt, R&D, advertising, size) are jointly significant. |
| M5 | Pooled OLS with NYSE/AMEX interaction | Robust $F$-test of all slopes $= 0$ | $F = 473.82$ | (8, 27,557) | $< 0.001$ | Extended model with interaction remains strongly jointly significant. |
| M6 | Firm fixed-effects (with year dummies) | Cluster-robust $F$-test of all slopes (incl. year dummies) | $F = 42.32$ | (37, 1,961) | $< 0.001$ | Covariates and time effects are jointly significant after absorbing firm fixed effects. |
| FE test | Firm fixed effects (non-robust FE) | $F$-test that all firm effects $u_i = 0$ | $F = 9.40$ | (1,961, 25,567) | $< 0.001$ | Reject $H_0$ of no firm heterogeneity; firm fixed effects are strongly warranted relative to pooled OLS. |

*Notes:* $F$-statistics and degrees of freedom as reported by Stata; $p$-values below 0.001 are denoted $< 0.001$.

# 6 Discussion and Extension

The estimates across models tell a consistent story: investment intensity is positively related to Tobin's $q$, but the magnitude of this link depends on how much heterogeneity and endogeneity I control for. In the simple pooled log–log regression shown in Table 2, column (2), the elasticity of investment with respect to $q$ is 0.240. In the pooled specification with controls and the NYSE/AMEX interaction, reported in Table 2, column (3), the elasticity for non-listed firms is 0.277, while for listed firms it is 0.237 after accounting for the $-0.040$ interaction term. The $R^2$ also rises from 0.108 to 0.162 between columns (2) and (3), so Tobin's $q$ remains central even after conditioning on standard firm-level determinants.

Allowing the elasticity to differ by listing status reveals economically meaningful heterogeneity. As Table 2, column (3), shows, the estimated elasticity is 0.277 for firms not listed on NYSE/AMEX and 0.237 for listed firms. This pattern suggests that investment at smaller or less prominent firms is more sensitive to changes in valuation, consistent with the idea that these firms face tighter financial frictions and react more strongly when market value improves.

The panel estimates highlight the importance of unobserved firm characteristics. Once I include firm and year fixed effects (M6), the elasticity rises to about 0.280 (s.e. $\approx 0.014$), and the Hausman test strongly rejects random effects in favor of fixed effects ($\chi^2(37) \approx 310$, $p < 0.001$). This indicates that pooled OLS understates the responsiveness of investment to $q$ because it ignores time-invariant firm traits—such as management quality, technology, or corporate culture—that are correlated with both valuation and investment.

The IV results provide an upper-bound perspective on causality. In the fixed-effects IV model (M7), instrumenting log $q$ with pstar, h0, and h1 yields an elasticity of about 0.339 (s.e. $\approx 0.059$). In a complementary pooled 2SLS regression with year dummies, the elasticity is about 0.220 (s.e. $\approx 0.027$), and the first-stage statistics indicate strong instrument relevance: the partial $R^2$ on $\ln(\text{vala})$ is about 0.034, and the robust first-stage $F$ is approximately 187.1 ($p < 0.001$). At the same time, the over-identification test rejects the joint exogeneity of the instruments at the 5 percent level ($\chi^2(1) \approx 4.79$, $p \approx 0.029$), while the endogeneity tests do *not* reject the null that $\ln(\text{vala})$ is exogenous ($p \approx 0.18$–$0.19$). This pattern suggests that the IV estimates should be interpreted cautiously: they likely bracket the true causal effect from above, while the fixed-effects estimate around 0.28 is a more conservative and credible central value.

Several extensions could deepen the analysis. A first extension would be to allow sector-specific elasticities—interacting $\ln(\text{vala})$ with industry dummies—to test whether high-tech, R&D-intensive, or regulated sectors react more strongly to valuation shocks than traditional manufacturing or utilities. A second extension would be to explore non-linearities around the theoretical benchmark $q = 1$, for example by estimating models that allow the elasticity to differ when $q$ is below, near, or well above one, to see whether firms only respond strongly once they are clearly "in the money." A third extension would be to move toward a dynamic panel framework that includes lagged investment and lagged $q$, capturing adjustment costs and expectations more explicitly and connecting the reduced-form results more closely to formal $q$-theory. These extensions would all build directly on the existing specification while offering sharper tests of how and when Tobin's $q$ translates into real investment.

# 7 Conclusion

In the preferred fixed-effects OLS specification with firm and year dummies and financial controls (M6), the elasticity of investment with respect to log Tobin's $q$ is about 0.28 (s.e. $\approx 0.014$). In the corresponding fixed-effects IV model (M7), where log Tobin's $q$ is instrumented by the present discounted value of dividends and the asset-pricing terms h0 and h1, the elasticity rises to about 0.34 (s.e. $\approx 0.059$). With an average Tobin's $q$ of roughly 1.62 in the sample, this IV elasticity of 0.34 implies a semi-elasticity of about 0.21 in a levels specification—that is, a coefficient of approximately 0.21 on Tobin's $q$ in levels in a regression of ln(investment/assets) on $q$.

Taken together, these estimates imply that a 1 percent increase in Tobin's $q$ is associated with roughly a 0.25–0.30 percent increase in investment intensity, depending on the specification, while a one-unit increase in $q$ around its mean (for example, from 1.5 to 2.5) corresponds to about a 0.21 increase in log investment/assets. This directly answers the central research question posed in the introduction: Tobin's $q$ does have a causal effect on firms' investment-to-assets ratios, but the magnitude is moderate rather than dominant.

Economically, the results support the core intuition of $q$-theory: firms that the market values more highly relative to their asset base tend to invest more. At the same time, the modest size of the elasticity and the significant roles of cash flow, leverage, R&D, advertising, and firm size indicate that Tobin's $q$ is only one of several channels through which financial conditions shape real investment. Compared with the existing literature—such as Blundell et al. (1992) and Bond et al. (2003), who typically report smaller coefficients on $q$ in similar investment equations—my preferred fixed-effects and FE–IV estimates suggest a somewhat stronger, but still far from overwhelming, sensitivity of investment to valuation.

The diagnostic and instrument-validity tests point to heteroskedasticity and raise some concerns about instrument exogeneity: first-stage $F$-statistics indicate that the instruments are strongly relevant, but the over-identification test rejects perfect exogeneity at conventional levels, while endogeneity tests do not reject exogeneity of Tobin's $q$. For this reason, the FE–IV estimates with log $q$ and their implied levels interpretation are best viewed as upper bounds, with the fixed-effects OLS elasticity around 0.28 serving as a more conservative central estimate.

Overall, the evidence from the Hall and Hall panel data confirms that Tobin's $q$ is a useful summary measure of investment opportunities but does not fully determine corporate investment behaviour. Future work that incorporates sectoral heterogeneity, non-linearities around $q = 1$, and dynamic adjustment (for example, through lagged investment and valuation terms) could further clarify when market-based signals are most informative and how they interact with technological constraints and financial frictions in shaping firms' investment decisions.

# References

A. B. Abel and O. J. Blanchard. The present value of profits and cyclical movements in investment. *Econometrica*, 54(2):249–273, 1986.

A. B. Abel and J. C. Eberly. A unified model of investment under uncertainty. *American Economic Review*, 84(5):1369–1384, 1994.

R. Blundell, S. Bond, M. Devereux, and F. Schiantarelli. Investment and Tobin's Q: Evidence from company panel data. *Journal of Econometrics*, 51(1–2):233–257, 1992.

S. Bond, J. A. Elston, J. Mairesse, and B. Mulkay. Financial factors and investment in Belgium, France, Germany, and the United Kingdom: A comparison using company panel data. *Review of Economics and Statistics*, 85(1):153–165, 2003.

W. C. Brainard and J. Tobin. Pitfalls in financial model building. *American Economic Review*, 58(2):99–122, 1968.

R. J. Caballero and E. M. R. A. Engel. Explaining investment dynamics in U.S. manufacturing: A generalized (S, s) approach. *Econometrica*, 67(4):783–826, 1999.

J. G. Cummins, K. A. Hassett, and S. D. Oliner. Investment behavior, observable expectations, and internal funds. *American Economic Review*, 96(3):796–810, 2006.

T. Erickson and T. M. Whited. Measurement error and the relationship between investment and Q. *Journal of Political Economy*, 108(5):1027–1057, 2000.

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

S. Gilchrist and C. P. Himmelberg. Evidence on the role of cash flow for investment. *Journal of Monetary Economics*, 36(3):541–572, 1995.

C. J. Hadlock and J. R. Pierce. New evidence on measuring financial constraints: Moving beyond the KZ index. *Review of Financial Studies*, 23(5):1909–1940, 2010.

B. H. Hall and R. E. Hall. The value and performance of U.S. corporations. *Brookings Papers on Economic Activity*, 1993(1):1–50, 1993.

F. Hayashi. Tobin's marginal Q and average Q: A neoclassical interpretation. *Econometrica*, 50(1):213–224, 1982.

S. N. Kaplan and L. Zingales. Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics*, 112(1):169–215, 1997.

J. H. Stock and M. W. Watson. *Introduction to Econometrics*. Pearson, 4th edition, 2019.

L. H. Summers. Taxation and corporate investment: A Q-theory approach. *Brookings Papers on Economic Activity*, 1981(1):67–140, 1981.

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

T. M. Whited. Debt, liquidity constraints, and corporate investment: Evidence from panel data. *Journal of Finance*, 47(4):1425–1460, 1992.

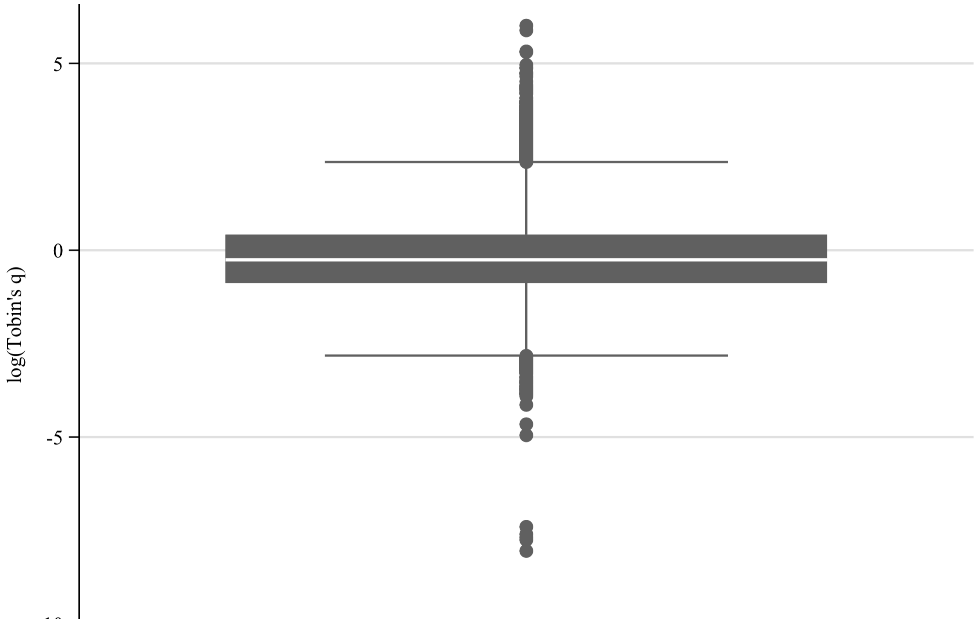
# Appendix

## Appendix 1: Summary statistics for main variables (estimation sample, $N = 27{,}566$)

| Variable | Description | $N$ | Mean | Std. dev. | Min | Max |
|---|---|---|---|---|---|---|
| inva | Investment / beginning-of-year assets | 27,566 | 0.099 | 0.081 | 0.000 | 2.380 |
| vala | Tobin's $q$: market value / book assets | 27,566 | 1.619 | 5.348 | 0.000 | 407.094 |
| cfa | Cash flow/assets | 27,566 | 0.248 | 0.331 | −5.726 | 17.785 |
| debta | Long-term debt / assets (leverage) | 27,566 | 0.276 | 0.300 | −0.009 | 9.545 |
| rnda | R&D / assets (R&D intensity) | 27,566 | 0.045 | 0.163 | 0.000 | 12.853 |
| adva | Advertising/assets | 27,566 | 0.027 | 0.105 | 0.000 | 6.496 |
| sales | Annual sales (million USD) | 27,566 | 729.995 | 2,729.344 | 0.001 | 71,643.380 |
| nyseamex | 1 if listed on NYSE or AMEX, 0 otherwise | 27,566 | 0.697 | 0.459 | 0.000 | 1.000 |
| ln_inva | Log of investment / assets (positive inva) | 27,566 | −2.578 | 0.753 | −8.047 | 0.867 |
| ln_vala | Log of Tobin's $q$ (positive vala) | 27,566 | −0.195 | 1.030 | −8.047 | 6.009 |
| ln_sales | Log of annual sales (million USD) | 27,566 | 4.897 | 1.770 | −6.908 | 11.179 |

*Notes:* Estimation sample, $N = 27{,}566$ firm-year observations.

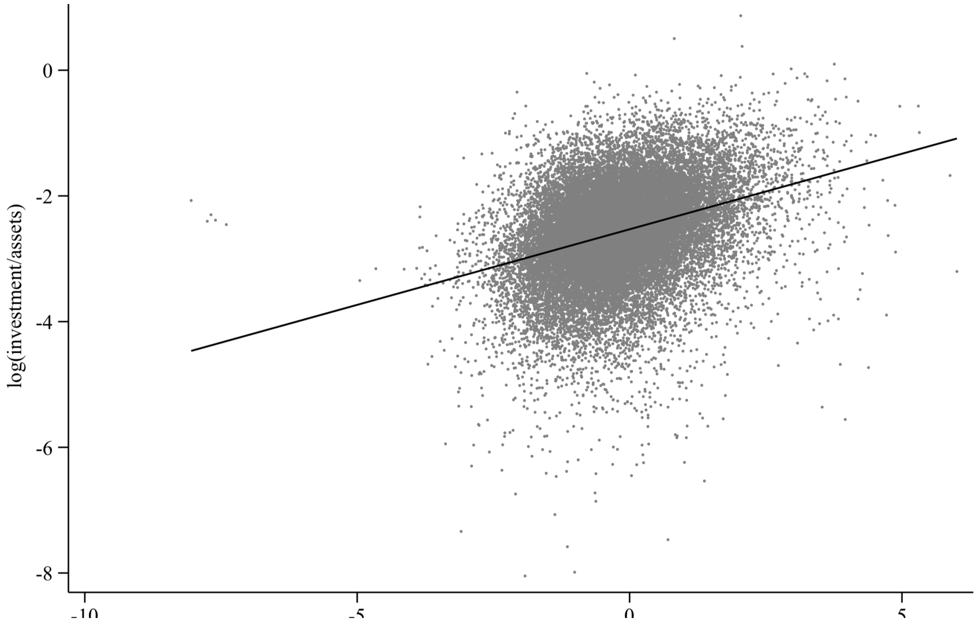
## Appendix 2: Box plot of log Tobin's $q$



## Appendix 3: Correlation matrix (Pearson correlations; main variables and instruments, estimation sample, $N = 27{,}566$)

|  | ln_inva | ln_vala | cfa | debta | rnda | adva | ln_sales | nyseamex | pstar | h0 | h1 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ln_inva | 1.000 | 0.329 | 0.169 | 0.042 | 0.143 | −0.002 | 0.070 | −0.091 | 0.055 | −0.029 | 0.028 |
| ln_vala | 0.329 | 1.000 | 0.431 | −0.122 | 0.322 | 0.170 | −0.183 | −0.172 | 0.050 | −0.053 | 0.055 |
| cfa | 0.169 | 0.431 | 1.000 | −0.093 | 0.096 | 0.603 | 0.012 | −0.120 | 0.066 | −0.044 | −0.003 |
| debta | 0.042 | −0.122 | −0.093 | 1.000 | 0.015 | 0.034 | −0.023 | −0.005 | −0.064 | −0.069 | −0.070 |
| rnda | 0.143 | 0.322 | 0.096 | 0.015 | 1.000 | 0.067 | −0.177 | −0.159 | −0.018 | −0.135 | −0.091 |
| adva | −0.002 | 0.170 | 0.603 | 0.034 | 0.067 | 1.000 | 0.019 | −0.050 | 0.036 | −0.092 | −0.072 |
| ln_sales | 0.070 | −0.183 | 0.012 | −0.023 | −0.177 | 0.019 | 1.000 | 0.373 | 0.410 | −0.081 | 0.030 |
| nyseamex | −0.091 | −0.172 | −0.120 | −0.005 | −0.159 | −0.050 | 0.373 | 1.000 | 0.132 | 0.285 | 0.249 |
| pstar | 0.055 | 0.050 | 0.066 | −0.064 | −0.018 | 0.036 | 0.410 | 0.132 | 1.000 | −0.159 | −0.217 |
| h0 | −0.029 | −0.053 | −0.044 | −0.069 | −0.135 | −0.092 | −0.081 | 0.285 | −0.159 | 1.000 | 0.786 |
| h1 | 0.028 | 0.055 | −0.003 | −0.070 | −0.091 | −0.072 | 0.030 | 0.249 | −0.217 | 0.786 | 1.000 |

## Appendix 4: Scatterplot of log investment/assets and log Tobin’s $q$

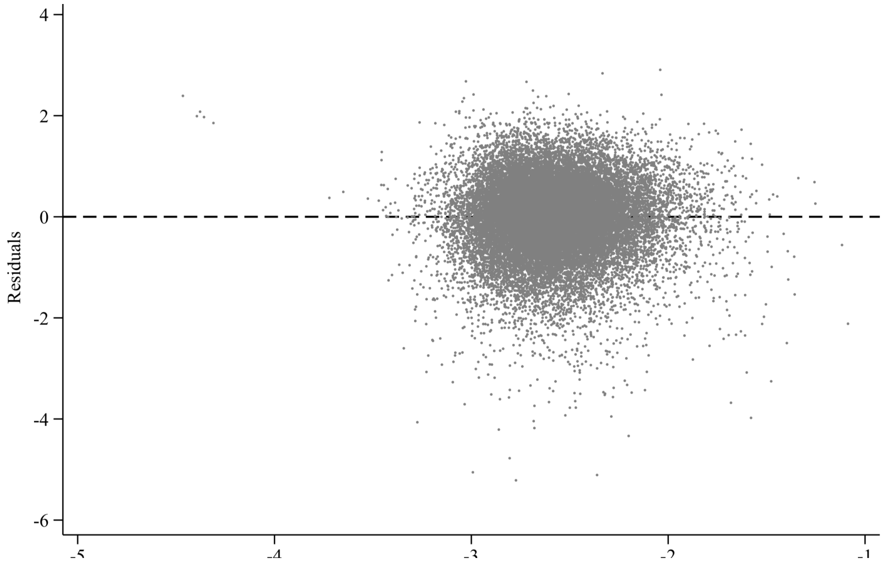


## Appendix 5: Simple pooled regressions of investment intensity on Tobin’s $q$ in three functional forms: (1) linear, (2) quadratic, and (3) log–log

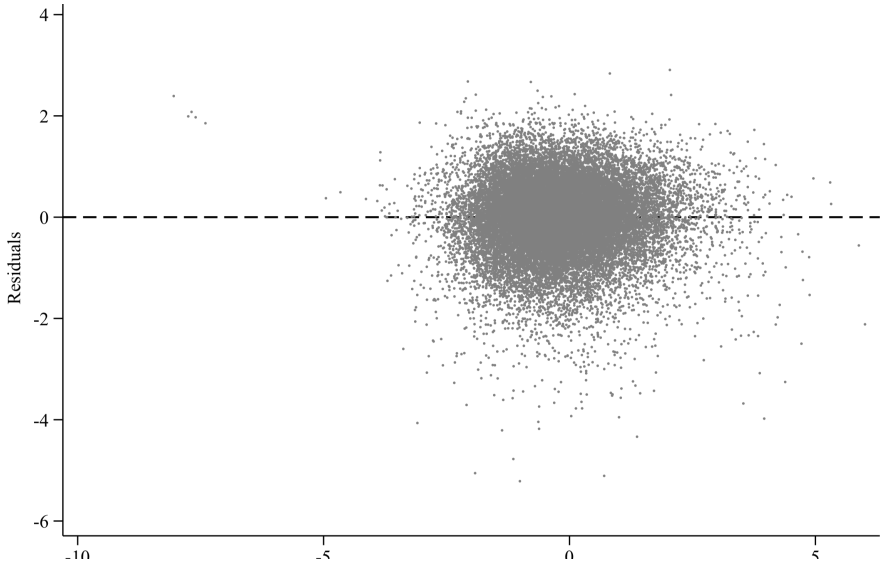
| Variable | (M1) Linear | (M2) Quadratic | (M3) Log–log |
|---|---|---|---|
| vala | 0.003 | 0.006 | — |
| | (0.001) | (0.000) | |
| vala2 | — | −0.000 | — |
| | | (0.000) | |
| ln_vala | — | — | 0.240 |
| | | | (0.005) |
| Constant | 0.094 | 0.089 | −2.531 |
| | (0.001) | (0.001) | (0.004) |
| $N$ | 27,566 | 27,566 | 27,566 |
| $R^2$ | 0.031 | 0.062 | 0.108 |
| RMSE | 0.079 | 0.078 | 0.711 |

*Notes:* Robust standard errors in parentheses. Models (1)–(2) use investment/assets in levels; model (3) uses log investment/assets.

## Appendix 6: Graphs of residuals and predicted values



Residuals versus fitted log investment/assets.



Residuals versus log Tobin's *q*.

## Appendix 7: Multicollinearity diagnostics: variance inflation factors (VIFs)

| Model | Specification (short label) | Variables used for VIF | Max VIF | Mean VIF |
|---|---|---|---|---|
| (1) M1 | Linear pooled OLS (inva on vala) | vala | 1.00 | 1.00 |
| (2) M2 | Quadratic pooled OLS (inva on vala, vala$^2$) | vala, vala2 | 2.57 | 2.57 |
| (3) M3 | Log–log pooled OLS (ln_inva on ln_vala) | ln_vala | 1.00 | 1.00 |
| (4) M4 | Pooled log–log + controls + oil dummies | ln_vala, cfa, debta, rnda, adva, ln_sales | 1.94 | 1.37 |
| (5) M5 | Pooled log–log + controls + NYSE/AMEX interaction + oil dummies | ln_vala, ln_vala_nyse, cfa, debta, rnda, adva, ln_sales, nyseamex | 3.53 | 1.83 |
| (6) M6 | FE model (firm, year FE) – structural RHS via pooled OLS | ln_vala, cfa, debta, rnda, adva, ln_sales, year dummies | – | – |
| (7) M7 | IV model – first-stage for ln_vala | pstar, h0, h1, cfa, debta, rnda, adva, ln_sales, year dummies | – | – |

*Notes:* Variance inflation factors (VIFs) are computed for the pooled OLS specifications (M1–M5). Models (M6) and (M7) are estimated using firm and year fixed effects and instrumental variables, respectively; in such specifications VIFs are not meaningful because the within-transformation and instrument structure create mechanical collinearity with limited within-firm variation. For the IV model (M7), instrument relevance is instead evaluated via the first-stage $F$-statistic.

## Appendix 8: Instrument validity tests for Tobin's $q$ (pooled IV model)

| Test | Name | Null hypothesis | Statistic | $p$-value | Interpretation |
|---|---|---|---|---|---|
| First-stage relevance | Instrument relevance for ln_vala | Instruments (pstar, h1) jointly irrelevant in first stage | $F(2, 27{,}527) = 187.08$ | 0.000 | Strong evidence that instruments predict ln_vala; no weak-instrument problem. |
| Overidentification test | Score $\chi^2$ $J$-test | Instruments jointly exogenous | $\chi^2(1) = 4.79$ | 0.0286 | Null of perfect exogeneity rejected at 5%; at least one instrument may be mildly invalid. |
| Endogeneity of ln_vala | Score $\chi^2$ test | Tobin's $q$ is exogenous | $\chi^2(1) = 1.83$ | 0.176 | Cannot reject exogeneity of ln_vala. |
| Endogeneity of ln_vala | Regression $F$-test | Tobin's $q$ is exogenous | $F(1, 27{,}527) = 1.72$ | 0.189 | Cannot reject exogeneity; IV not statistically required. |

*Notes:* Tests computed from the pooled 2SLS specification with year dummies and instruments pstar, h0, and h1.

## Appendix 9: Replication code (Stata)

This appendix reproduces the Stata do-file used to construct the variables, generate the figures, and estimate every model reported in the paper. The estimation sample is the unbalanced panel of 1,962 firms and 27,566 firm-year observations from Hall and Hall (1993).

```
*===============================================================*
* The Effect of Tobin’s q Ratio on Corporate Investment
* Replication do-file
* Data: Hall & Hall (1993), "pstar" / Invest1993 firm panel
*===============================================================*

import delimited "/Users/ducthanh/Desktop/Econometrics Data/Invest1993/Invest1993.txt"

*--- Panel setup -----------------------------------------------*
capture drop firmid
egen firmid = group(cusip), label
xtset firmid year

*--- Variable labels -------------------------------------------*
label var inva      "Investment / beginning-of-year assets"
label var vala      "Tobin’s q: total market value / assets"
label var cfa       "Cash flow / assets"
label var debta     "Long-term debt / assets"
label var rnda      "R&D / assets"
label var adva      "Advertising / assets"
label var sales     "Annual sales (million USD)"
label var nyseamex  "1 if firm listed on NYSE or AMEX"
label var pstar     "Present discounted value of dividends"
label var h0        "Instrument h0 from Hall & Hall (1993)"
label var h1        "Instrument h1 from Hall & Hall (1993)"
label var year      "Fiscal year"
label var cusip     "Firm identifier (CUSIP)"

*--- Constructed variables -------------------------------------*
capture drop ln_inva ln_vala ln_sales vala2 ln_vala_nyse
gen ln_inva  = ln(inva)  if inva  > 0
label var ln_inva "Log investment / assets"
gen ln_vala  = ln(vala)  if vala  > 0
label var ln_vala "Log Tobin’s q"
gen ln_sales = ln(sales) if sales > 0
label var ln_sales "Log sales"
gen vala2 = vala^2
label var vala2 "Tobin’s q squared"
gen ln_vala_nyse = ln_vala * nyseamex
label var ln_vala_nyse "Log q x NYSE/AMEX dummy"

order inva vala cfa debta rnda adva sales nyseamex ///
      ln_inva ln_vala ln_sales vala2 ln_vala_nyse ///
      pstar h0 h1 year cusip, first

*--- Variable descriptions and summary statistics --------------*
describe inva vala cfa debta rnda adva sales nyseamex ///
         ln_inva ln_vala ln_sales vala2 ln_vala_nyse ///
         pstar h0 h1

summarize inva vala cfa debta rnda adva sales nyseamex ///
          ln_inva ln_vala ln_sales
```

```
tabstat ln_inva ln_vala cfa debta rnda adva ln_sales nyseamex ///
        inva vala sales vala2 ln_vala_nyse pstar h0 h1, ///
        statistics(n mean sd min max) columns(statistics) format(%9.3f)

*--- Estimation sample -----------------------------------------*
capture drop sample
gen sample = !missing(ln_inva, ln_vala, cfa, debta, rnda, adva, ln_sales, nyseamex)

xtset firmid year
xtreg ln_inva ln_vala cfa debta rnda adva ln_sales if sample, fe cluster(firmid)

*--- Appendix 1: summary statistics (estimation sample) --------*
tabstat inva vala cfa debta rnda adva sales nyseamex ///
        ln_inva ln_vala ln_sales if e(sample), ///
        statistics(n mean sd min max) columns(statistics) format(%9.3f)

summarize inva vala cfa debta rnda adva sales nyseamex ///
          ln_inva ln_vala ln_sales if e(sample)

*--- Appendix 2: box plot of log Tobin's q ---------------------*
graph box ln_vala if e(sample), ///
    title("Box Plot of Log Tobin's q") ytitle("log(Tobin's q)")
graph export "appendix2_box_lnvala.png", replace

*--- Appendix 3: correlation matrix ----------------------------*
pwcorr ln_inva ln_vala cfa debta rnda adva ln_sales nyseamex ///
       pstar h0 h1 if e(sample), sig

*--- Appendix 4: scatterplot ln(inva) vs ln(vala) --------------*
twoway (scatter ln_inva ln_vala if e(sample)) ///
       (lfit ln_inva ln_vala if e(sample)), ///
       title("Scatterplot of log investment/assets and log Tobin's q") ///
       xtitle("log(Tobin's q)") ytitle("log(investment/assets)") legend(off)
graph export "appendix4_scatter_lninv_lnvala.png", replace

*===============================================================*
* Models M1-M3: simple pooled regressions
*===============================================================*
xtset firmid year
capture drop sample
gen sample = !missing(inva, vala, ln_inva, ln_vala, cfa, debta, rnda, adva, ln_sales)

* M1: linear in levels   inva = b0 + b1*vala + u
reg inva vala if sample, vce(robust)
estimates store M1

* M2: quadratic   inva = c0 + c1*vala + c2*vala^2 + v
capture confirm variable vala2
if _rc {
    gen vala2 = vala^2
    label var vala2 "Tobin's q squared"
}
reg inva vala vala2 if sample, vce(robust)
estimates store M2

* M3: log-log   ln(inva) = d0 + d1*ln(vala) + e
reg ln_inva ln_vala if sample, vce(robust)
estimates store M3
```

```
estimates table M1 M2 M3, b(%9.3f) se(%9.3f) stats(N r2 rmse) ///
    title("Simple pooled regressions of investment on Tobin's q")

* Predicted values and residuals (Appendix 6)
predict yhat, xb
predict uhat, resid

twoway scatter uhat yhat, yline(0) ///
    title("Residuals vs fitted log investment/assets") ///
    xtitle("Fitted log(investment/assets)") ytitle("Residuals")

twoway scatter uhat ln_vala, yline(0) ///
    title("Residuals vs log Tobin's q") ///
    xtitle("log(Tobin's q)") ytitle("Residuals")

*--- Diagnostics for M1-M3 -------------------------------------*
reg inva vala
estimates store M1_plain
estat ovtest                // Ramsey RESET
estat hettest               // Breusch-Pagan
estat imtest, white         // White test

reg inva vala, vce(robust)
estimates store M1

reg inva vala vala2
estimates store M2_plain
estat ovtest
estat hettest
estat imtest, white

reg ln_inva ln_vala
estimates store M3_plain
estat hettest
estat imtest, white
estat ovtest

*==============================================================*
* Model M4: pooled log-log with controls
*==============================================================*
reg ln_inva ln_vala cfa debta rnda adva ln_sales
estimates store M4_plain
estat ovtest
estat hettest
estat imtest, white
vif

reg ln_inva ln_vala cfa debta rnda adva ln_sales, vce(robust)
estimates store M4

*==============================================================*
* Model M5: pooled log-log with NYSE/AMEX interaction
*==============================================================*
capture drop ln_vala_nyse
gen ln_vala_nyse = ln_vala * nyseamex
estat ovtest
estat hettest
estat imtest, white
vif
```

```
reg ln_inva ln_vala cfa debta rnda adva ln_sales nyseamex ln_vala_nyse, vce(robust)
estimates store M5

* Elasticity for non-listed firms (nyseamex = 0)
lincom ln_vala
* Elasticity for NYSE/AMEX firms (nyseamex = 1)
lincom ln_vala + ln_vala_nyse
* Test whether the two elasticities differ
test ln_vala_nyse = 0

*===============================================================*
* Model M6: fixed vs random effects
*===============================================================*
xtset firmid year

* FE and RE without robust SEs (for Hausman test)
xtreg ln_inva ln_vala cfa debta rnda adva ln_sales i.year, fe
estimates store M6_fe_plain
xtreg ln_inva ln_vala cfa debta rnda adva ln_sales i.year, re
estimates store M6_re_plain

hausman M6_fe_plain M6_re_plain, sigmamore

* Preferred FE with firm-clustered robust SEs
xtreg ln_inva ln_vala cfa debta rnda adva ln_sales i.year, fe vce(cluster firmid)
estimates store M6

xtreg ln_inva ln_vala cfa debta rnda adva ln_sales i.year, re vce(cluster firmid)
estimates store M6_re

*===============================================================*
* Model M7: fixed-effects IV (endogeneity of Tobin's q)
*===============================================================*
ivregress 2sls ln_inva cfa debta rnda adva ln_sales i.year ///
    (ln_vala = pstar h0 h1), vce(robust)
estimates store model_iv_pooled
estat firststage
estat overid
estat endogenous

*--- Joint F-tests ----------------------------------------------*
reg ln_inva ln_vala cfa debta rnda adva ln_sales, vce(robust)
test cfa debta rnda adva

reg ln_inva ln_vala cfa debta rnda adva ln_sales nyseamex ln_vala_nyse, vce(robust)
test ln_vala ln_vala_nyse

xtreg ln_inva ln_vala cfa debta rnda adva ln_sales i.year, fe vce(cluster firmid)

*--- Oil-crisis dummies ----------------------------------------*
capture drop oil7374 oil7981
gen oil7374 = inrange(year, 1973, 1974)
label var oil7374 "Oil crisis dummy: 1973-74"
gen oil7981 = inrange(year, 1979, 1981)
label var oil7981 "Oil crisis dummy: 1979-81"
order oil7374 oil7981, after(year)

reg ln_inva ln_vala cfa debta rnda adva ln_sales oil7374 oil7981 if sample, vce(robust)
```

```
estimates store M4_oil

reg ln_inva ln_vala cfa debta rnda adva ln_sales nyseamex ln_vala_nyse ///
    oil7374 oil7981 if sample, vce(robust)
estimates store M5_oil
```