Journal of Financial Economics 120 (2016) 41–57

Contents lists available at ScienceDirect

Journal of Financial Economics

journal homepage: www.elsevier.com/locate/finec

The expected returns and valuations of private and public firms☆

Ilan Cooper*, Richard Priestley*

Department of Finance, BI Norwegian Business School, Norway

ARTICLE INFO

Article history:
Received 22 March 2013
Revised 30 March 2015
Accepted 1 April 2015
Available online 22 January 2016

JEL Classification:
G0
G12
G31

Keywords:
Real investment
Systematic risk
Mispricing
q theory
Investment returns
Cost of capital
Private firms
Public firms

ABSTRACT

Characteristics play a similar role in describing returns in private firms as in public firms. This evidence suggests a causal effect of optimal investment underlying the role of characteristics, as private firms do not have stock prices to over- or under-react on. Common factor models largely describe the cross section of investment returns of both types of firms, suggesting that the common factors are likely aggregate risk factors. Finally, the cost of capital and firm valuations are similar across private and public firms.

1. Introduction

While all previous assessments of risk, return, cost of equity capital, and valuation ratios have focused on public firms, the importance of private firms in the economy should not be underestimated. For instance, Asker, Farre-Mensa, and Ljungkvist (2015)) find that in 2007 private US firms accounted for 54.5% of aggregate nonresidential fixed investment, 67.1% of private sector employment, 57.6% of sales, and 20.6% of aggregate pre-tax profits. The vast majority of firms in the US are closely held corporations. The 2010 US Census reports seven million corporate tax filers, of which only about eight thousands are public firms. Thus, private firms are an important, but often neglected, part of the economy.

In this paper, we examine the determinants of the cross section of industry investment returns, derived from

☆ This paper was previously circulated as "The cross section of industry investment returns". We thank Doron Avramov, Azi Ben-Rephael, Xi Chen, Thierry Foucault, Fangjian Fu, Ignacio Garcia de Ollala Lopes, Andreea Mitrache, Øyvind Norli, Avi Wohl, Lu Zhang, and seminar participants at Ben-Gurion University, the Center for Corporate Governance Research workshop at the Norwegian Business School, University of Cyprus, University of Haifa, IE Madrid, the Technion, and participants at the 2013 European Finance Association and the 2013 World Finance Conference for helpful comments and suggestions. G. William Schwert (the editor) and an anonymous referee deserve special thanks. Andreea Mitrache provided superb research assistance. We are grateful for the Center for Corporate Governance Research at BI Norwegian Business School, and the Center for Asset Pricing Research at BI Norwegian Business School, for financial support. We thank Eric J. Bartelsman, Randy A. Becker and Wayne B. Gray for graciously making the Manufacturing Industry Productivity Database available. All remaining errors are our own.

* Corresponding authors.
E-mail addresses: ilan.cooper@bi.no (I. Cooper), richard.priestley@bi.no (R. Priestley).

http://dx.doi.org/10.1016/j.jfineco.2016.01.023

the q theory of investment (Cochrane, 1991; Liu, Whited, and Zhang, 2009) within 10 groups of industries differing by the fraction of private and public firms in the industry. We use the National Bureau of Economic Research industry productivity database that aggregates both public and private firms and the Compustat database to sort industries into deciles according to the fraction of the sales (employees) of public firms in the industry to total industry sales (employees). We identify private industries as those industries in the two bottom deciles and public industries as those in the top decile.[1] Examining investment returns of industries that consist mainly of private firms allows us to address three important issues.

First, investment returns are equal to the weighted average cost of capital.[2] Therefore, if the role of characteristics in investment returns in a sample that includes primarily private firms is similar to their role in investment returns of a sample of mostly public firms, this evidence casts doubt on mispricing as an explanation for the role of these characteristics. The reasons for this is that private firms have no stock prices to over- or under-react on and their managers are less susceptible to misvaluation. Instead, the role of characteristics is likely to stem from their presence in the first order conditions of firms' optimal investment decisions.

Our identification scheme of private firms, and the likelihood that these firms do not overreact or underreact to market prices, enables us to interpret characteristic-based factors. If a factor is a true aggregate risk factor, then it should price all equity, whether it belongs to public or private firms, assuming equity holders of both public and private firms require a premium for bearing the factor's systematic risk. To date the literature has examined only the risk-return relation of public firms, and, therefore, it has not been possible to establish whether common risk factors are sources of aggregate uncertainty or are relevant only for firms that are publicly listed on the stock exchange. Many investment-based studies refrain from claiming that characteristics-based factors are risk factors. In contrast, given our identification of private firms, we are able to interpret the role of these factors.

Second, the investment approach renders it feasible for us to obtain estimates for the cost of capital and valuations of private firms. Cost of capital estimates for private firms are notoriously difficult to obtain because of the lack of stock prices. However, by using investment returns, we can obtain the first estimates of the cost of capital of private firms from asset pricing models. Most firms in the economy are private, and being able to obtain a risk-based measure of the cost of capital is crucial to optimal decision making for these firms. Our paper assesses the only means, to the best of our knowledge, of achieving this goal.

Third, following Belo, Xue, and Zhang (2013), we obtain valuation ratios (that is, Tobin's q) implied by firms' first order conditions with respect to investment. Subsequently, we compare the valuation ratios as well as the cross section of valuation ratios of private and public industries. To the best of our knowledge, ours is the first paper to examine the valuation of private firms and to compare them with those of public firms.

Our main findings can be summarized as follows. First, we show that characteristics that have been shown to describe the cross section of stock returns, namely the investment to capital ratio (I/K), the return on assets (ROA) (see Hou, Xue, and Zhang, 2015a; 2015b), size (which we measure as the stock of capital), and idiosyncratic volatility of returns can summarize the cross section of investment returns of both industry portfolios with a relatively large fraction of private firms and industry portfolios with a relatively small fraction of private firms. Therefore, because characteristics share a similar role in describing average investment returns for both private and public firms, their role is unlikely to stem from stock mispricing simply because private firms have no stock price. Instead, the role of characteristics appears to stem from their fundamental part in the first order conditions for investment decisions (Lin and Zhang, 2013).

Second, a four-factor model derived from the q theory of investment, similar to that in Hou, Xue, and Zhang (2015a), composed of the "market" investment return, an I/K factor, an ROA factor, and a size factor performs well in describing the cross section of investment returns of 20 characteristic-based industry portfolios. The portfolios are composed of five I/K portfolios, five ROA portfolios, five portfolios sorted by idiosyncratic volatility of returns, and five portfolios sorted by the size of the capital stock.

The model performs well in terms of small pricing errors and a large cross-sectional $\bar{R}^2$. This is the case irrespective of the fraction of private firms in each portfolio. Therefore, because the risk factors affect both public and private firms, they are likely to be true aggregate risk factors in that they are aggregate sources of uncertainty in the economy.

Third, based on the estimates from the four-factor model, we calculate the cost of capital (expected return) for all industries and industries with varying degrees of private firms in them.[3] The differences in these estimates across private and public firms are generally small, suggesting that private and public firms have similar costs of equity. No systematic difference exists in the cost of capital in the sense that private firms always have a higher (lower) cost of capital than public firms. Our findings of a similar cost of capital for public and public firms are consistent with Moskowitz and Vissing-Jørgensen (2002) who use estimates of private firm value and profits at the aggregate level and study the returns to aggregate entrepreneurial investment. Fourth, we find that private industries have valuation ratios and a cross-sectional

[1] The two bottom deciles consist of industries with only private firms.

[2] Cochrane (1991) demonstrates this theoretically for equity-only firms. Liu, Whited, and Zhang (2009) show that expected investment returns are equal to the expected weighted average cost of capital for portfolios sorted on characteristics that give large spreads in average stock returns.

[3] Due to lack of data on industries' capital structure in our database, we can provide evidence on the weighted average cost of capital but not on the cost of equity and the cost of debt separately.

variation of valuation ratios that are similar to those of public industries.

The rest of the paper is organized as follows. In Section 2, we illustrate the equivalent role of characteristics and covariances in returns and elaborate on the advantages of our identification scheme of private firms. Section 3 describes the data and variable construction. Section 4 describes the econometric methodology of estimating the adjustment cost parameters, and Section 5 presents the empirical findings. The paper concludes in Section 6.

2. Identification

In this section we describe the equivalence of the roles of characteristics and covariances in average returns. Our analysis follows closely Lin and Zhang (2013). Subsequently we argue that our scheme of identifying private firms enables us to infer the source of the role of characteristics in investment returns.

2.1. The equivalent role of characteristics and covariances

The role of firm characteristics in describing the cross section of average stock returns has led to the claim that mispricing is prevalent in the economy. Daniel and Titman (1997) show that characteristics dominate covariances in summarizing the cross section of average stock returns.[4] These findings are part of the backbone of the evidence suggesting investors exhibit behavioral biases [see the discussion in Barberis and Thaler, 2003].

However, Lin and Zhang (2013) show that in general equilibrium, just like covariances, firm characteristics are sufficient statistics for expected stock returns, and expected stock returns are determined endogenously jointly with covariances [as in the consumption approach of Lucas, 1978] and firm characteristics [as in the investment approach of Cochrane, 1991]. Therefore, the search for mispricing through running horse races of covariances against characteristics is pointless. Moreover, characteristics dominate covariances in return regressions because, as Lin and Zhang (2013) show, the former are measured more precisely. However, this says nothing about mispricing. Evidence that characteristics dominate covariances is consistent with both rational and irrational pricing.

We follow Lin and Zhang (2013) and show the equivalence between the role of characteristics and covariances. In the typical consumption economy with no production, the agent's first order consumption problem results in the following well known expression for expected returns:

$$E_t\left[M_{t+1} r^s_{i,t+1}\right] = 1, \tag{1}$$

where M_{t+1} is the stochastic discount factor and $r^s_{i,t+1}$ is the gross return on stock i. Cochrane (2005) shows how to use the definition of covariance to write expression (1) in terms of a beta pricing model:

$$E_t\left[r^s_{i,t+1}\right] - r_f = \beta^M_i \lambda_M, \tag{2}$$

where $r_f = \frac{1}{E_t[M_{t+1}]}$ is the risk-free rate, $\beta^M_i = -cov(r^s_{i,t+1}, M_{t+1})/var(M_{t+1})$ is the loading of $r^s_{i,t+1}$ on M_{t+1}, and λ_M is the price of risk defined as $var(M_{t+1})/E_t[M_{t+1}]$.

Turning to a production economy with adjustment costs, Cochrane (1991) shows that stock returns can be written in terms of characteristics:

$$r^s_{i,t+1} = \frac{\pi_{i,t+1}}{1 + a\left(\frac{I_{i,t}}{K_{i,t}}\right)}, \tag{3}$$

where $\pi_{i,t+1}$ is firm i's productivity given a set of random aggregate shocks, $I_{i,\,t}$ is firm investment, $K_{i,\,t}$ is firm capital stock, and a is an adjustment cost parameter. Lin and Zhang (2013) focus on the equivalence between these two approaches:

$$r_f + \beta^M_i \lambda_M = E_t\left[r^S_{i,t+1}\right] = \frac{E_t[\pi_{i,t+1}]}{1 + a\left(\frac{I_{i,t}}{K_{i,t}}\right)}, \tag{4}$$

where the first term presents the expression for expected returns in terms of covariances and the final term in terms of characteristics. Rearranging makes the relation between covariances and characteristics clearer:

$$\beta^M_i = \left(\frac{E_t[\pi_{i,t+1}]}{1 + a\left(\frac{I_{i,t}}{K_{i,t}}\right)} - r_f\right) \Big/ \lambda_M. \tag{5}$$

In a general equilibrium framework with positive adjustment costs, expected stock returns, covariances, and characteristics all become endogenous. No causal relation exists among these variables. No causality runs from covariances to expected returns, from characteristics to expected returns, or vice versa. Therefore, showing that risk factors (covariances) or characteristics are important in stock return regressions does not mean that they describe expected returns. We can say nothing about the rationality of prices from these approaches. However, we can say nothing about irrationality either. The point is that characteristics can show up in the cross section of returns because of their role in the firm's first order investment decision or because of mispricing.

2.2. Identification

Our approach of examining industries that are composed mostly of private firms can be of help in identifying the driving forces behind the role of characteristics in describing the cross section of average stock returns. The reason is that private firms do not have stock prices to over- or under-react on. Therefore, private firms are less dependent on investor sentiment and less subject to investor misvaluations. Thus, we are able to shed more light on the ongoing debate on the role of characteristics in the cross section of average returns.

[4] More recent examples are Daniel, Hirshleifer, and Teoh (2002), Barberis and Thaler (2003), Richardson, Tuna, and Wysocki (2010), Dechow, Khimich, and Sloan (2011) and Hirschleifer, Hou, and Teoh (2012).

Consider Cochrane (1991), who shows that investment returns are equal to stock returns of an unlevered firm. The equivalence between stock returns and investment returns allows us to use investment returns for private firms. This enables us to address two central and important issues. First, if characteristics and loadings on risk factors are important in determining expected returns in a similar manner for private and public firms' investment returns, then the role of characteristics in general is likely to be due to the first order production decisions of firms and not due to mispricing. That is, to the extent that managers of private firms are less affected by investor sentiment or valuation mistakes regarding their firms than investors in the stock market and managers of public companies, finding that characteristics drive the cross section of investment returns among private firms would lend some support to the idea that it is the fundamental first order investment decision that describes the role of characteristics in the cross section of stock returns.

Second, what is the cost of capital for private firms and public firms? This issue has not been addressed before in a risk-return framework. Asset pricing tests that use private firms as part of the sample have a further advantage. If a factor that is related to returns is a true risk factor then a necessary condition is that it is a source of aggregate uncertainty affecting all firms in the economy. To our knowledge, the extant literature has focused asset pricing tests entirely on returns of public firms because of the availability of stock returns. Consequently, assessing whether these factors are an aggregate source of uncertainty is not possible. By including private firms, we are able to assess whether risk factors are an aggregate source of uncertainty.

At first blush, though, it could be thought that the findings we present regarding the role of characteristics and mispricing should be considered cautiously. The reason is that the lack of stock prices does not necessarily imply that investment returns are not affected by overvaluation or undervaluation of the firm. For example, if a certain characteristic indicates that a public firm's stock is overpriced and subsequent stock returns are abnormally negative, then the same characteristic could be associated with abnormally high real investment due to managers' overvaluation of investment projects followed by negative abnormal investment returns for private firms. However, to the extent that managers of firms, and especially of private firms, are less affected by investors' misvaluation concerning the firm than investors in the stock market, our results are consistent with a rational-based explanation for the role of characteristics in summarizing expected stock returns.

Our claim that the results are most consistent with a rational-based explanation are predicated on a number of factors that lead us to believe that the investment returns of private firms are less likely to be affected by investors' misvaluations. First, when managers possess private information on which they base their expectations and rational decisions they are likely to ignore investors' misvaluations. Given that private firms are likely to be characterized by more asymmetric information, the influence of investor sentiment is further diminished for these firms. This is collaborated by Hribar and Quinn (2013), who examine the trading patterns of managers and find evidence that they can see through market sentiment.

Second, as noted by Polk and Sapienza (2009), if the market misprices firms according to their level of investment, managers can try to boost short-run share prices by catering to current sentiment. Managers with shorter shareholder horizons should cater more. Stein (1996) argues that managers with short horizons should be aggressively investing when investors are overly optimistic. However, this mechanism is unlikely to exist within private firms. Asker, Farre-Mensa, and Ljungkvist (2015)) present evidence consistent with managers of public firms being short-termist and managers of private firms not being short-termist.

Third, while managers of private firms could still raise capital through private placements when their firms are overvalued, being non-short-termist implies they will use the proceeds for investment in T-bills instead of undertaking negative net present value (NPV) projects (Stein, 1996). Fourth, Cooper and Priestley (2011) find that the investment-future stock return relation can be explained without recourse to arguments based on overinvestment or investor overreaction. They find that differences in systematic risk between high and low investment firms can describe the differences in average stock returns between high and low investment firms.

Overall, while we cannot fully rule out that investment returns of private firms are affected by sentiment or other behavioral biases, certainly they are less likely to be. Therefore, our findings that the same characteristics and risk factors are relevant for both private and public firms points to the conclusion that the role of characteristics in both private and public firms' investment returns and the previous reported role of them in stock returns is unlikely to be related solely to mispricing.

3. Data and variable construction

We use the Bartelsman, Becker and Gray NBER-CES Manufacturing Industry Productivity Database, available on the NBER website, as well as the Compustat database.[5] The NBER database contains annual four-digit standard industrial classification (SIC) industry-level data on output, investment, capital stock, and other industry-related variables for all four-digit manufacturing industries in the US for the period 1958–2009. The data cover 459 manufacturing industries and are collected from various government sources, with many of the variables taken directly from the US Bureau of the Census Annual Survey of Manufacturers (ASM) and Census of Manufacturers. The ASM is a survey of approximately 60 thousand establishments, carried out by the Census Bureau. Bartelsman and Gray (1996) provide a detailed description of the database.

Our primary variable of interest is the rate of return on investment. We follow Liu, Whited, and Zhang (2009)

[5] The NBER Database is available at http://www.nber.org/nberces/.

and assume a quadratic adjustment cost function and derive the investment return as follows:

$$r^I_{i,t+1} = \frac{(1-\tau_{t+1})\left[MPK_{t+1} + \frac{a}{2}\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)^2\right] + \tau_{t+1}\delta_{i,t+1} + (1-\delta_{i,t+1})\left[1 + (1-\tau_{t+1})a\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)\right]}{\left[1 + (1-\tau_{t+1})a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right]} \quad (6)$$

where MPK is the derivative of the firm's profit function with respect to capital, K is the stock of capital, I is investment, δ is capital depreciation, and a is an adjustment cost parameter. A larger value of a implies that the industry is facing higher adjustment costs of investment. See the Appendix A.

As Liu, Whited, and Zhang (2009) note, the investment return given in Eq. (6) is the ratio of the marginal benefit of an additional unit of installed capital (marginal q) to the marginal cost of installing an extra unit of capital. The term $(1-\tau_{t+1})[MPK_{t+1}]$ is the marginal after-tax profit produced by an extra installed unit of capital. The term $(1-\tau_{t+1})[\frac{a}{2}(\frac{I_{i,t+1}}{K_{i,t+1}})^2]$ is the marginal after-tax reduction in adjustment costs caused by having an extra unit of installed capital. The term $\tau_{t+1}\delta_{i,t+1}$ is the marginal depreciation tax shield, and the last term in the numerator of Eq. (6) is the marginal continuation value of an extra unit of capital net of depreciation.[6]

We follow Gilchrist and Himmelberg (1995) and assume that the profit function is homogenous of degree one, implying that marginal profit, MPK, is the ratio of realized earnings to the firm's stock of capital.

To calculate industry investment returns, we need several data items and estimates. We use the real capital stock series from the NBER database for the capital stock K. Investment, I, is given by total capital expenditures, deflated by a deflator for that series to obtain investment in real terms, where both capital expenditure per industry and the investment deflator are from the NBER database. To calculate MPK and ROA, we also need earnings. We define earnings by subtracting total payroll from value added, and deflating this difference by the value of shipment deflator.

We are also interested in the valuation ratio, namely, Tobin's q, derived from the q-theory of investment. We follow Belo, Xue, and Zhang (2013) and derive Tobin's q from the first order condition with respect to investment of the firm's optimization problem. See the Appendix A. Under our quadratic adjustment cost specification Tobin's q is

$$q_{it} = 1 + (1-\tau_t)a\left(\frac{I_{it}}{K_{it}}\right). \quad (7)$$

Both stock and flow variables at the NBER database are recorded at the end of year t. However, the model requires flow variables subscripted t to be measured over the course of year t. Therefore, for the numerators of the ratios MPK, I/K, and ROA of year t, which are all flow variables, we use the end of year t values, and for the denominators of these ratios, which all are stock variables, we use values at the end of year $t-1$. MPK as well as ROA in year t are defined as the end of year t earnings of the industry divided by the end of year $t-1$ stock of capital of the industry. The investment-to-capital ratio, I/K, in year t is defined as the ratio of the industry's investment in year t to its capital stock in the end of year $t-1$.

We follow Liu, Whited, and Zhang and measure τ_t, the corporate tax rate, as the statutory corporate tax rate. The source for the tax data is the Commerce Clearing House annual publications.

We use Compustat data on depreciation and amortization (item DP) to compute industry-level rates of depreciation as follows. For each two-digit SIC industry in each year, we sum the depreciation of all firms in that industry and divide by the sum of capital stocks of all firms on Compustat in the industry. In each year, each four-digit SIC industry is assigned the depreciation rate of the two-digit industry for that year. We use the item DP from the Compustat database due to the lack of depreciation data at the NBER database.

For the industry-specific adjustment cost parameter a, we apply a generalized method of moments (GMM) estimation and estimate the valuation equation. We winsorize industry characteristics, namely, the size of the capital stock, the investment-to-capital ratio, and the return on assets, at the 1% level to reduce the impact of outliers and potential estimation errors.[7]

3.1. Identifying private and public industries

We identify industries with mostly private firms and industries consisting mostly of public firms as follows. For each year, we sort industries into deciles by the ratio of the

[6] The price of an installed unit of capital is equal to its marginal value (marginal q), which under optimality equals the marginal cost of investment given by $a\left(\frac{I_{i,t+1}}{K_{i,t+1}}\right)$. Thus, the last term in the numerator of Eq. (6) reflects the value of the undepreciated extra unit of capital.

[7] Not winsorizing yields very similar results for all of the empirical tests conducted in the paper. A potential problem when using the NBER database to calculate industry investment returns is the fact that the data are only for US-based variables. That is, there is no information in this database on the stock of capital of US industries held abroad, as opposed to Compustat which includes data on total firm capital held domestically and abroad. However the required return on investment in the stock of capital held in the US should not be affected by the exclusion of capital held in other countries for the following reason. If a firm undertakes an investment project in the US, it will require a rate of return on that investment that either corresponds to the risk of the project, or is related to some behavioral biases the firms' managers have. Thus, the risk-return relation for such projects can be studied independently of capital held in foreign countries. This is similar to examining the cross section of average stock returns in a subsample of the Center of Research in Security Prices (CRSP) database. For example, in a subsample that contains only NYSE stocks. Any asset pricing model would contend that average returns of firms in that subsample of firms are related to their riskiness or to some characteristics.

Table 1
Descriptive statistics.
This table reports industry-year (i.e. panel data) statistics from the National Bureau of Economic Research (NBER) Manufacturing Industry Productivity Database. The statistics reported are mean, standard deviation, and skewness, as well as the 5th, 25th, 50th, 75th, and 95th percentiles for investment-to-capital ratio (*I*/*K*), return on assets (*ROA*), size of the real capital stock in real 1987 billions of dollars (*K*), Tobin's *q* (*q*) and investment returns (r^I). Panel A reports these statistics for the sample that contains all 459 four-digit standard industrial classification code manufacturing industries in the NBER database. The mean, standard deviation (in parentheses), and skewness (in square brackets) for 10 groups (the first two deciles are in the column titled "Private") sorted by the fraction of the sales of listed firms in the industry to total industry sales are presented in Panel B. "Private" refers to the two deciles with the lowest fraction, and "Public" refers to the summary statistics of the decile with the highest fraction. The sample period is 1960–2009. The p-Values in Panel B [*P*(*diff*)] are for the hypothesis that the means of private and public industries are the same and are computed by a bootstrap approach.

Panel A. All industries

Variable	Mean	Standard deviation	Skewness	5%	25%	50%	75%	95%
I/*K*	7.72%	4.32%	2.42	2.53%	4.90%	6.98%	9.55%	15.17%
ROA	0.86	1.55	46.66	0.23	0.46	0.69	1.03	1.90
K (billion $)	2.45	5.80	7.36	0.11	0.44	0.97	2.24	8.29
r^I	8.56%	49.78%	29.24	−32.43%	−9.64%	5.34%	20.59%	54.33%

Panel B. Portfolios sorted by the fraction of sales of listed firms to total industry sales

Variable	Private	3	4	5	6	7	8	9	Public	*P*(*diff*)
I/*K*	7.54% (4.78%) [3.06]	7.55% (4.78%) [3.05]	7.56% (4.67%) [3.09]	7.69% (4.73%) [3.45]	8.19% (4.86%) [4.22]	7.92% (4.65%) [4.66]	7.72% (4.11%) [1.97]	7.64% (4.13%) [1.76]	7.33% (3.97%) [1.08]	0.03
ROA	0.79 (0.60) [2.81]	0.79 (0.60) [2.81]	0.79 (0.59) [2.79]	0.80 (0.58) [3.15]	0.82 (0.59) [3.56]	0.81 (0.55) [3.91]	0.88 (0.64) [3.82]	0.94 (1.22) [11.78]	1.10 (4.45) [19.13]	0.00
K (billion $)	0.94 (1.32) [3.93]	0.95 (1.32) [3.88]	1.11 (1.86) [8.82]	1.48 (3.07) [9.19]	1.89 (4.03) [6.89]	1.92 (3.26) [11.56]	3.89 (8.94) [5.46]	3.94 (7.88) [5.36]	5.25 (10.03) [3.99]	0.00
r^I	8.93% (48.41%) [31.67]	8.91% (48.21%) [31.70]	8.82% (47.71%) [33.22]	7.62% (31.65%) [4.28]	9.84% (29.03%) [1.59]	8.41% (28.73%) [2.65]	7.26% (25.58%) [1.10]	9.82% (38.41%) [6.75]	12.60% (107.26%) [17.78]	0.11

sum of the sales of the public firms in the industry to total industry sales. Industries with mostly private firms are identified as the industries in the lower decile groups. For the lowest two deciles in each year no firms appear in the Compustat database. Hence these deciles consist of purely private industries and we term these industries as private industries. We term the highest decile group as public industries. As a robustness check, we also sort industries into deciles by the ratio of the number of employees of public firms in the industry to total number of employees in the industry.

We use sales data from Compustat, aggregated over all firms in each four-digit SIC industry for the sales of public firms in each industry and we use the non-deflated value of shipment series from the NBER database for total industry sales.[8]

3.2. Descriptive statistics

We now turn to examining some simple summary statistics of the data. Our sample period is 1960–2009 for all of our empirical tests for the following reasons. The sample starts in 1960 because the denominator of the investment return includes the lagged investment-to-capital ratio. For example, *I*/*K* of 1959 appears in the denominator of the investment return for 1960. *I*/*K* of 1959 is defined as the investment of 1959 divided by the capital stock of 1958. Because the data for all items in the NBER database start in 1958, we can construct investment returns from 1960 onward. Our sample ends in 2009 because the most recent year for which data are available at the NBER database is 2009.

Panel A of Table 1 reports the descriptive statistics of the 459 industries. The average of the investment-to-capital ratio over all 22,839 industry years in the sample is 7.72% with a standard deviation of 4.32%. The investment-to-capital ratio is positively skewed as the median is somewhat smaller than the mean (6.98% versus 7.72%). *ROA* exhibits a high skewness of 46.66, in which the median of 0.69 is smaller than the mean of 0.86. The standard deviation of *ROA* is 1.55.

Table 1 also reports the descriptive statistics for the capital stock, *K*. The capital stock is measured in 1987 dollars and is the real capital stock calculated in the NBER database using the perpetual inventory method. The mean capital stock in our sample is $2.45 billion. The standard deviation of the capital stock is high (5.80), and it is positively skewed with skewness of 7.36. The last row of the table reports summary statistics for investment returns (see Section 4.1).

The results for the decile groups sorted by the fraction of sales of publicly listed firms in the industry to total industry sales are presented in Panel B. *I*/*K* is higher for private industries than for public industries (7.54% for the bottom two decile industries versus 7.33% for the top decile industries and the difference is statistically significant). However, no monotonic pattern emerges moving from the lowest deciles to the highest deciles. The volatility and skewness of *I*/*K* decline in general with the fraction of sales of public firms in the industry.

[8] There are many more matches of the SIC codes in the NBER database with the SIC codes from CRSP than with SIC the codes from Compustat. Moreover, all the SIC codes in Compustat that appear in the NBER database also appear in CRSP. Therefore we first match the SIC codes of the NBER database with CRSP, and then we extract Compustat data of those industries using the CRSP/Compustat merged database.

Panel B of Table 1 also shows that *ROA* is smaller for private industries than for public industries. The *ROA* of decile 1 is 0.79, and it increases as the fraction of sales of public firms in the industry rises. The *ROA* of the top decile is 1.10, and the difference between public and private industries is highly statistically significant.

The size of the stock of capital of industries rises substantially as the fraction of public firms in the industry rises, from 0.94 billion dollars for decile 1 to 5.25 billion dollars for decile 10, and the difference is highly statistically significant with a p-Value of 0.00. This pattern indicates that industries with a higher fraction of publicly listed firms are larger than those consisting of mostly private firms. The volatility is in general larger for public industries than for private industries. Finally, the last row of Panel B shows the average annual investment returns of private and public industries (see Section 4.1).

In summary, the I/K ratio and investment returns are similar across private and public firms, *ROA* is higher for pubic industries, and public firms are larger in terms of the capital stock, which is perhaps to be expected.

4. Econometric methodology

To obtain investment returns, we estimate the industry-specific adjustment cost parameter a using GMM to fit the valuation equation moment. We use the investment model specified in Belo, Xue, and Zhang (2013) and consider quadratic adjustment costs. When estimating the parameters at the industry level, specifying convex adjustment costs adds one more parameter to be estimated, the curvature, leading to an unidentified equation (we have only one moment condition but two parameters: the slope and the curvature of the adjustment costs). However, because we specify quadratic adjustment costs, we have only one parameter to estimate. With one moment condition and one parameter for each four-digit industry, the estimation is exactly identified and the moment fits perfectly. We test whether average Tobin's q in the data equals the average q predicted by the model.

The valuation moment condition is

$$E\left[q_{i,t} - \left(1 + (1-\tau_t)a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right)\frac{K_{i,t+1}}{A_{i,t}}\right] = 0, \quad (8)$$

and the valuation error from the empirical moment is defined as

$$e_i^q \equiv E_T\left[q_{i,t} - \left(1 + (1-\tau_t)a\left(\frac{I_{i,t}}{K_{i,t}}\right)\right)\frac{K_{i,t+1}}{A_{i,t}}\right] = 0. \quad (9)$$

Following Belo, Xue, and Zhang (2013), we estimate the adjustment cost parameter, $b \equiv (a)$, by minimizing a weighted combination of the sample moment (8), denoted by g_T. The GMM objective function is a weighted sum of squares of the model errors, that is, $g_T' W g_T$, where W is the identity matrix. Let $D = \frac{\partial g_T}{\partial b}$ and S be a consistent estimator of the variance–covariance matrix of the sample errors g_T. We estimate S using a standard Bartlett kernel with a window length of three. The estimate of b, denoted $\widehat{b}$, is asymptotically normal with variance–covariance matrix $var(\widehat{b}) = \frac{1}{T}(D'^{-1}D'WSWD(D'^{-1}))$. To construct standard errors for the model errors on each four-digit industry or a group of industries, we use $var(g_T) = \frac{1}{T}[I - D(D'^{-1}D'W)]S[I - D(D'^{-1}D'W)]$ which is the variance–covariance matrix for the model errors, g_T. We use a χ^2 test to assess whether the model errors are jointly zero. The χ^2 test is given by $g_T'[var(g_T)]^+ g_T \sim \chi^2$ (#moments − #parameters), where superscript + denotes the pseudo-inversion.

We conduct the GMM estimation at the four-digit industry level. To assess the overall performance of the model, we estimate the adjustment costs using the group of four-digit industries with non-missing items for the sample period 1963–2009.

4.1. Data for the GMM estimation and summary statistics for the investment returns

For the purpose of estimating the adjustment cost parameter, we use only the industries that on average over the sample period had the largest (in the top 25% of the four-digit SIC code industries for which data are available in both the NBER database and the Compustat database) fraction of sales of listed firms to total industry sales. Thus, these industries consist mostly of public firms. We thereby minimize the measurement error due to using Tobin's q when some of the firms in the industry are unlisted because of the items required to compute Tobin's q is the market value of equity which we take from Compustat.

For the investment-to-capital ratio, we use the NBER data instead of Compustat data for the following reasons. First, these two data sets are different. For example, the average firm investment-to-capital ratio in Compustat for manufacturing firms in the period 1963–2009 is 0.29, and the average industry investment-to-capital ratio in the NBER database is 0.08. The difference between the two could stem, for example, from different depreciation methods. By far the most common depreciation method applied by firms (and hence appears in the Compustat database) is the straight-line method. The NBER bases its depreciation patterns on empirical evidence of used asset prices in resale markets wherever possible. For most asset types, geometric patterns are used because the available data suggest that they more closely approximate actual profiles of price declines than straight-line patterns (see Fraumeni, 1997).

Second, an advantage of using the NBER data is that the variables are given in real terms. Hence, these variables could be less susceptible to measurement errors as opposed to Compustat data, which are given in historical cost terms.

The group that consists of the top 25% fraction of sales of listed firms includes one hundred industries. We exclude industries with fewer than two firms on average per year. In doing this, we follow Belo, Xue, and Zhang (2013). This reduces our sample to 76 industries.

Subsequent to estimating the valuation equation, we assign the estimated parameters of the industries we use in the GMM estimation to the other industries as follows. For each SIC code, we assign the estimated parameter of the public industry with four-digits that is closest to that four-digit industry. For example, the four-digit SIC code industry 3412, which is not among the 76 industries for which we estimate the parameters, is assigned the *a* estimate of the industry with SIC 3411, which is among the 76 industries for which we conduct the GMM estimation. This procedure ensures that industries are assigned the parameter values estimated for industries in the same industry group. The SIC code 3412 industry is metal shipping barrels, drums, kegs and pails, and the SIC code 3411 industry is metal cans. Both industry 3411 and industry 3412 belong to the same industry group: metal cans and shipping containers.

To estimate the valuation Eq. (8), we also need the ratio of capital to total assets. Because total assets are not given in the NBER database, we use for this ratio the net capital stock (Compustat item PPENT) and total assets (Compustat item AT) from the Compustat database.

We follow Belo, Xue, and Zhang (2013) in matching the timing of the variables. We include all firms with fiscal year ending in the second half of the calendar year. Tobin's q used in the valuation equation is market value of equity plus debt to total assets (item AT). Total debt, $B_{i,t+1}$, is long-term debt (item DLTT in Compustat) plus short-term debt (Compustat item DLC) for the fiscal year ending in the calendar year $t-1$.

We aggregate the firm-level variables constructed from Compustat data, specifically Tobin's q and the capital-to-assets ratio, at the two-digit SIC level and assign the four-digit industries' variables with the two-digit variables of the industries they belong to. For example, for each two-digit industry in each year, we sum the market value of equity plus debt of all the firms in that industry and divide by the sum of total assets of all firms in that industry. Subsequently, we assign to each four-digit industry the Tobin's q of the two-digit industry that it belongs to.

The reason we resort to using the two-digit variables is that many of the accounting Compustat data items needed have missing values at the four-digit level for many of the four-digit industries. Aggregating at the two-digit level enables us to estimate the parameters for more industries. Moreover, using the NBER database, we find that the cross-sectional variation in the investment-to-capital ratio of four-digit industries within the two-digit industries they belong to is very small relative to the mean investment-to-capital ratio. Hence, using two-digit accounting variables is plausible.

After estimating the adjustment cost parameter, we are able to construct the investment returns for each of the 459 industries. Summary statistics for the investment returns are reported in Table 1. The last row of Panel A shows that for the entire NBER database the mean investment return is 8.56% and the volatility of investment returns is 49.78%. The investment return varies between −32.43% for the 5th percentile and 54.33% for the 95th percentile.

The last row of Panel B of Table 1 shows that the average annual investment returns of public firms (12.60%) is higher than that of private industries (8.93%). However, the difference is not statistically significant at conventional significance levels. Moreover, the pattern of average investment returns is non-monotonic as the fraction of sales of pubic firms in the industry to total industry sales rises.

5. Empirical results

The empirical results are arranged as follows. Section 5.1 reports the GMM estimation results for the 76 industries for which we estimate the adjustment cost parameters. Section 5.2 presents results on the determinants of the cross section of investment returns at the four-digit manufacturing industry level. In the cross-sectional regressions, we focus on the following characteristics. The investment-to-capital ratio (I/K) and the return on assets (*ROA*) summarize the cross section of average stock returns (Hou, Xue, and Zhang, 2015a). We also examine whether size, which we measure as the size of the real capital stock, and idiosyncratic volatility describe the cross section of average investment returns. Idiosyncratic volatility is measured as the standard deviation of the residuals from regressions of industry returns on four factors. We describe the factors in detail below.

Following the factor model in Hou, Xue, and Zhang (2015a), we conduct asset pricing tests by examining the cross-sectional patterns of investment returns when using four investment return-based risk factors (see Section 5.3). These factors are a "market" investment return factor, which we define as the equal-weighted investment return of all industries in our sample, an I/K factor, an *ROA* factor, and a size factor. In Section 5.4, we investigate whether the cost of capital calculated from the asset pricing model varies between public and private firms within the manufacturing sector. Section 5.5 reports estimates of Tobin's q so as to consider differences in valuation ratios across private and public firms.

5.1. GMM estimation results

Table 2 reports the estimates of the adjustment cost parameter, a, along with the corresponding t-statistics. Across all 76 four-digit industries, the estimates are positive and significant. The mean of the estimates of a across the industries is 81.55. To interpret the magnitude of the adjustment costs, we follow Belo, Xue, and Zhang (2013) and report in Table 2 the fraction of lost sales due to adjustment costs $\frac{C}{Y}$, where $C(I_{it}, K_{it}) = \frac{a}{2}\left(\frac{I_{it}}{K_{it}}\right)^2 K_{it}$ is the adjustment cost function, $a > 0$ is the adjustment cost parameter, and Y is sales. We computer the fraction of sales lost as follows. First, for each of the selected four-digit industries we compute the time-series of adjustment costs. Second, for each year we divide the adjustment costs by sales (value of shipment from the NBER database) to obtain the ratios of adjustment costs-to-sales and take the average of these ratios in the time-series. This average is reported in Table 2.

We also report the average ratio of adjustment costs to sales across all industries. On average, the implied adjustment costs represents 12.21% of sales. The cross-sectional standard deviation of the average fraction of lost sales is 9.94%. The magnitude of the implied adjustment costs varies largely across the industries. Belo, Xue, and Zhang (2013) report an average of sales lost due to adjustment costs of 5.94% on average across the Fama and French 30 industries, which is within one standard deviation of our

Table 2
Generalized method of moments (GMM) estimation.
This table reports the one-step GMM results at the four-digit SIC industry level from estimating the valuation moment given by Eq. (8). The estimate of adjustment costs is denoted by *a*. The corresponding *t*-statistics are reported along with $\frac{C}{Y}$, which is the ratio in percent of the implied capital adjustment costs over sales. The sample period is 1963–2009.

SIC	a	$t(a)$	$\frac{C}{Y}$
2043	58.97	5.05	8.74
2061	75.91	5.44	18.14
2063	66.88	5.50	15.80
2086	71.90	8.38	9.46
2121	113.12	7.45	8.35
2321	117.71	6.35	5.34
2329	101.36	6.43	6.40
2341	146.84	5.71	5.05
2522	47.80	6.71	7.93
2542	55.09	10.22	5.96
2621	14.27	5.18	4.99
2672	12.10	7.03	1.90
2731	53.48	6.11	5.08
2761	70.01	6.58	10.75
2771	73.57	7.51	10.01
2812	53.58	7.54	21.56
2813	67.51	9.83	35.59
2823	122.20	4.72	16.62
2834	48.46	6.75	11.64
2835	33.26	4.67	17.83
2891	56.26	7.68	7.63
2952	9.12	4.28	0.90
3069	27.96	5.60	3.49
3143	234.20	4.93	8.39
3149	278.91	3.35	11.78
3221	26.05	6.03	4.51
3241	22.66	4.18	10.09
3316	11.99	4.48	1.02
3411	73.14	4.97	6.38
3423	68.12	6.43	7.15
3431	121.11	10.75	7.41
3442	64.98	7.68	5.56
3494	72.61	5.73	7.72
3499	67.80	7.33	7.22
3531	77.71	5.26	8.39
3532	89.12	4.41	7.25
3533	75.18	3.43	13.06
3546	79.86	5.74	8.97
3549	67.54	6.23	7.04
3559	65.44	5.80	9.61
3561	75.93	6.17	7.49
3564	76.93	7.38	6.69
3567	76.09	6.70	5.42
3569	76.48	6.07	9.25
3572	44.18	4.75	63.03
3578	79.83	7.33	13.47
3581	91.02	6.61	10.70
3589	62.23	6.91	6.62
3613	77.93	7.21	6.69
3629	90.69	8.32	9.19
3634	92.86	5.38	6.93
3643	69.92	6.73	8.13
3651	82.07	5.51	9.30
3661	75.68	7.84	10.26
3663	46.57	7.04	11.55
3669	47.89	6.38	8.35
3674	36.43	5.58	48.07
3679	46.63	5.69	11.47
3699	51.02	5.42	9.10
3716	29.57	6.90	1.76
3751	30.54	4.89	4.58
3822	142.43	5.76	18.09
3823	125.28	5.06	17.80
3825	117.88	4.46	28.48
3826	110.93	3.46	24.76
3829	122.41	5.35	18.90
3841	75.24	6.43	24.14
3842	105.17	5.06	21.18
3843	109.94	9.28	17.08
3845	64.42	4.12	20.37
3851	109.63	4.08	28.90
3873	198.06	7.53	11.45
3914	145.21	4.08	23.54
3931	123.38	6.12	10.46
3942	211.60	4.12	11.49
3949	81.71	6.71	12.57

(*continued on next page*)

average estimate. Bloom (2009) surveys the estimates of convex adjustment costs to be between zero and 20% of revenue. Thus, our estimates are in line with those reported in previous studies.

5.2. Characteristics and the cross section of private and public firms' investment returns

In Table 3, for each of the groups of industries sorted by the fraction of sales of public firms in the industry to total industry sales, we run year-by-year cross-sectional Fama and MacBeth regressions of industry investment returns on industry characteristics.

Panel A of Table 3 reports the results for univariate cross-sectional regressions of investment returns on the one-year-lagged investment-to-capital ratio. That is, we regress investment returns in year t on the ratio of investment in year $t-1$ to capital in year $t-2$. The first column presents the results for deciles 1 and 2, which contain only private firms. Consistent with the result for stock returns (Xing, 2008), the coefficient on the investment-to-capital ratio is negative for private industries. The coefficient on lagged I/K is −2.21 and it is statistically significant with a t-statistic of −7.22. The $\overline{R}^2$ in this regression is 6.75%. The size of the coefficient (in absolute value) in general declines as the fraction of sales of public firms in the industry rises. However, all the coefficients across the 10 groups of industries are negative and statistically significant, with t-ratios ranging from −2.45 to −9.85. The coefficient on lagged I/K for public industries (decile 10) is −1.01 (with a t-statistic of −2.45), and it is statistically significantly different from the coefficient of −2.21 for private industries. The size of the coefficient in Xing (2008, Table 3) is considerably larger (−4.75) in absolute value relative to our estimates (−1.01 to −2.19). This could be because we use investment returns and industry portfolios, and Xing uses individual stock returns. The $\overline{R}^2$s range from 3.94% to 8.13%, versus 1% in Xing (2008). Overall, Panel A shows that the lagged investment to capital ratio effect is important for all firms and in particular for private firms.

The finding regarding the role of I/K in describing the cross-sectional variation in investment returns among private industries is interesting because one of the behavioral explanations for the investment effect in stock return is a slow reaction of the market to overinvestment by empire-building managers (Titman, Wei, Wei, and Xie,

Table 3
Cross-sectional Fama and MacBeth regressions of investment returns on characteristics: sales.
This table reports coefficients from Fama and MacBeth cross-sectional regressions for 10 industry groups sorted by the fraction of sales of the listed firms in the industry to total industry sales. The group titled “Private” contains in each year the industries in the two bottom decile of the fraction, whereas the group titled “Public” consists of the industries in the top decile fraction in each year. The frequency of the data is annual and the sample period is from 1960 to 2009. The table reports average intercepts and slopes from the cross sectional regressions. *t*-statistics are in parentheses. $\overline{R}^2$ is the average $\overline{R}^2$ of the cross sectional regressions. The p-Values in the last column of each panel are for the hypothesis that the regression coefficients and the $\overline{R}^2$ of private and public industries are equal and are computed by a bootstrap approach.

	Private	3	4	5	6	7	8	9	Public	*P*(*diff*)
Panel A. Investment-to-capital ratio, (*I*/*K*)										
$\hat{\gamma}_0$	0.25	0.25	0.23	0.21	0.21	0.22	0.17	0.22	0.22	0.44
	(8.38)	(8.34)	(8.42)	(9.67)	(9.48)	(10.15)	(8.44)	(8.41)	(8.30)	
$\hat{\gamma}_{I/K}$	−2.21	−2.19	−2.06	−1.67	−1.70	−2.06	−1.25	−1.68	−1.01	0.02
	(−7.22)	(−7.16)	(−7.84)	(−9.22)	(−8.52)	(−9.85)	(−5.40)	(−5.99)	(−2.45)	
$\overline{R}^2$	6.75%	6.68%	6.21%	7.23%	6.83%	8.13%	5.13%	3.94%	3.98%	0.03
Panel B. Return on assets (*ROA*)										
$\hat{\gamma}_0$	0.00	0.00	0.00	−0.00	0.01	0.03	−0.02	−0.05	−0.05	0.02
	(0.16)	(0.16)	(0.02)	(−0.22)	(0.74)	(2.09)	(−1.14)	(−3.13)	(−2.65)	
$\hat{\gamma}_{ROA}$	0.11	0.11	0.10	0.09	0.07	0.04	0.10	0.16	0.14	0.22
	(6.11)	(6.09)	(5.91)	(7.19)	(4.22)	(3.15)	(7.09)	(8.69)	(8.42)	
$\overline{R}^2$	3.05%	3.02%	2.84%	3.36%	2.94%	0.26%	8.65%	17.50%	27.17%	0.00
Panel C. Size (*K*)										
$\hat{\gamma}_0$	0.11	0.11	0.10	0.08	0.08	0.09	0.07	0.08	0.10	0.84
	(6.52)	(6.49)	(6.12)	(5.98)	(5.62)	(5.75)	(6.08)	(4.07)	(6.75)	
$\hat{\gamma}_K$	−2.51	−2.46	−1.94	−1.02	−0.75	−1.12	−0.03	−0.27	0.28	0.00
	(−6.00)	(−5.90)	(−6.20)	(−3.77)	(−3.69)	(−4.95)	(−0.40)	(−0.81)	(2.08)	
$\overline{R}^2$	0.34%	0.34%	0.41%	-0.72%	−1.32%	−0.65%	−0.71%	4.93%	−1.32%	0.00
Panel D. Idiosyncratic volatility (*IVOL*)										
$\hat{\gamma}_0$	0.03	0.03	0.03	0.06	0.06	0.04	0.04	−0.03	−0.08	0.19
	(0.92)	(0.92)	(0.96)	(2.61)	(2.52)	(1.43)	(2.61)	(−0.99)	(−1.04)	
$\hat{\gamma}_{IVOL}$	0.18	0.18	0.16	0.04	0.06	0.13	0.13	0.58	0.70	0.14
	(1.27)	(1.26)	(1.13)	(0.34)	(0.64)	(1.13)	(1.60)	(3.49)	(2.12)	
$\overline{R}^2$	3.56%	3.51%	3.55%	5.40%	1.78%	2.19%	1.82%	8.90%	8.19%	0.17
Panel E. Multiple regressions										
$\hat{\gamma}_0$	0.14	0.14	0.15	0.17	0.15	0.14	0.13	0.07	0.08	0.14
	(4.12)	(4.08)	(4.12)	(6.01)	(5.26)	(4.45)	(4.85)	(2.50)	(3.56)	
$\hat{\gamma}_{I/K}$	−2.82	−2.80	−2.72	−2.34	−2.48	−2.76	−2.37	−2.79	−2.47	0.17
	(−14.86)	(−14.70)	(−14.82)	(−11.36)	(−10.44)	(−11.50)	(−9.91)	(−10.50)	(−15.09)	
$\hat{\gamma}_{ROA}$	0.16	0.16	0.17	0.15	0.16	0.13	0.16	0.21	0.18	0.37
	(10.24)	(10.17)	(10.21)	(12.66)	(8.13)	(8.05)	(9.13)	(15.37)	(12.31)	
$\hat{\gamma}_K$	0.02	0.01	−0.23	−0.23	−0.15	−0.17	0.09	0.35	0.14	0.82
	(0.03)	(0.01)	(−0.62)	(−1.25)	(−0.83)	(−0.61)	(1.16)	(1.78)	(3.14)	
$\hat{\gamma}_{IVOL}$	0.09	0.09	0.00	−0.15	−0.05	0.15	−0.10	0.13	0.03	0.66
	(0.63)	(0.63)	(0.02)	(−1.34)	(−0.50)	(1.34)	(−1.20)	(1.10)	(1.39)	
$\overline{R}^2$	17.19%	17.06%	16.83%	18.31%	14.90%	14.57%	19.76%	29.91%	34.18%	0.00

2004). This explanation is less likely to hold for private firms, for which agency conflicts between managers and shareholders are less likely to be prevalent. The other behavioral explanation for the investment effect in stock returns is market overreaction to firm growth (Cooper, Gulen, and Schill, 2008). As there is no market price for private firms, this explanation could also be less likely to hold for the investment effect within private firms. Our results, and in particular those showing the size of the coefficient is greater for the sample with a higher fraction of private firms, where mispricing could be thought to be less prevalent, lend support to the rational-based explanation of the investment effect. This is consistent with recent findings by Cooper and Priestley (2011) that the spread in stock returns between low investment firms and high investment firms can be largely summarized by loadings on macroeconomic risk factors.

Panel B of Table 3 reports the results for *ROA*. Hou, Xue, and Zhang (2015a) show that the *q*-theory of investment implies a positive relation between return on equity (*ROE*) and future stock returns. Given a certain level of investment, a firm’s riskiness must increase with *ROE* to justify the level of investment. The intuition is as follows. Consider two firms with a given investment-to-capital ratio. As investment is determined by expected future cash flows and by risk, the firm with higher *ROE*, that is, higher expected cash flows, must also have higher risk to explain that its investment-to-capital ratio is not higher. The same intuition applies to *ROA*, which we use in our tests because we lack data on industries’ capital structure. Hou, Xue, and Zhang (2015b) show that the *q*-factor model’s performance is robust to the use of ROA in place of ROE.

Hou, Xue, and Zhang (2015a) show that the risk premium on a stock return factor defined as the excess return of high *ROE* stocks over low *ROE* stocks is 0.58% per month and is statistically significant. In Panel B of Table 3, the coefficients on *ROA* range from 0.04 to 0.16 and are all statistically significant, with *t*-ratios between 3.15 and 8.69. The difference between the coefficients of private and public firms is not statistically significant, with a p-Value of 0.22. The adjusted $\bar{R}^2$s range from 0.26% to as high as 27.17%.

The results for size, namely, the size of the capital stock, are presented in Panel C. The coefficients on size are multiplied by 10^5 as the size of the capital stock is very large relative to returns (the mean industry capital stock is 2.45 billion dollars in our sample). With the exception of decile 10, the size coefficients are all negative and most are statistically significant. The size of the coefficient on size, in absolute value, in general declines as the fraction of public firms in the industry rises, indicating that the effect is stronger for private industries. The adjusted $\bar{R}^2$s are very low and suggest that size is the least important characteristic of the cross section of investment returns.

Panel D of Table 3 presents the results when we regress the current year's investment returns on idiosyncratic volatility (*IVOL*). *IVOL* is defined as the standard deviation of the residuals from time series regressions of industry returns on the four factors, namely the market, *I*/*K*, *ROA*, and size factors using the full sample of annual observations from 1960 to 2009. We describe the factors in details in Section 5.3. Across all 10 deciles, the coefficient on *IVOL* is positive. *IVOL* seems to play a larger role within public industries, where the coefficient is 0.70 (with a *t*-ratio of 2.12), than within private industries, where the coefficient is only 0.18 and is not statistically significant, but the difference is statistically insignificant with a p-Value of 0.14. Moreover, the pattern of the coefficient on *IVOL* is non-monotonic, as the coefficients on deciles 5 and 6 are relatively low. The adjusted $\bar{R}^2$s range from 1.78% to 8.90%. Overall, managers seem to require higher expected investment returns as idiosyncratic risk rises.

Panel E of Table 3 presents multiple regression results, in which the regressors are the variables used in the univariate regressions in the previous panels. The signs of the coefficients on *I*/*K* and *ROA* remain unchanged, and their statistical significance and magnitude are high. Moreover, the magnitude rises as the average coefficients on *I*/*K* and *ROA* across the 10 groups in Panel E are −2.64 and 0.16, respectively, and the corresponding averages in the univariate regressions are −1.80 and 0.10, respectively. As opposed to the univariate regressions, most the signs of the coefficients on size are positive and some of the signs of the coefficients on *IVOL* are negative. *IVOL* positively describes the cross section of investment returns for private industries and is unrelated to returns for pubic industries but the difference between the coefficients on private and public firms is statistically indistinguishable from zero and the coefficients themselves are not statistically significant. The adjusted $\bar{R}^2$s are large, ranging from 14.57% to 34.18%, indicating that the characteristics, jointly, have reasonable explanatory power.

As a robustness check, we repeat the cross-sectional Fama and MacBeth regressions for deciles of industries formed by the fraction of the number of employees of public firms in the industry to the total number of employees in the industry.[9] Table 4 presents the findings. The coefficients on *I*/*K* are all negative, ranging from −1.42 to −2.21 and statistically significant, with *t*-ratios ranging from −5.97 to −10.66 . In contrast to the results in Table 3, the effect of *I*/*K* does not vary considerably across the deciles. The coefficients on *ROA* are all positive and highly statistically significant.

As in Panel C of Table 3, with the exception of the size coefficient for decile 10, the coefficients are negative and some are statistically significant. With the exception of decile 8, the coefficients on *IVOL* is positive. The effect of *IVOL* for public industries is stronger than for private industries (the coefficients on *IVOL* for public and private industries are 0.45 and 0.26, respectively). However the difference is statistically indistinguishable from zero, with a p-Value of 0.47.

The multiple regression results presented in Panel E of Table 4 are largely similar to those when industries are sorted according to the fraction of sales of listed firms to total industry sales in Panel E of Table 3. All of the coefficients on *I*/*K* are negative and highly statistically significant. The *ROA* coefficients are all positive and strongly significant.

Overall, the results in Table 4 are largely similar to those in Table 3 and provide additional evidence of the role of characteristics in summarizing the cross section of average investment returns The results in Tables 3 and 4 show that characteristics that are important for summarizing the cross section of stock returns are also important for describing the cross section of investment returns of private and public firms. Given the large size of the private company sector in the economy, our results are important and lend support to the risk-based explanations for the role of characteristics in summarizing average stock returns based on the investment first order condition.

5.3. Characteristics and investment returns across all industries

In this subsection, we present the results of cross-sectional Fama and MacBeth regressions for the entire sample of 459 industries. Examining the entire sample is beneficial due to the large sample size and consequently the higher power of the tests. Thus, it serves as a robustness check. The results are shown in Table 5. The second to fifth columns present univariate regression results, and the multiple regression results appear in the sixth column. *I*/*K* has a negative sign, with a coefficient of −1.71, and it is statistically significant with a *t*-ratio of −15.54. The $\bar{R}^2$ in the regression is 5.12%. The following column shows that *ROA* helps to describe the cross section of industry investment returns, with a coefficient of 0.14 and a *t*-ratio of 11.38. The $\bar{R}^2$ is relatively large at 17.48%. As opposed to the negative coefficients in Tables 3 and 4, the coefficient on size, presented in the third column of Table 5, is positive but statistically insignificant. The coefficient on

[9] For the first three deciles, the fraction of employees of public firms in the industry to total industry employees is zero.

Table 4
Cross-sectional Fama and MacBeth regressions of investment returns on characteristics: employees.
This table reports coefficients from Fama and MacBeth cross sectional regressions for 10 industry groups sorted by the fraction of employees of the listed firms in the industry to total industry employees. The group titled "Private" contains in each year the industries in the two bottom deciles of the fraction, whereas the group titled "Public" consists of the industries in the top decile fraction in each year. The frequency of the data is annual and the sample period is from 1960 to 2009. The table reports average intercepts and slopes from the cross-sectional regressions. *t*-statistics are in parentheses. $\overline{R}^2$ is the average $\overline{R}^2$ of the cross sectional regressions. The p-Values in the last column of each panel are for the hypothesis that the regression coefficients and the $\overline{R}^2$ of private and public industries are equal and are computed by a bootstrap approach.

	Private	4	5	6	7	8	9	Public	$P(diff)$
Panel A. Investment-to-capital ratio (I/K)									
$\hat{\gamma}_0$	0.23 (9.62)	0.23 (9.57)	0.22 (9.59)	0.23 (9.62)	0.22 (10.17)	0.18 (8.95)	0.20 (10.30)	0.25 (8.38)	0.93
$\hat{\gamma}_{I/K}$	−1.92 (−9.44)	−1.91 (−9.37)	−1.85 (−10.11)	−1.95 (−9.73)	−2.00 (−10.66)	−1.42 (−5.97)	−1.62 (−9.04)	−2.21 (−7.22)	0.07
$\overline{R}^2$	5.94%	5.92%	5.72%	7.61%	8.01%	5.60%	4.38%	6.75%	0.10
Panel B. Return on assets (ROA)									
$\hat{\gamma}_0$	−0.01 (−1.03)	−0.01 (−0.99)	−0.01 (−0.40)	0.01 (0.34)	0.02 (1.42)	−0.03 (−1.66)	−0.04 (−2.14)	−0.04 (−1.98)	0.34
$\hat{\gamma}_{ROA}$	0.12 (8.53)	0.12 (8.46)	0.10 (6.12)	0.08 (4.31)	0.05 (3.01)	0.11 (8.45)	0.13 (7.67)	0.14 (9.11)	0.34
$\overline{R}^2$	10.01%	9.96%	9.20%	3.63%	1.22%	4.19%	7.51%	23.83%	0.01
Panel C. Size (K)									
$\hat{\gamma}_0$	0.08 (5.57)	0.08 (5.52)	0.08 (4.99)	0.11 (6.52)	0.08 (5.41)	0.07 (5.75)	0.08 (5.17)	0.08 (3.71)	0.42
$\hat{\gamma}_K$	−0.01 (−0.03)	−0.00 (−0.02)	−0.12 (−0.62)	−2.51 (−6.00)	−0.86 (−2.24)	−0.62 (−3.08)	−0.32 (−1.20)	1.34 (1.99)	0.00
$\overline{R}^2$	3.16%	3.17%	2.66%	0.34%	−1.01%	−1.30%	−0.06%	1.22%	0.50
Panel D. Idiosyncratic volatility ($IVOL$)									
$\hat{\gamma}_0$	0.01 (0.39)	0.01 (0.38)	0.02 (0.56)	0.07 (2.90)	0.03 (0.99)	0.06 (3.42)	0.05 (2.38)	−0.00 (−0.05)	0.80
$\hat{\gamma}_{IVOL}$	0.26 (1.87)	0.26 (1.87)	0.21 (1.49)	0.03 (0.29)	0.14 (1.07)	−0.00 (−0.03)	0.08 (0.83)	0.45 (1.92)	0.47
$\overline{R}^2$	3.60%	3.54%	3.07%	5.48%	4.29%	0.47%	4.39%	4.27%	0.79
Panel E. Multiple regressions									
$\hat{\gamma}_0$	0.12 (3.72)	0.12 (3.68)	0.15 (4.66)	0.19 (6.79)	0.10 (2.43)	0.10 (3.40)	0.12 (4.66)	0.10 (4.49)	0.59
$\hat{\gamma}_{I/K}$	−2.67 (−17.23)	−2.65 (−17.07)	−2.72 (−16.51)	−2.82 (−9.54)	−2.85 (−10.31)	−2.40 (−10.48)	−2.65 (−11.43)	−2.62 (−11.38)	0.86
$\hat{\gamma}_{ROA}$	0.16 (11.85)	0.16 (11.79)	0.17 (8.17)	0.17 (8.11)	0.17 (8.97)	0.19 (12.75)	0.20 (11.07)	0.19 (14.79)	0.15
$\hat{\gamma}_K$	0.13 (1.41)	0.13 (1.50)	0.15 (1.14)	−0.59 (−2.64)	0.42 (1.44)	0.43 (0.99)	0.31 (1.26)	0.27 (1.70)	0.45
$\hat{\gamma}_{IVOL}$	0.10 (0.71)	0.10 (0.71)	−0.06 (−0.35)	−0.12 (−1.04)	0.17 (1.25)	−0.07 (−0.74)	−0.17 (−1.98)	0.02 (0.56)	0.57
$\overline{R}^2$	23.25%	23.05%	24.23%	21.59%	15.51%	14.47%	21.02%	31.16%	0.14

Table 5
Aggregate cross sectional regressions with characteristics.
This table reports coefficients from Fama and MacBeth cross-sectional regressions of industry investment returns on industry characteristics using the entire sample of 459 manufacturing industries available on the National Bureau of Economic Research Manufacturing Industry Productivity Database. The frequency of the data is annual and the sample period is from 1960 to 2009. The table reports average intercepts and slopes from the cross-sectional regressions. *t*-statistics are in parentheses. $\overline{R}^2$ is the average $\overline{R}^2$ of the cross-sectional regressions.

$\hat{\gamma}_0$	0.22 (13.30)	−0.03 (−2.29)	0.08 (5.90)	−0.01 (−0.33)	0.11 (5.45)
$\hat{\gamma}_{I/K}$	−1.71 (−15.54)				−2.61 (−22.13)
$\hat{\gamma}_{ROA}$		0.14 (11.38)			0.17 (14.26)
$\hat{\gamma}_K$			0.16 (1.17)		0.01 (0.19)
$\hat{\gamma}_{IVOL}$				0.37 (3.36)	0.09 (0.90)
$\overline{R}^2$	5.12%	17.48%	1.09%	3.80%	28.51%

IVOL is 0.37, and it is statistically significant with a *t*-ratio of 3.36, implying that idiosyncratic risk entails a risk premium, consistent with the results of Fu (2009) for stock returns.

The multiple regression results are in the last column of Table 5. The coefficient on I/K is now higher, at −2.61, and is statistically significant, with a *t*-ratio of −22.13. The coefficient on *ROA* also rises from 0.14 to 0.17, and its statistical significance remains high. The size coefficient is 0.01 and is statistically insignificant, and the coefficient on *IVOL* falls from 0.37 to 0.09 and loses its statistical significance. The adjusted $\overline{R}^2$ is 28.51%, indicating reasonable explanatory power.

Overall, the results for the entire cross section of 459 industries are consistent with the role of the characteristics that have previously been examined in the cross section of stock returns and reflect their importance among both private and public firms.

5.4. Asset pricing tests

In this subsection, we assess whether the four factors from the model of Hou, Xue, and Zhang (2015a) can describe the cross section of average investment returns of 20 portfolios formed according to *I*/*K*, *ROA*, size, and idiosyncratic volatility using the cross-sectional regression approach of Fama and MacBeth (1973). To construct the factors, we use the entire sample of 459 industries in the NBER database. We use the market portfolio, formed by equal-weighting the investment returns of all industries. The *I*/*K* factor is defined as the investment return in year t of the low 33% investment-to-capital industries in year $t-1$ over the return on the top 33% investment-to-capital industries in year $t-1$. The *ROA* factor is defined as the year t return of the top 33% *ROA* industries in year $t-1$ over the bottom 33% *ROA* industries in year $t-1$. The size factor is defined as the year t investment return of the bottom 33% industries based in the size of their capital stock, k, in year $t-1$ over the bottom top 33% industries based on the size of their capital stock, k, in year $t-1$. The average investment return on the market portfolio is 8.54% with a t-ratio of 7.66%. The *I*/*K* factor earns a premium of 15.70% and is statistically significant with a t-ratio of 14.70. The average investment return on the *ROA* factor is 12.31% with a t-ratio of 12.85. The average investment return on the size factor is 2.89% with a t-ratio of 1.97.

The asset pricing tests are undertaken using the Fama and MacBeth (1973) procedure. In the first step, a time series regression is employed to estimate the factor loadings (betas) of the portfolio returns. The second step runs cross-sectional regressions of investment returns on the estimated betas to estimate the prices of risk. The use of annual data rules out the typical rolling regression approach to estimate betas for each period. Instead, we use full sample estimates to obtain factor loadings (betas), and in the second step we estimate a cross-sectional regression of average investment returns in each year on the factor loadings estimated over the full sample.

Industries are then ranked according to the fraction of public firms sales to total sales of each industry, and we separate them into quintiles. Within each quintile, we form five portfolios on each of the following: *I*/*K*, *ROA*, idiosyncratic volatility and size. Quintile 1 contains the firms with the lowest fraction of public firms, which is zero in our sample, that is, purely private firms, and quintile 5 contains industries with the highest fraction of public firms.[10] This sorting procedure allows us an approximate comparison of the determinants of the expected returns of private and public firms. For the 20 test assets in each quintile, the following cross-sectional regression is estimated:

$$r_i = \lambda^0 + \lambda^m \widehat{\beta}_{i,m} + \lambda^{I/K} \widehat{\beta}_{i,I/K} + \lambda^{ROA} \widehat{\beta}_{i,ROA} + \lambda^k \widehat{\beta}_{i,k} + e_i, \tag{10}$$

where r_i is the average investment return for the ith portfolio, λ^0 is a constant that should equal the risk-free rate, λ^m is the price of risk of the market factor, $\widehat{\beta}_{i,m}$ is the beta with respect to the market factor, $\lambda^{I/K}$ is the price of risk associated with the *I*/*K* factor, $\widehat{\beta}_{i,I/K}$ is the beta with respect to the *I*/*K* factor, λ^{ROA} is the price of risk associated with the *ROA* factor, $\widehat{\beta}_{i,ROA}$ is the beta associated with the *ROA* factor, λ^k is the price of risk associated with the size factor, $\widehat{\beta}_{i,k}$ is the beta associated with the size factor, and e_i is the residual.

We also report the cross-sectional $\bar{R}^2$, which, following Jagannathan and Wang (1996) and Lettau and Ludvigson (2001), is calculated as $\bar{R}^2 = [Var_c(\bar{r}_i) - Var_c(\bar{e}_i)]/Var_c(\bar{r}_i)$, where Var_c is the cross-sectional variance, $\bar{r}_i$ is the average excess investment return and $\bar{e}_i$ is the average residual. We also assess the performance of the model by calculating the square root of the squared pricing error across all 20 portfolios. Finally, we report a statistic that tests whether the pricing errors are jointly zero. This is a Chi-square test given as $\widehat{\alpha}' cov(\widehat{\alpha})^{-1}\widehat{\alpha}$, where $\widehat{\alpha}$ is the vector of average pricing errors across the 20 portfolios and cov is the covariance matrix of the pricing errors.

Panel A of Table 6 reports the estimated prices of risk for the five quintiles. The first row reports quintile 1, which has only private firms. The market price of risk is estimated to be 2.3% per annum and is not statistically significant with a corresponding t-statistic of 1.19. The price of risk associated with the *I*/*K* factor is estimated to be 10.7% per annum with a t-statistic of 8.27, the price of risk associated with the *ROA* factor is 8.1% per annum with a t-statistic of 5.21, and the estimated price of risk for the size factor is −2.1% per annum but is not statistically significant. The intercept in the cross sectional regression is 6.8% per annum, close to the mean of the risk-free rate of 5.4% for this sample period. The cross-sectional $\bar{R}^2$ is 88%, indicating a good fit. In addition, the pricing errors are economically small, with an average over all portfolios of 1.5% per annum. However, the test of jointly zero pricing errors is rejected.

The remaining rows in Table 6 report the estimated prices of risk for the four quintiles. with increasing amounts of public firms in the portfolios. A great deal of consistency exists in both the estimated prices of risk and the measures of model performance across all five quintiles. For example, comparing the private firms in quintile 1 with the public firms in quintile 5, the estimates and statistical significance of the intercept, market return factors, the *I*/*K* factor, and the *ROA* factor differ very little. It is only the estimate of the price of risk associated with the size factor that changes from −2.1% to 3.7%. However, in both cases the estimates are not significantly different from zero.

Table 6 indicates that a four-factor asset pricing model can successfully summarize the cross section of a sample of portfolios of industries that contain mainly private firms and mainly public firms. To our knowledge, this is the first time that private firms' expected returns are related to systematic risk factors that have been shown to be important in summarizing the cross section of listed firms' expected stock returns. The findings from the four factor model

[10] We use quintile groups in this subsection as opposed to decile portfolios in Section 5.2 (which uses individual industries for the Fama and MacBeth regressions) to ensure a sufficient number of industries within each of the 20 portfolios we form for each group.

Table 6
Cross-sectional regressions with risk factors.
We perform a set of cross sectional regressions of investment returns on factor loadings. The four-factor model, based on Hou, Xue, and Zhang (2015a), is

$$r_i = \lambda^0 + \lambda^m \widehat{\beta}_{i,m} + \lambda^{I/K} \widehat{\beta}_{i,I/K} + \lambda^{ROA} \widehat{\beta}_{i,ROA} + \lambda^k \widehat{\beta}_{i,k} + \epsilon_i,$$

where r_i is the investment return, $\widehat{\beta}_{i,MKT}$ is the factor loading on the market investment return portfolio, $\widehat{\beta}_{i,I/K}$ is the factor loading on the I/K investment return portfolio, $\widehat{\beta}_{i,ROA}$ is the factor loading on the *ROA* investment return portfolio, and ϵ_i is the residual. The factor loadings are estimated over the full sample period. The table reports the constant and the estimated prices of risk (t-values in parentheses). Quintiles are sorted by the fraction of the sales of listed firms in the industry to total industry sales. "Private" refers to the quintile with the lowest fraction, whereas "Public" refers to the quintile with the highest fraction. $R^2 = [Var_c(\bar{r}_i) - Var_c(\bar{e}_i)]/Var_c(\bar{r}_i)$, where Var_c is the cross-sectional variance, $\bar{r}_i$ is the average investment return and $\bar{e}_i$ is the average residual. $\bar{R}^2$ is the adjusted R^2. We define the pricing error for a given portfolio i as the difference between the actual investment return and the expected investment return according to the cross-sectional test. *pe* represents the square root of the aggregate squared pricing errors across all portfolios in each division (p-Value in brackets). The test assets are 20 portfolios, five each according to *I/K, ROA*, lagged investment, and size of the capital stock. The sample period is 1960–2009.

Panel A. Quintiles sorted by sales

γ_0	γ_{MKT}	$\gamma_{\frac{I}{K}}$	γ_{ROA}	γ_K	$\bar{R}^2$	pe_{ALL}	χ^2_{ALL}
Quintile 1: Private							
0.068 (4.52)	0.023 (1.19)	0.107 (8.27)	0.081 (5.21)	−0.021 (1.11)	0.882	0.015	52.498 [0.00]
Quintile 2							
0.073 (4.58)	0.020 (1.05)	0.129 (10.50)	0.079 (5.36)	−0.023 (1.19)	0.886	0.014	39.292 [0.00]
Quintile 3							
−0.089 (5.44)	0.157 (8.21)	0.132 (8.10)	0.069 (5.00)	0.163 (6.17)	0.625	0.019	46.607 [0.00]
Quintile 4							
−0.078 (4.72)	0.162 (8.16)	0.142 (7.90)	0.068 (4.65)	0.118 (5.23)	0.450	0.024	42.614 [0.00]
Quintile 5: Public							
0.059 (1.42)	0.034 (0.95)	0.122 (7.16)	0.076 (3.14)	0.037 (1.48)	0.900	0.023	39.094 [0.00]

Panel B. Aggregate

−0.019 (1.29)	0.105 (5.73)	0.132 (11.79)	0.083 (8.02)	0.047 (2.74)	0.873	0.014	69.668 [0.00]

indicate that these factors are a source of aggregate uncertainty in the sense that they have a similar role in a sample of mainly private firms and a sample of mainly public firms. We can confirm the aggregate nature of the risk factors by assessing the performance of the model using all industries and thus ignoring the split between private and public firms.

Panel B of Table 6 reports the estimated prices of risk from the 20 test portfolios when we use all industries. The market price is estimated to be 10.5% per annum and is statistically significant. The price of risk associated with the I/K factor is 13.2% per annum with a t-statistic of 11.79, and the estimated price of risk associated with the *ROA* factor is 8.3% per annum with a t-statistic of 8.02. The estimated price of risk associated with the size factor, which was statistically significant only for two of the quintiles, is 4.7% per annum with a t-statistic of 2.74. The cross-sectional $\bar{R}^2$ is 87%, similar to the findings when looking at both private and public firms separately. The pricing errors for the four factor model are low with an average across all 20 portfolios of 1.4% per annum. However, the Chi-square test rejects the null hypothesis that the 20 pricing errors are jointly zero.[11]

[11] As a robustness check, we repeat the asset pricing tests using quintiles sorted by the fraction of the number of employees of listed firms in the industry to the total number of employees in the industry. The results are largely similar and are available in the Online Appendix.

The evidence presented here shows that a factor model motivated from the q-theory of investment is able to successfully summarize the cross-sectional differences in the 20 portfolios formed on four characteristics that include a substantial number of unlisted firms. This is an important finding because it rules out, at least to some extent, the possibility that characteristics are driven by the mispricing of stocks. A large part of the sample has no stock price and, therefore, investors cannot under- or overvalue many of these assets based on their characteristics. Coupled with the likely scenario that managers of unlisted firms are less likely to be affected by investor sentiment, the results point to the conclusion that, first, the fundamentals risk factors are related to the risk and return characteristics of firms and, second, the risk and return characteristics of non-listed firms are similar to those of listed firms.

5.5. The cost of capital for listed and unlisted firms

We now examine whether the cost of capital, namely expected investment returns that are calculated from the four factor model, vary across the quintiles. Here we aim to answer the question of whether the cost of capital (expected investment returns) also varies between public and private firms. Therefore, in Table 7, we report the cost of capital for private and public industries across the portfolios formed on characteristics.

The second and third columns of Table 7 report average and expected investment returns for industries in quintile

Table 7
Expected and actual investment returns.
This table reports the average investment returns (AR) and the expected investment returns (ER) from the four-factor model based on Hou, Xue, and Zhang (2015a) for quintile groups based on the fraction of sales of public firms in the industry to total industry sales. For each quintile the actual returns and expected returns are reported for five portfolios sorted by the investment-to-capital ratio (*I*/*K*), the return on assets (*ROA*), size (measured as the size of the real capital stock) and idiosyncratic volatility (*IVOL*). The sample period is 1960–2009.

	Quintile 1		Quintile 2		Quintile 3		Quintile 4		Quintile 5		Q1–Q5
Portfolio	AR	ER	AR	ER	AR	ER	AR	ER	AR	ER	ER difference
Low *I*/*K*	0.235	0.241	0.221	0.228	0.161	0.143	0.158	0.107	0.147	0.131	0.110
2	0.123	0.103	0.117	0.095	0.113	0.109	0.103	0.065	0.155	0.133	−0.030
3	0.075	0.071	0.075	0.084	0.079	0.067	0.068	0.114	0.152	0.166	−0.095
4	0.048	0.024	0.048	0.018	0.028	0.066	0.032	0.065	0.127	0.157	−0.133
High *I*/*K*	−0.042	0.001	−0.036	0.002	−0.027	−0.012	−0.016	0.002	−0.002	0.005	−0.004
Low *ROA*	0.029	0.058	0.022	0.057	0.019	0.058	0.018	0.035	0.018	0.039	0.019
2	0.062	0.042	0.064	0.047	0.060	0.072	0.045	0.081	0.045	0.061	−0.019
3	0.080	0.075	0.076	0.063	0.067	0.074	0.070	0.079	0.048	0.069	0.006
4	0.089	0.077	0.093	0.087	0.083	0.079	0.093	0.034	0.081	0.094	−0.017
High *ROA*	0.176	0.186	0.169	0.171	0.125	0.091	0.126	0.076	0.384	0.323	−0.137
Small	0.162	0.146	0.156	0.147	0.097	0.124	0.121	0.097	0.100	0.131	0.015
2	0.093	0.099	0.085	0.076	0.086	0.087	0.053	0.076	0.058	0.042	0.057
3	0.067	0.077	0.070	0.079	0.071	0.072	0.079	0.095	0.086	0.031	0.046
4	0.066	0.070	0.067	0.077	0.057	0.059	0.055	0.044	0.090	0.068	0.002
Large	0.051	0.047	0.046	0.045	0.043	0.029	0.023	0.045	0.249	0.312	−0.265
Low *IVOL*	0.102	0.064	0.101	0.071	0.087	0.030	0.072	0.029	0.053	0.069	−0.005
2	0.036	0.063	0.039	0.054	0.049	0.069	0.065	0.058	0.058	0.045	0.018
3	0.069	0.071	0.071	0.075	0.078	0.063	0.057	0.069	0.051	0.061	0.010
4	0.077	0.070	0.074	0.067	0.048	0.078	0.041	0.068	0.106	0.087	−0.017
High *IVOL*	0.134	0.151	0.128	0.141	0.123	0.089	0.112	0.107	0.294	0.271	−0.120

1, which contains private firms. The remaining columns, excluding the final one, report the average and expected investment returns for the four remaining quintiles. Some clear patterns emerge in both actual and expected investment returns. For all the four characteristics, the portfolios have average investment returns and expected investment returns that match up well, consistent with the small pricing errors reported in the cross sectional tests. What is interesting is when we compare the expected returns between samples that have different proportions of public firms. For example, the final column reports the difference in expected return between quintile 1 and quintile 5, the closest we can get to comparing private and public firms. We find similar expected investment returns for all but a few of the extreme portfolios. This indicates the expected investment returns between portfolios that include more public firms are similar to those that include more private firms.

No systematic differences exist in the expected returns across the portfolios with a different fraction of public firms that would indicate a private firm effect in the cost of capital. For example, there is no private firm effect in the sense that all the expected investment returns of private firms are always higher (lower) than those of public firms. Any differences that are observed are likely to be a result of a difference in the value of a particular characteristic, for example, a higher (lower) *I*/*K* ratio, instead of being due to the firms being public or private.[12] What is interesting with the findings in Table 7 is that no systematic differences exist in the costs of equity capital between the portfolios that include more or fewer public and private firms. This is an important finding and provides new evidence that the private equity premium is similar to the public equity premium. Two interesting implications derive from this result. First, risk-adjusted estimates of the cost of capital for private firms, notoriously difficult to obtain, can be estimated from the investment returns of these firms. Second, because the cost of capital from the investment return approach is similar for public and private firms, given a characteristic, private firms can use public firms' stock returns that have a similar characteristic to proxy their cost of capital, especially if they do not have an extreme value of a particular characteristic.

The results that private and public firms have the same cost of capital can seem surprising given the lack of liquidity of private firms and the potential under-diversification of their owners. However, the findings are consistent with Moskowitz and Vissing-Jørgensen (2002) who use estimates of private firms' value and profits and study the returns to entrepreneurial investment. They find that, in spite of poor diversification, the returns to private equity are not systematically higher than the return to public equity.

5.6. Valuation of private and public firms

Belo, Xue, and Zhang (2013) present a new methodology for equity valuation, which is based on the *q*-theory of investment and arises from the perspective of managers' supply of capital and find strong empirical support for it.

12 Similar results are obtained when splitting industries into private and public using the number of employees and are available in the Internet Appendix.

Table 8
Tobin's *q*.
The table reports means, standard deviations, skewness, and the 5th, 25th, 50th, 75th, and 95th percentiles of Tobin's *q* for the entire sample of 459 manufacturing industries for which data are available at both the National Bureau of Economic Research Manufacturing Industry Productivity Database and Compustat, as well as for decile groups sorted by the fraction of listed firms in the industry to total industry sales. The column titled Private represents the two bottom deciles, and the top decile column is titled "Public". Tobin's *q* is given in Eq. (7). The p-Value in the last row is for the hypothesis that the mean *q* of the group in the column is the same as the Tobin's *q* for the private group and is computed by a bootstrap approach. The sample period is 1960–2009.

	All industries	Private	3	4	5	6	7	8	9	Public
$\bar{q}$	1.42	1.46	1.46	1.47	1.41	1.56	1.50	1.29	1.32	1.35
Std(*q*)	1.02	1.24	1.17	1.19	1.16	1.21	1.10	0.79	0.75	0.82
Skewness	4.27	5.60	5.15	5.07	3.79	3.93	4.08	2.26	2.03	1.76
5%	0.50	0.48	0.49	0.48	0.44	0.60	0.52	0.50	0.49	0.44
25%	0.83	0.84	0.83	0.84	0.77	0.91	0.87	0.79	0.82	0.80
50%	1.15	1.15	1.13	1.15	1.11	1.22	1.21	1.09	1.16	1.16
75%	1.68	1.74	1.71	1.71	1.61	1.79	1.79	1.54	1.60	1.68
95%	3.20	3.53	3.43	3.50	3.39	3.58	3.40	2.81	2.74	2.98
P(*diff*)	0.01	–	0.92	0.42	0.02	0.00	0.11	0.00	0.00	0.00

This approach is particularly suitable for the valuation of private firms, for which stock market value is not available.

We now examine several characteristics of Tobin's *q* of private and public industries, derived from the *q*-theory of investment. We first compare the valuation ratios of decile groups sorted by the fraction of sales of public firms in the industry to total industry sales. This comparison is interesting because private and public firms differ along many dimensions, each of which could affect firms' valuations. Subsequently, we examine higher moments of *q* as well as its cross sectional distribution.

The expression for Tobin's *q* is based on the firm's first order condition for optimal investment decisions, and is given in Eq. (7). Table 8 presents the results. The mean Tobin's *q* of the entire sample, shown in the first column, is 1.42. The other columns report in general that Tobin's *q* declines with the fraction of public firms in the industry, although the pattern is non-monotonic and is not very strong. The mean Tobin's *q* of private industries is 1.46, and public industries have a mean Tobin's *q* of 1.35, and the difference between the means is statistically significant with a p-Value of 0.00, as seen in the last row which presents the p-Values of the difference.

The following row in the table shows that the standard deviation of *q* declines with the percentage of public firms in the industry, as does the skewness as seen in the third row. Further information regarding the cross sectional distribution of *q* within each decile is given in the following rows. The 5th percentile of the distribution of *q* is approximately 0.5 throughout the 10 groups, and the 95th percentile is between 2.74 and 3.58. Overall, the percentiles are similar across the deciles.

Overall, Table 8 shows that the valuations of private and public industries are largely similar, as are the cross sectional distributions of *q*. To the best of our knowledge, ours is the first paper to examine private firms' valuations. Applying this valuation methodology to private firms is important in its own right and can be useful in investment decisions in private equity.

6. Conclusion

This paper examines the determinants of the cross-sectional variation in average investment returns for industry portfolios composed mainly of privately held firms and for industries consisting of mostly publicly listed firms. Investment returns are derived from the *q*-theory of investment (see Liu, Whited, and Zhang, 2009). We use the NBER Manufacturing Industry Productivity Database to calculate investment returns at the aggregate industry level, which includes both public and private firms. The NBER database contains detailed data on real capital stock, real investment, and sales for all 459 manufacturing industries from 1958 to 2009.

We find that characteristics that are important determinants of the cross section of stock returns, namely, the investment-to-capital ratio, return on assets, size, and idiosyncratic volatility, also describe the cross sectional variation of both public and private firms' investment returns. Given that private firms have no stock price and if the managers of private firms are less susceptible to investor sentiment and misvaluations, our results lend some support for a rational-based interpretation of the role of characteristics in the cross section of returns. Nevertheless, our results cannot rule out the possibility that investor misvaluation spills over to private firms by affecting their investment behavior, although we argue that this is unlikely to happen.

We also test the performance of the four factor model of Hou, Xue, and Zhang (2015a) using 20 characteristic-based single-sorted portfolios as test assets. The multifactor model performs well in describing the cross section of investment returns of private and public firms separately and together. This is a noteworthy finding, as it is the first test of an asset pricing model over all assets, including private firms. For a candidate risk factor to be a true risk factor, it must be an aggregate factor that affects all firms. We show that these four factors affect all firms and not only public firms.

The asset pricing tests have economically important implications for cost of capital calculations for private firms. The cost of capital for private firms is difficult to measure using risk based measures. This is because of the lack of stock prices for these firms. We show that investment returns can be used to calculate the cost of capital. Moreover, an alternative way to calculate the cost of capital is to use proxy firms from the public market and their stock returns. While this method has been used in the past, we show that it is a reliable benchmark. Finally, we show that the valuation ratios, namely, Tobin's q, and their cross-sectional variation are similar for public and private firms.

Appendix A

A1. Derivation of Eqs. (6) and (7)

Our derivation largely follows the proof of Proposition 1 in Liu, Whited, and Zhang (2009)

Let $\Pi(K_{it}, X_{it})$ denote the firm's profit function after costlessly adjustable factors of production have been optimized over, where K_{it} is the firm's stock of capital and X_{it} represents a vector of exogenous aggregate and firm-specific shocks. The capital stock's law of motion is $K_{it+1} = (1 - \delta_{it})K_{it} + I_{it}$, where δ_{it} is the capital's depreciation rate and I_{it} is investment. Investing incurs adjustment costs, and the adjustment cost function is assumed to be $C(I_{it}, K_{it}) = \frac{a}{2}(\frac{I_{it}}{K_{it}})^2 K_{it}$, where $a > 0$ is the adjustment cost parameter.

The firm's payout is $D_{it} = (1 - \tau_t)[\Pi(K_{it}, X_{it}) - C(I_{it}, K_{it})] - I_{it} + B_{it+1} - r^B_{it}B_{it} + \tau_t \delta_{it} K_{it} + \tau_t(r^B_{it} - 1)B_{it}$, where D_{it} is dividends, τ_t is the corporate tax rate, B_{it+1} is the amount of debt issued at the beginning of period t, and r^B_{it} is the gross corporate bond returns on B_{it}. $\tau_t \delta_{it} K_{it}$ is the depreciation tax shield and $\tau_t(r^B_{it} - 1)B_{it}$ is the interest tax shield.

The firm maximizes its cum-dividend market value of equity $V_{it} = \max_{\{I_{it+s}, K_{it+s}, B_{it+a}\}^{\infty}_{s=0}} E_t[\sum_{s=0}^{\infty} M_{t+s} D_{it+s}]$, where M_{t+1} is the stochastic discount factor from t to $t+1$. The firm's constraints are the law of motion of capital and a transversality condition that prevents firms from borrowing an infinite amount to distribute to shareholders: $\lim_{Tarrow\infty} E_t[M_{t+T}B_{it+T+1}] = 0$.

Let q_{it} be the Lagrangian multiplier associated with the law of motion of capital constraint. q_{it}, therefore, represents the marginal value of capital. From maximizing the firm's value with respect to I_{it} and K_{it+1} it follows that $q_{it} = 1 + (1 - \tau_t)\frac{\varrho C(I_{it}, K_{it})}{\varrho I_{it}} = 1 + (1 - \tau_t)a\frac{I_{it}}{K_{it}}$ [which is the same as Eq. (7)] and $q_{it} = E_t[M_{t+1}\{(1 - \tau_{t+1})[\frac{\varrho \Pi(K_{it}, X_{it})}{\varrho K_{it+1}} - \frac{\varrho C(I_{it}, K_{it})}{\varrho K_{it+1}}] + \tau_{t+1}\delta_{it+1} + (1 - \delta_{it+1})q_{it+1}\}]$. Dividing both sides of the first order condition with respect to K_{it+1} and substituting for q_{it} from the first order condition with respect to I_{it} yields $E_t[M_{t+1} r^I_{it+1}]$, where r^I_{it+1} is the same as in Eq. (6).

References

Asker, J., Farre-Mensa, J., Ljungkvist, A., 2015. Corporate investment and stock market listing: a puzzle? Review of Financial Studies 28, 342–390.

Barberis, N., Thaler, R., 2003. A survey of behavioural finance. In: Constantinides, G.M., Harris, M., Stulz, R.M. (Eds.), Handbook of the Economics of Finance. Elsevier/North Holland, Amsterdam, pp. 1053–1123.

Bartelsman, E.J., Gray, W., 1996. The NBER manufacturing productivity database. Technical working paper. National Bureau of Economic Research, Cambridge, MA.

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

Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77, 623–685.

Cochrane, J.H., 1991. Production-based asset pricing and the link between stock returns and economic fluctuations. Journal of Finance 46, 209–237.

Cochrane, J.H., 2005. Asset Pricing. Princeton University Press, Princeton, NJ.

Cooper, I., Priestley, R., 2011. Real investment and risk dynamics. Journal of Financial Economics 101, 185–205.

Cooper, M.J., Gulen, H., Schill, M.J., 2008. Asset growth and the cross section of stock returns. Journal of Finance 63, 1609–1651.

Daniel, K.D., Hirshleifer, D., Teoh, S.H., 2002. Investor psychology in capital markets: evidence and policy implications. Journal of Monetary Economics 49, 139–209.

Daniel, K.D., Titman, S., 1997. Evidence on the characteristics of cross-sectional variation in stock returns. Journal of Finance 52, 1–33.

Dechow, P.M., Khimich, N.V., Sloan, R.G., 2011. The accural anomaly. Unpublished working paper. University of California, Berkeley, CA.

Fama, E., MacBeth, J., 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81, 607–636.

Fraumeni, B.M., 1997. The measurement of depreciation in the U.S. national income and product accounts. Survey of Current Business 11, 7–23.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. Journal of Financial Economics 91, 24–37.

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

Hirschleifer, D., Hou, K., Teoh, S.H., 2012. The accrual anomaly: risk or mispricing? Management Science 58, 320–335.

Hou, K., Xue, C., Zhang, L., 2015a. Digesting anomalies: an investment approach. Review of Financial Studies 28, 650–705.

Hou, K., Xue, C., Zhang, L., 2015b. A comparison of new factor models. Unpublished working paper. Ohio State University, Columbus, OH.

Hribar, P., Quinn, P., 2013. Managers and investor sentiment. Unpublished working paper. University of Iowa, Iowa City, IA.

Jagannathan, R., Wang, Z., 1996. The conditional CAPM and the cross-section of expected returns. Journal of Finance 51, 3–53.

Lettau, M., Ludvigson, S., 2001. Resurrecting the (C)CAPM: a cross-sectional test when risk premia are time-varying. Journal of Political Economy 109, 1238–1287.

Lin, X., Zhang, L., 2013. The investment manifesto. Journal of Monetary Economics 60, 351–366.

Liu, L.X., Whited, T.M., Zhang, L., 2009. Investment-based expected stock returns. Journal of Political Economy 117, 1105–1139.

Lucas Jr., R.E., 1978. Asset prices in an exchange economy. Econometrica 46, 1429–1445.

Moskowitz, J.T., Vissing-Jø rgensen, A., 2002. The returns to entrepreneurial investment: a private equity premium puzzle? American Economic Review 92, 745–778.

Polk, C., Sapienza, P., 2009. The stock market and corporate investment: a test of catering theory. Review of Financial Studies 22, 187–217.

Richardson, S., Tuna, I., Wysocki, P., 2010. Accounting anomalies and fundamental analysis: a review of recent research advances. Journal of Accounting and Economics 50, 410–454.

Stein, C.J., 1996. Rational capital budgeting in an irrational world. Journal of Business 69, 429–455.

Titman, S., Wei, K.C., Wei, J., Xie, F., 2004. Capital investment and stock returns. Journal of Financial and Quantitative Analysis 39, 677–700.

Xing, Y., 2008. Interpreting the value effect through the q-theory: an empirical investigation. Review of Financial Studies 21, 1767–1795.