Why are Most Firms Privately Held? [*] [†]

Joan Farre-Mensa
Harvard Business School

First draft: November 2010. This draft: March 2011

[*] I am indebted to my advisor, Alexander Ljungqvist, for his guidance and support throughout this project. Special thanks go to my committee members, John Asker, Douglas Gale, and Boyan Jovanovic, for invaluable comments and suggestions. I would also like to thank Viral Acharya, Yakov Amihud, Stephane Bonhomme, Alex Chinco, Rustom Irani, Andrew Karolyi, Gordon Phillips, Jacob Sagi, Mike Weisbach, participants at the FMA Doctoral Student Consortium and the NYU Financial Economics Workshop as well as seminar audiences at NYU Stern, Hofstra, Virginia (Darden), Georgia Tech, Rochester, HBS, Cornell, Ohio State, UNSW, Federal Reserve Board, UC Davis, IESE, CEMFI, Toulouse, and Warwick. I am grateful to Sageworks Inc. for access to their database on private companies.

[†] Address for correspondence: Harvard Business School, Rock Center 218, Soldiers Field, Boston MA, 02163. Phone 617-495-6963, e-mail jfarremensa@hbs.edu.

Why are Most Firms Privately Held?

Abstract

Even among large U.S. firms, most choose to remain private rather than listing on a stock market. I show that an important reason for this choice is public firms' inability to disclose information selectively. This leads to a 'two-audiences' problem: Disclosure reduces information asymmetries among investors but also potentially benefits product-market competitors. Being public involves a trade-off between this disclosure cost and the benefit of a lower cost of capital due to the greater liquidity with which shares can be traded on a stock market. Using a rich new dataset on private U.S. firms, I show that firms in industries with high disclosure costs and high information asymmetry are more likely to remain private while firms in industries that require a large scale to operate efficiently are more likely to be public. I then establish a new stylized fact: Public firms hold more cash than private firms, particularly when they operate in industries in which disclosing information to competitors is most costly. This fact is robust to several ways of addressing the endogeneity of the going-public decision, including matching, exploiting within-firm variation, and instrumental variables. Consistent with my model, I find that public firms hoard cash in order to mitigate the disclosure costs associated with raising capital.

Key words: Private companies; IPOs; Cash; Liquidity management; Disclosure costs; Regulation Fair Disclosure.

JEL classification: D21; D22; G32; G38; L26.

A stock market listing provides access to a deep pool of low cost capital, and yet over 70% of U.S. firms with more than 500 employees decide to remain private rather than listing on a stock market, with private firms accounting for over 60% of U.S. production.[1] Interestingly, the data show that there is substantial variation in the share of private firms across industries. My goal in this paper is to study why so many firms choose to stay private, how their choice is affected by the characteristics of the industry where they operate, and how this choice affects their behavior – in particular, my analysis will focus on their liquidity management. The emergence of a rich new dataset on private U.S. firms allows me to examine these questions both theoretically and empirically.

I study the trade-off between taking a firm public and keeping it private, focusing on two key differences between stock-market listed (or 'public') firms and unlisted (or 'private') firms. First, the ability of public firms' shareholders to trade their shares on a stock market means that their savings are relatively liquid. This allows public firms to raise capital at a lower cost than private firms, all else equal (Pagano, Panetta, and Zingales (1998)). Second, public firms are subject to disclosure requirements that do not apply to private firms, and which go well beyond the quarterly release of their financial statements. In particular, the Securities and Exchange Commission (SEC) Regulation "Fair Disclosure" prohibits public firms from selectively disclosing material information to outside shareholders. Thus, information must be disclosed to all market participants or not at all. This implies that public firms face a 'two-audiences' problem when communicating with investors: More information reduces investors' information asymmetry about the quality of the firm and so reduces the firm's cost of capital (see Botosan (1997) or Easley and O'Hara (2004)), but public disclosure is costly because the information may be exploited by product-market competitors.

There is ample evidence that the disclosure costs faced by public firms are substantial and that firms see them as one of the main costs of being public. A case in point is Google's 2004 IPO filing which contains the argument that "as a smaller private company, Google kept business information closely held, and we believe this helped us against competitors."[2] Surveying CFOs, Brau and Fawcett (2006) report

[1] Data on the number and sales of firms in the U.S. economy is available at http://www.census.gov/econ/susb/. The CRSP-Compustat database is the source for the number of stock market-listed firms.
[2] See Google's S1 filing, available at http://www.sec.gov/Archives/edgar/data/1288776/000119312504073639/ds1.htm.

that “disclosing information to competitors” and “SEC reporting requirements” are among the five most important reasons why private firms remain private in the U.S.

In order to analyze the trade-off between the liquidity benefit and the two-audiences problem associated with being public, I model the problem of an entrepreneur who needs to raise equity to develop a project. The entrepreneur can raise equity privately or take the firm public and raise capital through an IPO. The project later requires an additional investment whose size only becomes known once the project is already in place. The firm can fund this additional investment from its precautionary cash, in case it set some aside at the outset of the project, or by issuing additional equity in an asymmetric information environment. For unlisted firms, this information asymmetry is costless because they can selectively convey their value to private investors. On the other hand, information asymmetry makes it costly for public firms to raise additional capital because they face a ‘two-audiences’ problem. Indeed, public firms have to choose between issuing equity through a seasoned equity offering (SEO), which involves substantial public disclosure that eliminates the information asymmetry but can benefit competitors; or through a private investment in public equity (PIPE), which requires no information disclosure at the cost of selling the shares at a low (pooling) price.

If the additional investment is funded, the project yields a deterministic return which is paid as a taxed dividend to shareholders, together with any unused precautionary cash. However, before the dividend is paid, shareholders may be hit by a liquidity shock that forces them to sell their shares. I capture the liquidity benefit of being public by assuming that, while public-firm shareholders hit by the liquidity shock can sell their shares at no cost, private-firm shareholders incur a convex trading cost.

My model provides two sets of empirical predictions. The first examines how the characteristics of a firm’s industry affect the trade-off between going public and staying private. Specifically, my model predicts that the probability of a firm being public increases in the minimum efficient scale required to operate in its industry and decreases in the industry’s disclosure costs and degree of information asymmetry. I test these predictions using standard data on public firms (e.g., Compustat) and a novel database on private U.S. firms created by Sageworks Inc.[3] The (firm level) data are consistent with the predictions of my model. In addition, I find support for these same predictions at the industry level. Using data from the U.S. Census, I find a larger share of public firms in those industries where my model predicts that an entrepreneur's return to being public is higher.

My second prediction states that it is optimal for public firms (but not for private firms) to hoard cash. By hoarding cash, public firms avoid having to raise capital every time they need to make an additional investment. Raising additional capital is costly for public firms, because their inability to communicate selectively with investors implies that they are subject to a 'two-audiences' problem every time they raise capital. The Sageworks data allow me to compare the cash policies of public and private firms, and I indeed find that public firms hold more cash than private firms. The difference is large: public firms hold almost twice as much cash (as a percentage of assets) as private firms. This is a new stylized fact, and it is not driven by the 2007-2009 financial crisis as it is robust to ending my panel in 2006. Furthermore, I find that the cash difference between public and private firms increases in industry disclosure costs, the industry's return on capital, and cash flow volatility. My model shows that these three interaction effects are consistent with my novel hypothesis that public firms hoard cash in order to minimize the disclosure costs associated with raising capital.[4]

When comparing the cash policies of public and private firms, I face an important identification challenge: as highlighted by my model, firms choose whether to go public or remain private. In my baseline empirical analysis, I explicitly control for the industry characteristics that my model predicts affect both the going-public decision and a firm's cash policy. However, my model abstracts away from firm idiosyncratic factors likely to affect an entrepreneur's decision to go public, such as acquisition opportunities or the entrepreneur's desire to maintain decision-making control (Brau and Fawcett (2006)). If these idiosyncratic factors are also correlated with firms' cash policies, my analysis will be subject to

[3] Sageworks obtains the data from a large number of accounting firms which input data for all their (private) corporate clients. My version of the database consists of a seven year panel starting in 2002 consisting of over 270,000 private firm-years. Throughout, I exclude from my sample small private firms which have no realistic choice of being public.

[4] Several authors have analyzed the 'two-audiences' problem faced by public firms. Campbell (1979) and Bhattacharya and Ritter (1983) are seminal references, while more recent studies can be found in Bhattacharya and Chiesa (1995), Yosha (1995), Maksimovic and Pichler (2001), and Spiegel and Tookes (2009). However, mine is the first paper to highlight that disclosure costs are a first-order determinant of the cash policy of public firms.

endogeneity.

I follow three different approaches to address this potential endogeneity problem. The first tests whether there are other *observable* firm and industry characteristics (besides those predicted by my model) that affect both a firm's listing decision and its cash policy. Matching methods are an effective way to control for this selection on observables, and my results are robust to using both propensity score matching and to matching on industry and size. Of course, it is impossible to match on *unobservables*, so I also use a second test that exploits within-firm variation in listing status and thus allows me to difference away time invariant firm-level unobservables. Specifically, I analyze changes in cash holdings among firms that go public for the sole purpose of allowing existing shareholders to cash out.[5] This restriction to non-capital raising IPOs ensures that the IPO proceeds do not mechanically increase the firms' cash holdings. Nonetheless, I find that firms hold more cash after going public than they did before.

The within-firm test relies on a small sample of 98 IPOs, as non-capital raising IPOs are relatively rare. In addition, the test may be biased if the IPO decision is triggered by a change in unobservables that also lead to a change in cash policy (e.g., firms that go public to undertake acquisitions). To address these concerns, my third test uses the firm's state of location and its industry's minimum efficient scale to instrument a firm's listing status, using the combined Compustat-Sageworks dataset. The use of geographical location as an instrument is motivated by the findings of Loughran (2008) who provides direct evidence on the impact of location on a firm's ability to issue equity; the use of minimum efficient scale is justified by the prediction of my model that scale is positively related to the probability of being public but otherwise unrelated to firms' cash policies (after controlling for firm size and other standard controls in the cash literature).

The instrumental variables (IV) test suggests that endogeneity leads me to *under*estimate the level difference in cash holdings between public and private firms. This is consistent with the notion that the owner-managers of the (large) private firms that are more likely to remain private are those with a strong desire for control, as suggested by the findings in Brau and Fawcett (2006). Intuitively, an owner-manager

[5] This test is not possible in the Sageworks data because all firm data are held anonymously. Thus, I use Thomson Financial's SDC database to identify the IPO sample used in this particular test, which spans from 1980 to 2007.

with a strong desire for control will prefer to hold larger cash holdings than a manager with an average desire for control, because cash avoids the need to have frequent negotiations with banks or private shareholders. On the other hand, the IV results show that endogeneity has little, if any, effect on the estimates of the interaction between being public and industry disclosure costs, returns on capital, and cash flow volatility.

The notion that endogeneity is not driving the observed tendency of public firms to hoard cash is reinforced by my comparison of public and private firms' cash flow sensitivity of cash. Almeida, Campello, and Weisbach (2004) argue that financially constrained firms have a higher cash flow sensitivity of cash. Exploiting within-firm variation, I find that private firms exhibit a lower cash flow sensitivity of cash than public firms of similar size. This suggests that private firms behave as if they are less financially constrained than comparable public firms, which is consistent with disclosure costs being an important financing constraint for public firms.

My analysis focuses on the trade-off between the disclosure cost and the liquidity benefit of being public. Yet this is not to say that these are the only differences between public and private firms. Another salient difference is agency problems (Jensen and Meckling (1976)). These are common in public firms, given that a stock market listing requires at least some separation of ownership and control. Private firms, on the other hand, are subject to little, if any, separation of ownership and control (Asker, Farre-Mensa, and Ljungqvist (2010a)) and so have fewer agency problems. Agency considerations are, however, unlikely to explain why public firms hold more cash than private firms. Harford, Mansi, and Maxwell (2008) show that, among public firms, those with weaker corporate governance hold less cash than those with stronger governance. This suggests that the difference in cash holdings between public and private firms would actually be larger, on average, if public firms were free of agency problems.

My paper makes two main contributions. First, it contributes to our understanding of the trade-off between public and private ownership (see Pagano, Panetta, and Zingales (1998)) by showing how this trade-off depends on the characteristics of the firm's industry. Several recent papers have analyzed the relation between product-market competition and the decision to go public (see Chod and Lyandres (2010), Chemmanur and He (2009), Chemmanur, He, and Nandy (2010), and Hsu, Reed, and Rocholl

(2010)). However, most prior literature has examined this question at the firm level, while the distribution of public and private firms across industries in the U.S. economy remains largely unexplained. Chod and Lyandres (2010) and my paper are the first to follow a macro approach and empirically analyze how the share of public firms in an industry is related to the industry's characteristics.[6]

Second, mine is the first paper to analyze differences in cash policies between public and private firms, allowing me to establish a new stylized fact: Public firms hold substantially more cash than private firms.[7] The relevance of this new result is underscored by the fact that extant theory and empirical evidence do not offer a clear prediction of the difference in cash holdings between public and private firms. In particular, the literature has shown that, among public firms, those with greater access to capital markets and higher credit ratings tend to hold less cash (see, e.g., Opler, Pinkowitz, Stulz, and Williamson (1999), Acharya, Davydenko, and Strebulaev (2008)). In light of this finding, one might have expected public firms to actually hold less cash than private firms. My results suggest that this extrapolation would be misleading as public and private firms fundamentally differ in their disclosure requirements. Indeed, my paper shows that the inability of public firms to selectively communicate with outside investors plays a key role in explaining their propensity to hoard cash.

1. Modeling the Public vs. Private Trade-off

The purpose of my model is to analyze how the value of being public or private depends on the characteristics of the industry in which a firm operates. To do so, I study how a firm's public or private status affects its precautionary cash and financing policies. The predictions of the model will then guide my empirical analysis of the trade-off between going public and staying public, the distribution of public firms across industries in the U.S. economy, and the cash policies of public and private firms.

1.1. Model set-up

The model has four periods. The timeline is as follows: In period 0, a privately held firm needs to raise k units of capital to invest in a project; k can be interpreted as the minimum efficient scale required to operate in the industry. In period 1, the project requires an additional investment of random size. In

[6] Chod and Lyandres' (2010) analysis highlights that being public allows firms to better diversify idiosyncratic risk, and thus public-firm managers can adopt more aggressive product-market strategies than private firms.
[7] This finding is confirmed by Gao, Harford, and Li (2010) in a contemporaneous paper using a more limited sample of 3,836 private firms.

period 2, the firm's shareholders may be hit by a liquidity shock, forcing them to sell their shares of the firm. Finally, in period 3, the project produces a deterministic cash flow equal to zk, which is paid out as a (taxed) dividend to shareholders, and the firm is liquidated.

In period 0, the firm can raise capital in excess of k and carry it as precautionary cash, C.[8] There is a tax penalty for carrying precautionary cash that ends up not being used, because cash is taxed at rate τ when it is returned to shareholders as a dividend in period 3 (therefore, τ induces an opportunity cost of holding cash). Without loss of generality, I assume that neither the firm nor the entrepreneur have any initial capital; I also assume no time discounting. I abstract from capital structure considerations and assume that the firm can only raise capital by issuing equity, $E_0 = k + C$.[9] Capital markets are competitive: equity investors, assumed to be risk-neutral, supply capital at a competitive price and so make zero profits in expectation.

The entrepreneur can raise E_0 in two ways: she can either raise equity privately or she can take the firm public and raise capital through an IPO. I next describe the two fundamental differences between public and private firms that are at the heart of my model's trade-off: public firms are subject to disclosure costs, while private firms face a higher cost of capital, ceteris paribus, due to the relative illiquidity of their shares.

1.1.1. The disclosure cost of being public

The first fundamental difference between public and private firms is that public firms cannot communicate selectively with outside shareholders. The reason is that public firms have to comply with SEC Regulation Fair Disclosure (Regulation FD) which bans selective disclosure of material non-public information. Private firms are not subject to Regulation FD and thus face no restrictions on their communications with investors.

The inability of listed firms to communicate privately and selectively with investors affects their

[8] In the model, 'precautionary cash' refers to cash that is not needed to settle a firm's usual transactions (e.g., pay wages).
[9] In an extension of the model available upon request, I allow firms to issue a limited amount of non-risky debt. In this extension, debt capacity and precautionary cash are perfect substitutes. The implications of the model remain the same, with public firms holding more cash and debt capacity (i.e., less debt) than private firms.

ability to raise funding once the project has been initiated.[10] To capture the differential disclosure cost, and the resulting value of holding precautionary cash, I assume that all firms, whether public or private, suffer a stochastic cost overrun of size ρk in period 1. If the cost overrun is not funded, the project cannot be completed and the period 3 cash flow is 0. The firm does not learn the size of ρ until period 1, at which time it is observed by all market participants. The distribution of ρ is continuous on $(0,\infty)$, with density $f(\rho)$ that is common knowledge and non-increasing. If the magnitude of ρ is such that it is optimal for the firm to abandon the project, the firm can be liquidated in period 1.[11] After period 1, the firm requires no additional capital investment.

A firm can fund the cost overrun ρk in two ways: it can use its precautionary cash *C*, if it has any, or it can issue new equity. Given that dividends are taxed, it is always optimal for the firm to exhaust its cash holdings before issuing new equity in period 1.

If *C* is not sufficient to cover ρk, the firm needs to issue equity in period 1, $E_1 = \rho k - C$, to keep the project alive. However, period 1 investors face an information asymmetry. Unless the firm discloses some proprietary information about its project,[12] investors do not know whether it is optimal to continue with the project or to abandon it. To capture this information problem, I assume that investors cannot distinguish a viable firm from failed firms whose project is no longer viable. In order to streamline the presentation, the failed firms are not explicitly modeled here. In an appendix available on request, I solve an extended version of the model that includes failed firms. All results presented here are consistent with those obtained in the extended version.

If the firm has remained private, it can privately communicate to investors all information needed to separate itself from failed firms, and thus such separation is costless for private firms. However, if a public firm wants investors to be able to accurately value the firm, it has to raise equity through a seasoned equity offering (SEO), which requires substantial public disclosure. I capture the cost of publicly disclosing proprietary information, and in particular making it available to the firm's

[10] I assume that firms cannot change their listing status after period 0.
[11] I assume that the liquidation process allows shareholders to recover any precautionary cash carried by the firm.
[12] Campbell (1979) emphasizes that such proprietary information can be of two types: technological and strategic.

competitors, by assuming that the revenue of the firm's project decreases to $(1-\gamma)zk$ in the event of an SEO. This assumption can be motivated by the fact that the project's revenue depends not only on the firm developing a new product or strategy, but also on its competitors not developing a similar, competing product or strategy.[13] Thus, γ captures disclosure costs, i.e., the extent to which the disclosure of proprietary information about the project affects its profitability.

Alternatively, a public firm can sidestep an SEO and raise equity E_1 without disclosing proprietary information. In practice, this can be done through a so called PIPE (private investment in public equity).[14] In this case, investors cannot distinguish the firm from failed public firms. Hence, the firm faces a pooling equilibrium price that reflects investors' uncertainty regarding the project's viability. This price depends on the relative proportion of observably identical firms with and without viable projects in the industry. In particular, I assume that there is a fraction μ of failed public firms ready to take advantage of a pooling equilibrium if firms with viable projects raising E_1 choose to raise capital through a PIPE. The parameter μ is independent of E_1, and it is common knowledge.

The degree of information asymmetry between investors and public firms in a given industry can be interpreted as a key determinant of μ. In industries with more severe asymmetry, there will be proportionally more failed firms that are observably identical to viable firms in period 1, and thus μ will be higher.

1.1.2. The liquidity cost of being private

The second fundamental difference between public and private firms in my model is that the shares of private firms are illiquid and thus costly to trade. To capture this cost, I assume that a firm's shareholders may be hit by a liquidity shock in period 2 that prompts them to sell their shares. In particular, similar to Gorton and Pennacchi (1990), I assume that the utility function of a firm's shareholders (i.e., equity investors and the entrepreneur) over consumption e_t is as follows:

---

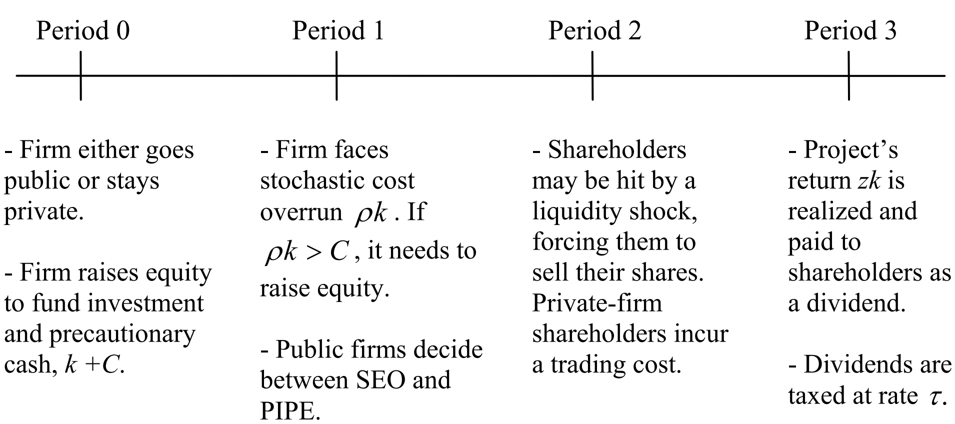
[13] For simplicity, my model does not consider any potential beneficial effects that disclosure may have on competitors. Also, an analysis of the welfare effects of disclosure is beyond the scope of the present model.

[14] To ensure compliance with Regulation FD, PIPE investors do not have access to material nonpublic information. PIPE investors receive a private placement memorandum containing the issuer's SEC filings, which are already public. Investors do not receive projections or other non-public information (Pinedo and Tanenbaum (2009)).

$$U(e_2, e_3) = \begin{cases} e_2, & \text{with probability } \lambda \\ e_2 + e_3, & \text{with probability } 1-\lambda \end{cases}$$

With probability $1-\lambda$, shareholders postpone their consumption until the firm pays a dividend in period 3. However, with probability λ, shareholders are hit by a liquidity shock that induces them to consume in period 2. This liquidity shock is immaterial for public-firm shareholders, since in period 2 the viability of the project is costlessly revealed in the firm's regulatory filings.[15] This allows public-firm shareholders to sell their shares at a price that fully reflects the firm's period 3 (after-tax) dividend. Yet the liquidity shock is costly for private-firm shareholders: since the firm's shares are not publicly traded, shareholders face a non-zero trading cost when selling their shares in period 2.[16] This trading cost can be interpreted as a search cost induced by the fact that shareholders cannot place their sell order through an exchange. I allow for the possibility that the marginal trading cost increases in the size of the trade (see Silber (1991) for evidence) and assume that private-firm shareholders face a positive cost $(\delta_1 + \delta_2 X)X$, with $\delta_2 > 0$, when selling shares of total value *X*.

Figure 1. Timeline of the model.



1.2. Solving the model: Optimal financing and precautionary cash policy

I solve the model using the Perfect Bayesian Equilibrium solution concept. A Perfect Bayesian Equilibrium consists of a strategy profile and a belief system such that the strategies of firms and

[15] More precisely, if the firm carried out an SEO in period 1, the firm's viability was revealed then; otherwise, it is costlessly revealed in the firm's regulatory filings at the beginning of period 2.
[16] Unused precautionary cash *C* can be paid out as a dividend in period 2 to minimize shareholders' trading costs.

investors are sequentially rational, given the belief system, and the belief system is consistent with Bayes' rule, given the strategy profile.

In both periods 0 and 1, investors finance a firm if and only if, given their beliefs regarding the firm's type[17] and the price of the shares the firm is issuing, they expect to obtain a non-negative return on their investment. To describe firms' equilibrium strategies, I distinguish the case of a firm that goes public in period 0 from that of a firm that stays private.[18] As is common in this set-up, I solve their problems backwards. In periods 3 and 2, a firm's problem is trivial, regardless of whether it is public or private: In period 3, the firm collects the cash flow from the project and pays it out as a dividend; in period 2, shareholders hit by a liquidity shock sell their shares of the firm, incurring a trading cost only if the firm is private. This leaves the period 1 and 0 strategies to be solved.

1.2.1. Period 1 optimal financing policy

Let $c = C/k$ denote a firm's precautionary cash holdings scaled by project size *k*. In period 1, a firm facing a cost overrun ρk that exceeds its precautionary cash *ck* has to issue equity to fund $(\rho - c)k$. If the firm is public, it has to choose between an SEO, disclosing the strategic information needed to convey to investors that the firm has a viable project, and a PIPE. The following lemma characterizes the financing policy that maximizes the value of the firm for its (current) shareholders:[19]

Lemma 1. *The optimal financing policy of a public firm that needs to raise* $(\rho - c)k$ *in period 1 is as follows:*

a) If $0 < \rho - c < \frac{1-\mu}{\mu}\gamma(1-\tau)z \doteq S$*, it is optimal to raise capital through a PIPE.*

b) If $S \leq \rho - c < (1-\tau)(1-\gamma)z - c \doteq L^{Pub} - c$*, it is optimal to raise capital through an SEO.*

c) If $L^{Pub} \leq \rho$*, it is optimal to liquidate the firm and return its precautionary cash ck to its shareholders.*

Figure 2 helps illustrate the intuition. When a public firm needs to raise a small amount of capital in

[17] In period 1, investors' beliefs are based on the amount of capital the firm raised in period 0, the amount raised in period 1, and the size of the cost overrun ρ. In period 0, they depend on the amount raised in period 0.
[18] As mentioned earlier, failed firms are not included in my exposition here.
[19] Proofs and technical derivations can be found in Appendix A.

period 1, a PIPE is preferable to incurring the disclosure cost of an SEO (i.e., the decrease in the project's profitability). However, PIPE investors need to be compensated for the possibility of investing in a failed firm. This implies a higher cost of capital in a PIPE than in an SEO. When the amount of capital needed exceeds the threshold Sk, the optimal strategy is to incur the disclosure cost and avoid being pooled with failed firms.[20] From the expression $S = \left((1-\mu)/\mu\right)\gamma(1-\tau)z$, it follows that the SEO threshold Sk increases in the disclosure cost γ and decreases in the information asymmetry μ, which determines the additional (pooling) cost of a PIPE relative to an SEO.

These predictions are consistent with the evidence in Gomes and Phillips (2007), who find that SEOs tend to be larger in size than PIPEs: from 1995 to 2003, the mean and median amount raised in SEOs as a percentage of firm value were 22% and 15%, respectively; in the case of PIPEs, the mean and median were 15% and 9%, respectively. In addition, Ali, Klasa, and Yeung (2010) find that firms in industries where disclosure costs are high prefer PIPEs over SEOs, which is consistent with the prediction that the SEO threshold Sk is increasing in the cost of disclosure.

In equilibrium, investors refuse to provide PIPE financing to a firm raising more than Sk units of capital, given that they believe that such a firm is a failed firm with probability one. Therefore, if a public firm needing more than Sk units of capital chose to carry out a PIPE, it would not be funded. As a result, attempting a PIPE would force such a firm into liquidation, with its shareholders only receiving the precautionary cash C carried by the firm, as shown in Figure 2.

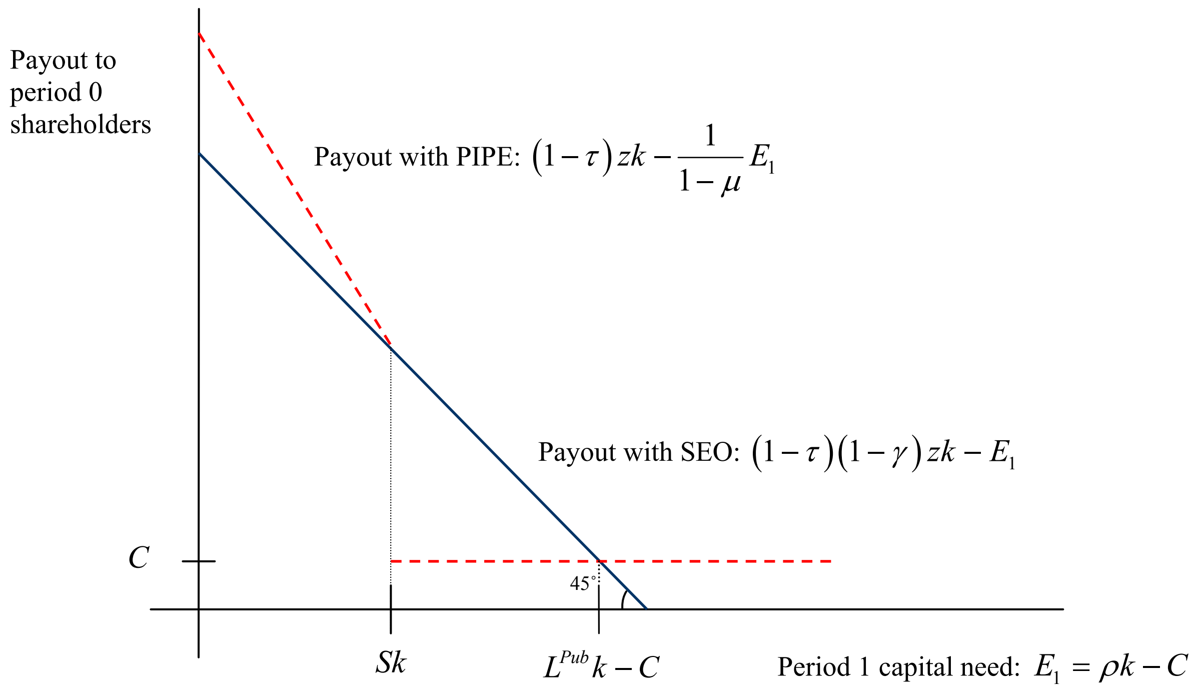
The figure also shows that when a public firm's cost overrun ρ is large, liquidating the firm and returning C to shareholders is actually an optimal strategy; specifically, liquidation is optimal when ρ exceeds L^{Pub}.

A private firm in need of $(\rho - c)k$ units of capital will raise equity privately, thus selectively disclosing to investors the information required to separate itself from failed private firms. This implies that, in equilibrium, failed private firms are never funded. Analogously as in the case of public firms,

[20] If $\gamma \geq \mu$, it is never optimal for a firm to carry out an SEO. Thus, I focus on the case $\gamma < \mu$ so that an SEO is a possible funding option. This ensures that PIPE investors are willing to fund the firm since $(\rho - c)k < (1-\mu)(1-\tau)zk$.

$L^{Pri} \doteq \left(1-\lambda\delta_1-\lambda\delta_2\left(1-\tau\right)zk\right)\left(1-\tau\right)z$ is such that if a private firm's cost overrun satisfies $\rho \geq L^{Pri}$, then it is optimal to liquidate the firm in period 1.

Figure 2. Financing options for a public firm in period 1.



1.2.2. Period 0 optimal precautionary cash policy

In period 0, a firm chooses its precautionary cash *c* to maximize the value of the firm to the entrepreneur. The following proposition characterizes the optimal precautionary cash policy depending on whether the firm is public or private.

Proposition 1.

It is optimal for a private firm to carry no precautionary cash. A public firm's optimal ratio of precautionary cash to project size $\hat{c}$ *is implicitly defined by the following equation:*

$$\tau\int_0^{\hat{c}} f\left(\rho\right)d\rho = \frac{\mu}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} f\left(\rho\right)d\rho \quad (1)$$

The intuition is as follows.[21] For both public and private firms, it is costly to hold precautionary cash.

[21] I focus on parameter values for which an SEO is the optimal strategy for at least some values of the cost overrun ρ, that is, $\hat{c}+S<L^{Pub}$.

If the cost overrun is small and the cash is not used, a tax is incurred when the cash is returned to shareholders as a dividend. Yet for private firms, precautionary cash offers no offsetting benefit: Private firms can selectively communicate with investors when they need to raise capital in period 1, so they are able to issue equity at a price that reflects the firms' value without disclosing proprietary information to competitors.[22] As a result, private firms optimally carry no precautionary cash. (Of course, private firms will still need to hold some cash to pay wages, suppliers, etc. In the model, I abstract away from this transaction motive to hold cash.)

For public firms, on the other hand, it is costly to raise equity in period 1. Therefore, public firms choose their precautionary cash such that its expected marginal cost equals its expected marginal benefit.[23] The marginal cost is the tax cost incurred when the cash is not used, as captured by the left-hand side of equation (1). The marginal benefit, captured by the right-hand side of equation (1), is avoiding the need to raise capital in period 1 via a PIPE at a pooling price that undervalues the firm. Equation (1) implies that $\hat{c} > 0$, i.e., it is optimal for public firms to carry some precautionary cash.

The cost of disclosing strategic information to competitors, γ, appears in equation (1) only indirectly through its effect on *Sk* (the minimum capital need for which it is optimal to carry out an SEO). This reflects the fact that the cost of an SEO is independent of the amount of equity that the firm needs to raise: Disclosure benefits competitors whether the firm raises a small amount or a large amount so regardless of the size of the SEO, the firm loses revenue γzk. Thus, there is no marginal benefit to carrying precautionary cash once the firm has decided that it is optimal to carry out an SEO, as captured by equation (1).

In order to analyze how a public firm's optimal precautionary cash holdings depend on the characteristics of the industry in which it operates, I apply the implicit function theorem to equation (1) (henceforth, cash always refers to cash scaled by size). Proposition 2 summarizes the results.

---

[22] The empirical evidence is consistent with the notion that it is common for private firms to raise equity. While I cannot test this implication using the Sageworks data because they do not include the statement of cash flows, one of the questions in the National Survey of Small Business Finances (NSSBF) speaks directly to it: "During the last 12 months, did [the firm] obtain any new equity investment from new or existing shareholders, excluding retained earnings?" In the latest NSSBF (2003), 8.3% of the largest private firms in the survey (20-499 employees) responded affirmatively.

[23] In the extended version of the model that explicitly includes failed firms, I show that (non-failed) public firms have no incentive to alter their precautionary cash policy to try to separate themselves from failed firms.

Proposition 2. *A public firm's optimal precautionary cash holdings are:*

a) increasing in the cost of revealing strategic information to competitors, γ *;*

b) increasing in the rate of return on the project, z.

With additional structure on the density of the cost overrun, the following comparative static can be added to Proposition 2:

c) Assuming that the cost overrun ρ *follows a normal distribution* $N\left(0,\sigma^2\right)$ *truncated at 0,*[24] *a public firm's optimal precautionary cash holdings are increasing in the volatility of the cost overrun, determined by* σ^2 *.*

The effect of the degree of information asymmetry in the industry, μ , on a public firm's optimal precautionary cash holdings is ambiguous. On the one hand, higher asymmetry increases the benefit of holding cash by increasing the PIPE cost $\mu/\left(1-\mu\right)$ in equation (1); on the other hand, higher asymmetry decreases the SEO threshold *Sk* and, as argued earlier, there is no marginal benefit to carrying precautionary cash once the firm has decided that it is optimal to carry out an SEO. Which effect dominates depends on the particular parameter values.

Also, note that it follows from equation (1) that the minimum efficient scale in an industry, *k*, has no effect on public firms' optimal cash holdings, $\hat{c}$.

1.3. Comparing the entrepreneur's returns to going public and staying private

Let r^{Public} and $r^{Private}$ denote the project's net return on investment of a public and private firm, respectively. This return is captured fully by the entrepreneur since outside investors make zero profits in expectation. The following proposition examines how the entrepreneur's return to going public depends on the characteristics of the industry in which the firm operates:

Proposition 3. *The difference between the entrepreneur's returns if she takes her firm public and if the firm stays private,* $r^{Public}-r^{Private}$*, is:*

---

[24] That is, $f\left(\rho\right)=\sqrt{2/\left(\pi\sigma^2\right)}\,e^{-\rho^2/\left(2\sigma^2\right)}$ if $\rho>0$ and $f\left(\rho\right)=0$ otherwise. Therefore, $E\left[\rho\right]=\sqrt{\left(2\sigma^2\right)/\pi}$ and $\mathrm{Var}\left[\rho\right]=\sigma^2\left(1-2/\pi\right)$. To separate the effect of the volatility of a firm's cost overrun from the expected value of its project, I assume that the return of the project is $z=\tilde{z}+E\left[\rho\right]$.

a) decreasing in the cost of revealing proprietary information to competitors, γ *;*

b) decreasing in the degree of information asymmetry between investors and public firms in the industry, μ *;*

c) increasing in the assets required by the firm to operate, k.

The intuition is as follows. The advantage of being private is that the firm can selectively communicate with investors in period 1. The greater the disclosure costs and information asymmetry in its industry, the more valuable is selective communication. The advantage of being public is that, if the firm's shareholders (equity investors and the entrepreneur) need to sell their shares in period 2, they do not have to pay the convex trading cost.

1.4. Empirical predictions

My first set of empirical predictions relate the choice between taking a firm public and keeping it private to the characteristics of its industry. If my model were literally true, all firms in an industry would either be public or private – there would be no industry with both public and private firms. What my model leaves out are the various idiosyncratic factors that influence an entrepreneur's decision to go public, such as her risk aversion (Chod and Lyandres (2010)), location (Loughran (2008)), or desire to maintain control (Brau and Fawcett (2006)). Let ε_i denote the net value of the omitted idiosyncratic factors. The distribution of ε is assumed independent of the industry characteristics that determine $r^{Public} - r^{Private}$. Thus, entrepreneur *i* will take her firm public if

$$r^{Public} - r^{Private} + \varepsilon_i > 0 \Leftrightarrow \varepsilon_i > -\left(r^{Public} - r^{Private}\right). \qquad (2)$$

Proposition 3 implies that the disclosure costs, information asymmetry, and minimum efficient scale in firm *i*'s industry affect the probability of firm *i* being public by altering $-\left(r^{Public} - r^{Private}\right)$, the threshold value of the idiosyncratic return ε_i beyond which it is optimal for the entrepreneur to take her firm public. From a macro perspective, in equilibrium I expect to find a larger share of public firms in industries in which the threshold $-\left(r^{Public} - r^{Private}\right)$ is smaller. Therefore, Proposition 3 yields the following prediction.

Prediction 1. *The probability of a firm being public and the share of public firms in an industry are:*

a) decreasing in the cost of revealing strategic information to competitors in the industry;

b) decreasing in the degree of asymmetric information between investors and public firms in the industry;

c) increasing in the minimum efficient scale of firms in the industry.

I will test this macro prediction both at the firm level, using data on public and private U.S. firms, and at the industry level, analyzing the distribution of public and private firms across industries in the U.S. economy. Following that, I will test two micro predictions concerning firms' precautionary cash policies that are implied by Propositions 1 and 2:

Prediction 2. *Public firms hold more cash, ceteris paribus, than private firms.*

Prediction 3. *The difference in cash holdings between public and private firms is:*

a) increasing in the industry cost of revealing strategic information to competitors;

b) increasing in the industry rate of return on capital;

c) increasing in the industry volatility of cash flows.[25]

Testing Predictions 2 and 3 requires data on both public and private firms in the same industry, in order to be able to compare the cash policies of similar firms that differ only on their listing status. This will ensure that precautionary reasons are the main driving force behind the observed differences in the cash policies of public and private firms.[26] The idiosyncratic factors ε that affect an entrepreneur's going-public decision provide the required variation in firms' public or private status within the same industry, as captured by equation (2). However, this raises a potential identification problem. If ε is correlated with the firm's cash policy, my empirical analysis of the cash policies of public and private firms will be subject to an endogeneity problem. I describe how I address this problem in the empirical part of the paper.

---

[25] Prediction (3.c) requires the assumption made in part (c) of Proposition 2. Also, because cost overruns are the only source of uncertainty in my model, their volatility determines the volatility of firms' cash flows.

[26] In particular, similar firms will have similar needs of cash to settle transactions. Other potential reasons that may drive the differences in cash between public and private firms, such as agency or tax considerations (Bates, Kahle, and Stulz (2009)), are addressed in the introduction and in Section 3.3, respectively.

2. Empirical Results: Industry Characteristics and the Public vs. Private Trade-off

Prediction 1 states that the probability of a firm being public increases in the industry's minimum efficient scale and decreases in the industry's disclosure costs and information asymmetry. In this section, I test this prediction first at the firm level and then at the industry level.

2.1. Firm-level analysis

2.1.1. Data and sample construction

My dataset combines data on U.S. public firms obtained from standard sources such as CRSP, Compustat, and I/B/E/S with data on U.S. private firms obtained from a new database provider called Sageworks. Sageworks contains accounting data from income statements and balance sheets, similar to Compustat, except that Sageworks covers private rather than public firms. (Further details and descriptive statistics can be found in Asker, Farre-Mensa, and Ljungqvist (2010b).)

Sageworks obtains the data from a large number of accounting firms which input data for *all* their clients directly into Sageworks' database. Therefore, selection does not operate at the firm level, which could give rise to selection concerns, but at the level of the accounting firm. Sageworks co-operates with most national mid-market accounting firms (those below the 'Big Four') and with hundreds of regional players, but with few of the thousands of local accountants who service the smallest U.S. businesses. This selection is not a concern in my analysis, given that the smallest firms in the economy have no realistic choice between staying private and going public. Records are never deleted in Sageworks, so there are no survivorship bias concerns.

My version of the Sageworks database covers fiscal years 2001 through 2008. I use 2001 to construct lags, leaving a seven-year panel with more than 270,000 firm-years. After excluding Canadian firms, firms with data quality problems (i.e., those violating basic accounting identities), and those with incomplete data, I am left with just under 85,000 private firms.

My public-firm sample consists of firms in the CRSP-Compustat database that are located and incorporated in the U.S., listed on a major U.S. exchange (NYSE, AMEX, or Nasdaq), have valid stock prices in CRSP, and have a CRSP share code of 10 or 11 (which screens out non-operating entities such as real estate investment trusts or mutual funds).

As is customary, I exclude financial firms (SIC 6), regulated utilities (SIC 49), and government entities (SIC 9) from both the public and private samples.

To ensure that firms in my sample are of sufficient size so that being public is a feasible option for them, I exclude 46,427 private firms and 272 public firms with average sales of less than $5 million in real 2005 dollars. (Brav (2009) uses a similar threshold with U.K. data.) My results are robust to using other thresholds (such as $1 million or $10 million in sales) and to excluding no public firms.[27]

2.1.2. Industry disclosure cost, information asymmetry, and minimum efficient scale

All industry variables are constructed at the five-digit NAICS level, equivalent to four-digit SIC.

Industry disclosure cost

I use two measures of industry disclosure cost. The first, *patenting intensity*, is based on the notion that disclosure costs are lower in industries in which firms rely on patents to protect their proprietary information. The following quote from Hall, Jaffe, and Trajtenberg (2001) helps explain why: "The whole idea of patents is that they constitute a 'package deal,' namely, the grant of temporary monopoly rights in exchange for *disclosure*. […] The inventor has to make a strategic decision to patent, as opposed to rely on secrecy or other means of appropriability." Levin et al. (1987) provide survey evidence consistent with the notion that there exists substantial cross-industry variation in the effectiveness of patents to protect intellectual property.

Patenting intensity is measured as follows. For each firm in an industry, I compute the log of one plus the number of patents the firm requested in 2002 (my first sample year) and then average across all firms in the industry. (Patent data comes from the NBER patent data project.) This measure is expected to be inversely related to industry disclosure costs. For expositional clarity, I change its sign so that it directly correlates with disclosure costs. A potential concern is that patenting intensity is strongly correlated with R&D intensity, which at the same time might be correlated with other factors affecting the trade-off between being public and private. To account for this, I control for an industry's R&D intensity when using patents to measure disclosure costs. For each industry-year, I measure R&D intensity as the average

[27] I use sales as a measure of size when constructing my sample (and later on when matching) because, unlike total assets, sales are not directly affected by the differences in cash holdings between public and private firms documented in Section 3. My results are robust to excluding small firms based on their total assets.

R&D expenditure by public firms in the industry (R&D expenses are not available in Sageworks).

The second measure, *information leakage*, draws on the work of Asker and Ljungqvist (2010) who show that concerns about disclosure of confidential information to product-market competitors determine firms' choice of which investment bank to have a relationship with. The measure builds on the notion that firms in industries with high disclosure costs are less willing to share an investment bank with other firms in their industry, because they are concerned about the risk of information leaking to their competitors when a bank is shared. I measure information leakage using the log of the maximum number of top 10 public firms (ranked by sales) in an industry that have shared an investment bank.[28] It ranges from eight firms sharing a bank in the residential building construction industry to zero firms sharing a bank in the semiconductor and other electronic component manufacturing industry. Like patenting intensity, this measure is inversely related to industry disclosure costs, and thus I also change its sign so that it directly correlates with disclosure costs.

Another potential way to measure disclosure costs would be to directly examine firms' disclosure policies. While information on company issued guidance (CIG) is readily available from standard databases (e.g., First Call), CIG is actually not a good measure of disclosure costs. Most CIG consists only of an estimate of the firm's future earnings, without the firm adding any qualitative information on its current or future projects. Data from other disclosure channels more prone to contain qualitative information (e.g., conference calls) are less readily available.

Industry information asymmetry

For each industry-year, I measure information asymmetry as the median forecast error made by financial analysts for public firms in the industry.[29] Forecast errors equal the absolute value of the difference between the outstanding consensus (i.e., mean) forecast for earnings per share (EPS) prior to the end of a firm's fiscal year and the firm's actual EPS, both taken from I/B/E/S. Following the accounting literature (e.g., Krishnaswami and Subramaniam (1999)), I standardize forecast errors by the

[28] Following Asker and Ljungqvist (2010), I code *n* firms in an industry as sharing a bank if a single bank has lead-managed one or more debt issues for each of the *n* firms over the previous five years. The measure is constructed from quarterly data from Thomson Financial. Specifically, for each industry, I use the maximum value of *n* over the 12 quarters ending at the beginning of my sample period (the fourth quarter of 2002). I set *n* to one if no firms in the industry share an investment bank. I then define information leakage as the natural logarithm of *n*.

[29] I use the industry median to account for the fact that the distribution of analysts' forecast errors is highly skewed.

firm's share price the day before the consensus forecast was calculated. For robustness, I also measure information asymmetry as the median dispersion of analysts' earnings forecasts across firms in the industry, using the standard deviation of earnings forecasts at the end of a firm's fiscal year.

Industry minimum efficient scale

For each industry-year, I measure an industry's minimum efficient scale as the median of the log of one plus property, plant, and equipment across all firms, public and private, in the industry.[30] This follows Salinger, Caves, and Peltzman (1990) and Sutton (1991), among others.

2.1.3. Empirical model and results

To analyze how the characteristics of a firm's industry affect its decision to be public or private, I estimate the following equation:

$$\begin{aligned} Public_{it} = {} & \beta_1 \, Industry\ disclosure\ cost_{it} + \beta_2 \, Industry\ information\ asymmetry_{it} + \\ & + \beta_3 \, Industry\ minimum\ efficient\ scale_{it} + \theta \, R\&D\ intensity_{it} + \alpha_t + \varepsilon_{it} \end{aligned} \tag{3}$$

where *Public* equals one if the firm is publicly traded and zero otherwise; the other covariates are described above.[31] I estimate equation (3) using a probit model; my results are robust to using either a logit or a linear probability model. Throughout, I report heteroskedasticity-robust standard errors clustered at the firm level.

The estimation results, reported in Table 2, are consistent with Prediction 1: The probability of a firm being public decreases in industry disclosure costs and information asymmetry, and it increases in the industry's minimum efficient scale. I report results for four specifications, combining my two measures of industry disclosure costs (patenting intensity and information leakage) with my two measures of information asymmetry (analysts' forecast errors and forecast dispersion).

In all four specifications, the three effects predicted by my model are not only statistically but also economically significant. To illustrate, column 1 shows that a one-standard-deviation increase in industry disclosure costs, measured by patenting intensity, is associated with a 6.1 percentage point decrease in the probability of a firm in the industry being public (from 25.9% at the mean of the independent variables to

[30] I obtain the data on private firms from Sageworks and, as elsewhere in the paper, I focus on firms with average real sales of at least $5 million. Note that this is not ideal, since Sageworks does not cover the whole universe of private firms.

[31] As argued before, it is important to control for R&D intensity if patenting intensity is used to measure disclosure costs.

19.8%). In the case of information asymmetry, a one-standard-deviation increase in the median forecast error in an industry is associated with a one percentage point decrease in the probability of being public. Finally, a one-standard-deviation increase in an industry's minimum efficient scale is associated with a sizeable 21.6 percentage point increase in the probability of a firm in the industry being public.

2.2. Industry-level analysis

Proposition 3 predicts that the return to being public increases in the minimum efficient scale and decreases in the disclosure costs and information asymmetry in a firm's industry. In this section, I test whether this prediction helps explain the share of public firms in an industry.

2.2.1. Measuring the share of public and private firms in an industry

There is no unambiguous measure of the share of public firms in an industry. A natural way is to measure the share of output produced by public firms. The U.S. Census Bureau collects data on the total output produced per industry every five years as part of the Economic Census.[32] Since the Census provides no breakdown by listing status, I measure the output of public firms by adding up their annual sales as reported in Compustat. Unfortunately, this introduces a potential measurement error. Compustat reports worldwide sales while the Census reports U.S. output. As a result, the share of public firms exceeds one in 5.5% of the industries. As pointed out by Gabaix (2010), who faces the same problem, current data sources do not provide a satisfactory solution to this problem.

Alternatively, I follow Chod and Lyandres (2010) and measure the share of public firms simply as the fraction of large firms in an industry that are public. The Census provides data on the total number of firms in each industry (see footnote 1), and I obtain the number of public firms from CRSP-Compustat. Like Chod and Lyandres, I exclude small firms (those with fewer than 100 or 500 employees) from the total. This avoids understating the economic importance of public firms in industries with thousands of small private firms and a few large public firms that account for most of the production.

The share of public firms across industries averages 39.8% when measured using output, and 14.2% and 6.2% when measured as the fraction of firms with at least 500 or 100 employees, respectively. The

[32] I use data from the 2002 Economic Census, which is the most recent Census with complete data. This is the same year I measure industry disclosure costs. The data is available at http://www.census.gov/epcd/www/concentration.html.

corresponding medians are 16.3%, 8.0%, and 3.3%.

2.2.2. Empirical model and results

To analyze whether my model can explain the distribution of public and private firms across industries in the U.S. economy, I estimate the following equation:

$$\log\left(\textit{Public Share}_j\right) = \beta_1 \textit{ Industry disclosure cost}_j + \beta_2 \textit{ Industry information asymmetry}_j + \\ + \beta_3 \textit{ Industry minimum efficient scale}_j + \theta \textit{ R\&D intensity}_j + \varepsilon_j \quad (4)$$

where *Public Share* is described above and modeled in logs to account for skewness. I estimate the equation by OLS and report robust standard errors. Each observation in equation (4) is a five-digit NAICS industry. As before, I exclude financial services, utilities, and the government sector.

The results, reported in Table 3, are again consistent with Prediction 1: There are proportionately more public firms in industries with lower disclosure costs, less information asymmetry, and higher minimum efficient scale. I present results for six different specifications. Columns 1 through 3 estimate equation (4) with my three measures of *Public Share*, using patenting intensity to measure the cost of disclosure in the industry. Columns 4 through 6 use information leakage to measure disclosure costs, and again report the estimation results for my three measures of public share.

The estimated effects of disclosure costs and information asymmetry are statistically significant in four of the six specifications, and the effect of minimum efficient scale is significant in all six. These three factors have an economically important effect on the industry distribution of public firms. To illustrate, in column 6, a one-standard-deviation increase in disclosure costs (here: information leakage) is associated with a 17.5% decrease in the share of public firms in the industry. The corresponding effects of information asymmetry (measured as forecast errors) and minimum efficient scale are an 11.5% decrease and a 37.9% increase in the share of public firms, respectively. The magnitudes of the estimated effects in the other columns are similar.

2.3. Discussion

The results presented in this section are consistent with Prediction 1. They support the notion that the trade-off between the disclosure cost and the liquidity benefit of being public is an important determinant of a firm's choice between going public and staying private. Clearly, other factors left out of my model also play an important role in the going-public decision, such as the desire to maintain decision-making control. These allow me to observe public and private firms in the same industry, enabling me to compare the cash policies of public and private firms in Section 3.

3. The Public vs. Private Trade-off at Work: Comparing Cash Policies

This section tests the predictions that disclosure costs induce public firms to hold more cash than private firms (Prediction 2) and that the difference in their cash holdings increases in industry disclosure costs, returns on capital, and cash flow volatility (Prediction 3). After describing my data and presenting the baseline models, I address a potential endogeneity problem. Because going public is a choice, public and private firms may differ in unobservable ways that correlate with their cash policies. Thus, an OLS regression of the effect of listing status on cash holdings may yield inconsistent results.

3.1. Variable definitions

3.1.1. Dependent variable

Following Bates, Kahle, and Stulz (2009) and others, my main measure of a firm's cash holdings is the ratio of cash (including short-term investments) to total assets. In robustness checks, I use the log of cash to sales and the log of cash to net assets (where net assets equals total assets minus cash).

3.1.2. Explanatory variables

In addition to a dummy variable capturing a firm's listing status, the main explanatory variables of interest are industry disclosure costs, returns on capital, and cash flow volatility. My model predicts that each of these should have a greater effect on the cash holdings of public firms than of private firms. In addition, I include the same common controls as Bates et al. (2009).

<u>Industry disclosure costs</u>

As before, my first measure of disclosure costs is *information leakage*. Patenting intensity, however, is problematic in the present context, as it affects cash holdings not only through disclosure costs but also through its effect on financial distress.[33] Therefore, I use *industry concentration* as my second measure of disclosure costs. This measure is motivated by prior evidence that firms in more concentrated industries

[33] Patents are an important part of intangible assets, which are positively correlated with costs of financial distress (Korteweg (2007)). And firms with high distress costs have been shown to hold more cash (e.g., Bates et al. (2009)).

disclose less information (see Bamber and Cheon (1998), Harris (1998), or Ali, Klasa, and Yeung (2010)), which suggests that in more concentrated industries the cost of disclosure is higher.[34]

Industry concentration is measured from the 2002 Economic Census. For each industry, I measure concentration as the fraction of sales made by the four largest firms in the industry (*C4*). My results are robust to using the eight largest firms (*C8*) and to using a Herfindahl-Hirschman index (HHI) for the 50 largest firms (though HHI is only available for manufacturing industries).

Industry returns on capital

For each industry-year, I measure industry returns on capital as the mean market-to-book for firms in the industry. (Only public firms in the industry contribute to the mean, since market-to-book is unavailable for private firms.)

Industry cash flow volatility

Following Bates et al. (2009), I define industry cash flow volatility as the standard deviation of industry cash flow to assets, computed as follows. For each firm-year, I compute the standard deviation of cash flow to assets for the previous ten years. I require at least three annual observations. I then average the firm cash flow standard deviations each year in each industry. Unlike elsewhere, I use three-digit NAICS codes to ensure I have enough firms in each industry (this also follows Bates et al. (2009)). Only public firms contribute to the mean, as there is no data on private firms that predate my sample.

Additional control variables

I include as additional control variables those used by Bates et al. (2009), following their definitions whenever possible. Specifically, I include cash flow to assets, working capital net of cash to assets, investment to assets, book leverage, and a dividend dummy. I also control for firm size (log of total assets), thus accounting for economies of scale in holding cash (Bates et al (2009)). For each industry-year, I include the mean of R&D to sales and the mean of acquisitions to assets, constructed using data for public firms only as R&D and acquisition expenditures are not available in Sageworks.

In addition, I control for industry information asymmetry, using the median analysts' forecast error.

[34] Concentration is highly correlated with the share of public firms in an industry as the largest firms in an industry are very likely to be public. For this reason, I do not use industry concentration to measure disclosure costs in the previous section.

My results in Section 2 show that information asymmetry is correlated with a firm's listing status, but my model predicts that its effect on cash holdings is ambiguous. My model also predicts that industry minimum efficient scale should have no effect on cash policies. I confirm this prediction in Table 5 below and so do not control for minimum efficient scale in my main cash regressions.

3.1.3. Descriptive statistics

Table 1 shows descriptive statistics (mean, median, and standard deviation) for the public and private firms in my baseline sample of firms with average annual real sales of at least $5 million during the sample period (2002 to 2008). Not surprisingly, public firms are substantially larger than private firms: The median public firm has total assets and sales of 312.7 and 282.1 million (real 2005) dollars, respectively, while total assets and sales are 5.1 and 12.3 million dollars for the median private firm. (See also the top graph in Figure 4 which shows the size distribution of public and private firms.) In addition, there are significant differences in the industry distribution of public and private firms. To ensure that the differences in their cash policies are not driven by these differences in scale and industry, in section 3.3.1 I match public and private firms on industry and size.

The statistics in Table 1 are consistent with my model prediction that public firms hold more cash than private firms: For public firms, the mean and median of cash over assets are 0.22 and 0.13, respectively; for private firms, the mean and median are 0.11 and 0.06. Of course, these descriptive statistics do not condition on anything, and thus this cannot be seen as a formal test of this prediction.

3.2. Baseline empirical strategy and results

3.2.1. Baseline model

The following equation models differences in public and private firms' cash policies and analyzes how these differences are driven by public firms' inability to communicate selectively with investors:

$$\begin{aligned} Cash_{it} = {} & \beta_1 Industry\ disclosure\ cost_{it} + \beta_2 Public_i \times Industry\ disclosure\ cost_{it} + \\ & + \gamma_1 Industry\ return\ on\ capital_{it} + \gamma_2 Public_i \times Industry\ return\ on\ capital_{it} + \\ & + \delta_1 Industry\ cash\ flow\ volatility_{it} + \delta_2 Public_i \times Industry\ cash\ flow\ volatility_{it} + \\ & + \lambda Public_i + \vec{\theta}\ Additional\ controls_{it} + \alpha_t + \varepsilon_{it} \end{aligned} \tag{5}$$

Table 4 presents the results. Columns 1 and 3 test Prediction 2 that public firms hold more cash than private firms by estimating equation (5) without interaction effects (i.e., I impose $\beta_2 = \gamma_2 = \delta_2 = 0$). Regardless of whether I measure disclosure costs using industry concentration (column 1) or information leakage (column 3), I find strong support for Prediction 2. The point estimates of λ show that public firms' cash holdings are 12.5 to 13.3 percentage points higher than those of private firms. Both estimates are highly significant ($p<0.001$).

While establishing that public firms hold more cash than private firms is of interest in itself, the purpose of my cash analysis is to show that this stylized fact is due, at least in part, to the disclosure costs associated with being public. To show this, columns 2 and 4 estimate equation (5) with all interaction effects included. The coefficients of interest, β_2, γ_2, and δ_2, capture the *additional* effect that industry disclosure costs, returns on capital, and cash flow volatility, respectively, have on the cash holdings of public firms relative to their effect on private firms. As shown by my model (Prediction 3), these three interaction effects directly test the hypothesis that the higher cash holdings of public firms are an optimal response aimed at minimizing disclosure costs.

The results are consistent with Prediction 3. In column 2, I find that a one-standard-deviation increase in disclosure costs, measured as industry concentration, is associated with a 0.4 percentage point decrease in private firms' cash holdings and a 1.3 percentage point increase in public firms' cash holdings. The difference of 1.7 percentage points is highly significant, both statistically and economically. It represents 13.0% of the cash holdings of the median public firm. In column 4, the estimated effect of disclosure costs is of a similar magnitude: A one-standard-deviation increase in disclosure costs, measured using information leakage, is associated with a 0.5 percentage point decrease in private firms' cash holdings and a 1.3 percentage point cash increase in public firms. The differential effect of 1.8 percentage points is again highly significant. These results are consistent with the prediction that disclosure costs induce public firms, but not private ones, to hoard cash.

Disclosure is more costly for firms in industries with high returns on capital, and thus my model predicts that the effect of returns on cash holdings will be larger for public firms. This is precisely what I find. In column 2, a one-standard-deviation increase in return on capital is associated with a 0.6 percentage point increase in private-firm cash holdings. For public firms, the effect is 2.9 percentage points. The corresponding effects in column 4 are 1.0 percentage point for private firms and 4.6 percentage points for public firms.

Firms in industries with more volatile cash flows should hold more cash to avoid having to raise capital when hit by a negative cash flow shock. This effect will be greater for public firms, since disclosure costs induce them to minimize the need of raising capital. My results support this prediction. In column 2, a one-standard-deviation increase in cash flow volatility is associated with a 1.7 percentage point increase in cash holdings for private firms and 5.9 percentage points for public firms. The corresponding effects in column 4 are 0.7 percentage points for private firms and 4.3 percentage points for public firms. As before, both interaction effects are highly significant.

A natural extension of my model suggests that negative cash flows induce public firms to hoard cash, to minimize disclosure costs incurred when raising capital to cover losses. My results support this prediction: Cash flow has a negative effect on the cash holdings of public firms but a positive effect for private firms. (In Section 4, I compare the cash flow sensitivity of cash in public and private firms.) The literature has argued that R&D spending is positively related to financial distress costs (Korteweg (2007), Bates et al. (2009)). Thus, the interaction effect of R&D with being public should be positively related to cash holdings.[35] The data support this. As for industry information asymmetry, its effect on cash holdings changes across specifications, consistent with the ambiguous effect predicted by my model. The remaining control variables behave in line with prior work (see Bates et al. (2009)).

3.2.2. Alternative specifications

In Table 5, I analyze the robustness of the results presented above to several alternative specifications. I organize the results in two panels: In Panel A, I add new control variables. In Panel B, I include the same control variables as in my baseline regressions in Table 4, but I either restrict the sample or define some variable differently (for brevity, I only report the estimates that directly test the predictions of my model). Also to conserve space, in both panels I only report the results of estimating my cash equation (5)

[35] Firms with high financial distress costs are more likely to try to avoid liquidation when facing a negative shock, which increases the value of precautionary cash holdings if the firm is public.

with all interaction effects included. (The last row in each column reports the public dummy estimate when the specification reported in that column is estimated with no interaction effects.)

Specifically, in columns 1 and 2 of Panel A I add to equation (5) a dummy variable that identifies firms with negative cash flow, both in levels and interacted with the public dummy. Regardless of whether I use industry concentration (column 1) or information leakage (column 2) to measure disclosure costs, I find that, unlike private firms, public firms with operating losses tend to hoard cash. This finding is consistent with the notion that disclosure costs make it costly for public firms to raise capital, and it confirms the interpretation given above to the estimated negative effect of cash flow on public firms' cash holdings. In columns 3 and 4 of Panel A, I test my model prediction that industry minimum efficient scale has no effect on firms' cash policies. The results confirm this prediction.[36]

I turn now to Panel B. In columns 1 and 2, I estimate equation (5) excluding 2007 and 2008 from the panel, to ensure that my results are not affected by the recent financial crisis. In columns 3 and 4, I limit my sample of private firms to those incorporated under Subchapter C of the Internal Revenue Code (the same legal form that the vast majority of public firms use), to ensure that my results are not driven by differences in the taxation regime of public and private firms. My results are robust to using either of these two subsamples, both when measuring disclosure costs as industry concentration (columns 1 and 3) and when using information leakage (columns 2 and 4).[37]

In columns 5 and 6, I use an alternative measure of industry concentration: the Herfindahl-Hirschman index (HHI). The Census only reports the HHI for manufacturing firms, so in column 5 my sample is reduced. In column 6, I use *C4, C8, C20,* and *C50* (available for all industries) to compute a simplified version of the HHI. My results are robust to either of these two approaches.

In columns 7 through 10, I check the robustness of my results to using the log of cash to sales and the log of cash to net assets as the dependent variable; as elsewhere, I report results with my two measures of disclosure costs. As in my baseline analysis, I find strong support for my model predictions (note that the point estimates are very different because the dependent variable is in logs). For instance, in column 7 I

[36] In unreported results, I interact minimum efficient scale with being public and I find that, regardless of how I measure disclosure costs, neither the level nor the interaction effect are significant at the 5% level.
[37] My results are also robust to excluding those few private firms (5%) that either use cash instead of accrual accounting or do not report which accounting method their use.

find that a one-standard-deviation increase in industry disclosure costs is associated with a 12.1% increase in the cash holdings of public firms, while in the case of private firms it is associated with a 5.4% decrease in their cash holdings.

3.3. Endogeneity of a firm's public or private status

An important concern with the results presented in the previous section is that a firm's listing status reflects an endogenous choice. Equation (5) explicitly controls for the industry characteristics that, according to my model, affect both the entrepreneur's listing decision and her cash policy. But if any unmodeled idiosyncratic factors affecting the listing decision also correlate with cash policy, the results in Section 3.2 will be inconsistent. I address this potential endogeneity problem in three ways. The first is a matching estimator that allows me to capture other *observable* firm and industry characteristics (than those predicted by my model) that affect the listing decision and that may correlate with a firm's cash policy in a way not captured by equation (5). (See Imbens and Wooldridge (2009) for a recent review of matching estimators.) In particular, a potential concern is that including firm size as a control in equation (5) might not be enough to fully capture the effect of size on firms' cash policies when firms are of very different scales.[38] Given that firm size is highly correlated with a firm's listing status, this would lead to endogeneity.

Matching methods are generally not robust to selection on *unobservables*. Thus, my second test exploits within-firm variation in listing status by analyzing changes in cash holdings among firms that go public for the sole purpose of allowing existing shareholders to cash out. This restriction ensures that the firm raises no capital in the IPO and so cash holdings do not increase mechanically. This test removes biases associated with systematic (ex-ante) differences between public and private firms, even if they are unobservable, but the endogeneity of the decision to go public still remains a concern. To address this concern, I will instrument the going-public decision using the location of a firm's headquarters and its industry's minimum efficient scale.

3.3.1. Selection on observables: Matching estimators

[38] For instance, large firms are more likely to have foreign operations, and the literature has argued that multinational firms hold more cash because of taxes (Foley et al. 2007). This might induce a non-linear effect of size on cash.

I use two ways to match public (in the matching terminology: treated) firms to private (untreated) firms: Local linear propensity score matching, with the propensity scores based on firm size, cash flow, sales growth, and industry; and nearest-neighbor matching on industry and size. (Given the panel structure of my dataset, I match at the firm rather than at the firm-year level.)

Local linear propensity score matching

Step one estimates propensity scores (Rosenbaum and Rubin (1983)) from the following probit:

$$Public_i = \beta_1 \operatorname{Log}\left(Sales\right)_i + \beta_2\ Cash\ flow_i + \beta_3\ Sales\ \mathrm{growth}_i + \sum_j \gamma_j NAICS5_{ji} + \varepsilon_i$$

$\operatorname{Log}\left(Sales\right)$, *Cash flow*, and *Sales growth* are averaged within-firm. The *NAICS*5 are five-digit NAICS industry dummies which here subsume all industry characteristics that affect a firm's decision to go public. See Michaely and Roberts (2007), Brav (2009), and Saunders and Steffen (2010) for a similar matching approach with data from the U.K. Some of these authors also include firm age in their probit models but this variable is not available in Sageworks.[39]

Public and private firms are then matched on their propensity scores using a local linear matching algorithm (LLM). LLM is a variation of kernel matching (see Heckman, Ichimura, and Todd (1997, 1998)). Similarly as kernel matching, LLM assigns to each private firm in the matched sample a weight that depends on the distance between its propensity score and the scores of public firms. The distance is evaluated using a kernel function. The larger the distance between the score of the private firm and the scores of public firms, the lower the weight.[40] LLM differs from kernel matching in that the weights contain a linear term in the propensity score in addition to the intercept. This is helpful when the non-treated observations fall on one side of the treated observations (Fan, 1992a,b), as is the case in my dataset (see the top graph in Figure 3).

Local linear matching uses all firms in the dataset, so it is important to impose a common support

---

[39] Ideally, the control variables used when estimating the propensity scores should simultaneously affect the participation decision (going public) and the outcome variable (cash holdings), while not being directly affected by participation (Caliendo and Kopeinig (2008)).Therefore, I do not include in the probit model variables such as leverage or investment, which are likely to be affected by a firm's listing status (see Brav (2009) and Asker, Farre-Mensa and Ljungqvist (2010a), respectively). Arguably, firm size is also affected by listing status. However, controlling for size differences between public and private firms is an important motivation to follow this matching approach, and thus I include size in the probit.

[40] I use a Gaussian kernel and set the bandwidth to 0.06, but my results are not sensitive to either of these choices.

restriction to ensure that the propensity scores of public and private firms overlap. I do this in two steps. First, as before, I use only firms with average real sales of at least $5 million. Second, I exclude from the match any firm whose propensity score either exceeds the 10th largest private-firm propensity score or falls below the 10th smallest public-firm propensity score. This robust variation of the maxima and minima criteria is suggested by Lechner (2002).

Figure 3 shows the distribution of propensity scores for public and privates firms both before matching (top graph) and after matching (bottom graph). The bottom figure illustrates how, by overweighting those private firms with high propensity scores, LLM produces a sample in which the distribution of propensity scores for public and private firms is nearly identical.

Table 6, Panel A presents the results of estimating equation (4) after matching (standard errors are bootstrapped to account for matching-induced variance).[41] The results mirror those in Table 4. Public firms hold around 11 percentage points more cash than private firms (columns 1 and 3) and disclosure costs appear to be an important force driving this difference (columns 2 and 4).

<u>Nearest-neighbor matching on industry and size</u>

A key feature of local linear matching is that all firms on the common support are part of the matched sample, regardless of whether a close match exists for them. An alternative algorithm that restricts the matched sample to public and private firms for which a close match exists is nearest-neighbor matching. Following Asker, Farre-Mensa and Ljungqvist (2010a), this matched sample is constructed as follows. Starting in 2002, for each public firm, I identify the private firm in the same four-digit NAICS industry and fiscal year closest in terms of sales such that $\max(Sales_{public}, Sales_{private}) / \min(Sales_{public}, Sales_{private}) < 2$. If no match can be found in a given fiscal year, the observation is discarded and a new match is attempted for that firm in the following year. Once a match is formed, it is kept intact for as long as both the public and private firms remain in the sample, to maximize the available time series for each firm. If a matching firm exits the panel, a new match is spliced in. Matching is with replacement to maximize the match rate.

[41] In each of the 500 bootstrap replications, a new set of propensity scores is computed, firms are matched, and equation (5) is estimated.

Figure 4 shows the distribution of log real sales for public and private firms in the baseline sample (top graph) and in the matched sample (bottom graph). The bottom figure shows that the overlap in the matched sample is near perfect, indicating that the match is good.

Table 6, Panel B presents the results of estimating equation (5) using this matched sample. The results continue to support Predictions 2 and 3. Public firms hold between 12.1 and 13.1 percentage points more cash than private firms (columns 1 and 3), in line with the estimates obtained in Table 4 without matching. In columns 2 and 4, the interaction effects of the public dummy with industry disclosure costs, returns on capital, and cash flow volatility are significant at the 5% level or better. The estimated magnitude of the interaction effect of disclosure costs is actually larger than in Table 4: The additional effect of a one-standard-deviation increase in disclosure costs on the cash holdings of public firms is 2.5 percentage points in both columns 2 and 4 (with disclosure costs measured using industry concentration and information leakage, respectively). The analogous estimated differences in Table 4 were 1.7 and 1.8, respectively.

Taken together, my matching results alleviate the concern that the results reported in Table 4 are an artifact of observable differences between public and private firms, and in particular, of differences in size or industry characteristics.

3.3.2 Within-firm changes in cash holdings around IPOs

Table 7 examines changes in cash holdings around the IPO for a particular sample of IPO firms: firms that go public for the sole purpose of allowing existing shareholders to cash out. This ensures that the IPO proceeds do not mechanically increase the firms' cash holdings, while differencing away firm unobservables.

Consistent with Prediction 2, my results in columns 1 through 3 show that firms increase their cash holdings following their IPO. To ensure that these findings are not driven by an economy-wide increase in cash holdings, in columns 4 and 5 I match each IPO firm to up to five already public firms (matching is on industry and size, and it is done the first year that the IPO firm is in the sample). I find that, before they go public, the IPO firms hold less cash than public firms, but their cash holdings become indistinguishable from those of other public firms once they are public. This rules out the concern that the results in columns 1 through 3 might be driven by economy-wide increases in cash.

3.3.3 Selection on unobservables: Instrumental variables

If public and private firms differ systematically in unobservable characteristics and if these characteristics also affect cash policy, my results may be biased. This is true even in my within-firm analysis if the decision to go public is triggered by a change in these unobservable characteristics. What might these unobservable characteristics be? Brau and Fawcett's (2006) survey indicates that two important reasons why private firms choose to stay private are the "desire to maintain decision-making control" and "to avoid ownership dilution". This suggests that the owner-managers of private firms that remain private value their ability to manage with little external oversight more than the owners who take their firms public. Arguably, this difference in preferences should induce the private firms in my sample to hold more cash than they would if their owners had an average desire for control. In such case my estimates of the difference in cash holdings between public and private firms would actually be biased downwards, so they should become larger when I instrument the going-public decision. Of course, other unobserved differences may bias my results in the other direction.

The ideal instrumental variable (IV) provides exogenous variation in the listing choice without otherwise being related to cash policy. I use two distinct instruments. The first consists of a set of dummy variables capturing the U.S. state of location of firms' headquarters. This instrument is motivated by the finding that investors are better able to obtain information on nearby firms and that this translates into higher returns when they invest in local companies (see, e.g., Coval and Moskowitz (1999), Huberman (2001), Ivkovic and Weisbenner (2005)). In related evidence, Malloy (2005) and Bae, Stulz, and Tan (2008) show that security analysts located closer to a firm produce more accurate forecasts than those located further away. This suggests that firms headquartered in areas with a larger pool of investors and security analysts are more likely to go public, all else equal, as they face a lower cost of producing information for their potential investors. Loughran (2008) provides direct evidence on the impact of location on a firm's ability to issue equity, showing that U.S. rural firms are less likely to conduct seasoned equity offerings than firms located in urban areas. As for the exclusion restriction, it seems unlikely that a firm's location will directly affect its cash policy, after controlling for the variables included in equation (5), other than through its effect on the listing decision.

Saunders and Steffen (2010) use a related instrument for the decision to go public in the U.K.: distance to London. This is suitable in a country like the U.K., where London is the only big financial centre and transportation hub. However, the U.S. is arguably more multimodal. Consider Loughran's (2008) argument: "Is a firm located in urban Los Angeles or one located in rural Bismarck, North Dakota, farther from institutional investors in New York City? Measured in miles, the company in Los Angeles is much farther away. Yet, it is also much easier for the institutional investor to reach. There are numerous direct flights from New York to Los Angeles every day. Getting to Bismarck is difficult, and once there, the analyst is almost certainly stuck for the night."

My second instrument is the industry's minimum efficient scale. Its use is motivated by the following prediction of my model: the higher the scale required to operate in an industry, the more likely firms in the industry are to be public. The empirical results in Section 2 support this prediction. On the other hand, my model predicts that the industry minimum efficient scale has no direct effect on precautionary cash holdings, which is consistent with the exclusion restriction.[42]

To implement my IV estimator, I first estimate the following equation by OLS:

$$\begin{aligned} Public_{it} &= \beta\, Industry\ disclosure\ cost_{it} + \gamma\, Industry\ return\ on\ capital_{it} + \delta\, Industry\ cash\ flow\ volatility_{it} + \\ &+ \vec{\theta}\, Additional\ controls_{it} + \sum_s \phi_s State_{sit} + \eta\, Industry\ minimum\ efficient\ scale_{it} + \alpha_t + \varepsilon_{it} \end{aligned}$$

I use a linear probability model to avoid identification in the second stage being driven by non-linearity in the first stage. The state dummies are jointly significant at the 1% level (the *F*-statistic is 6.1 when I use industry concentration to measure disclosure costs, and 4.0 when I use information leakage); the effect of industry minimum efficient scale is positive, as expected, and significant at the 1% level (the *t*-statistic is 6.4 if disclosure costs are measured with industry concentration and 8.2 if information leakage is used).[43] The *Additional controls* vector is the same as in equation (5).

---

[42] In section 3.2.2 above I show that, after controlling for all the variables included in equation (5) (including firm size), industry minimum efficient scale has no significant effect, neither statistically nor economically, on cash holdings.

[43] The *F* statistics of a joint test of the significance of my two sets of instruments (state dummies and minimum efficient scale) are 6.5 if disclosure costs are measured using industry concentration and 5.1 if information leakage is used. The fact that these *F*-stats are below 10 may raise concerns about the weakness of my instruments (Staiger and Stock (1997)), but the Kleibergen and Paap (2006) tests reported below reject both nulls of underidentification and weak identification.

Next, I estimate equation (5) by GMM, using the first-stage predicted probabilities, as well as their interactions with the control variables for which there is an interaction term in equation (5), as instruments for the public dummy and the interaction terms. This follows Wooldridge (2001).

Table 8 presents the estimation results. They continue to support the prediction that public firms hold more cash than private ones. The estimated difference is substantially larger than the estimates obtained in my previous specifications: 41.7 percentage points in column 1 and 36.0 percentage points in column 3. This is consistent with the notion that firms that choose to remain private are more likely to have a manager with a preference for control, and thus who values the independence provided by cash, than the average firm facing the choice between going public or staying private.

When I include the interaction effects, in columns 2 and 4, I again find that the difference in cash holdings increases in industry disclosure costs, returns on capital, and cash flow volatility. The point estimates are of similar size as in Table 4. Thus, endogeneity does not seem to bias the estimates of the interaction effects.

Table 8 also reports *rk* tests of the nulls of underidentification and of weak identification, based on Kleibergen and Paap (2006). These tests are robust to the fact that the errors in equation (5) are not *iid*. While the distribution of the weak identification test has not been fully worked out and thus the critical values should be approached with caution (Baum, Schaffer and Stillman (2007)), the large size of both test statistics in all four specifications suggests that neither underidentification nor weak identification is likely to be a concern in these models.

4. Additional Evidence: The Cash Flow Sensitivity of Cash of Public and Private Firms

Almeida, Campello, and Weisbach (2004) show that financial constraints can be captured by a firm's cash flow sensitivity of cash. Their argument builds on the notion that only those firms whose investments are constrained by capital market imperfections need to manage liquidity to maximize firm value. This implies that firms that anticipate financial constraints in the future should increase their cash holdings when cash flows are high, and thus their cash flow sensitivity of cash should be positive. In contrast, unconstrained firms' cash flow sensitivity of cash should be indeterminate.

The findings in Almeida et al. (2004) suggest an additional way of testing my model prediction that financing is costly for public firms due to their inability to selectively communicate with investors. Indeed, if disclosure costs constrain public firms' financing, then we should expect that, all else equal, public firms should exhibit a higher cash flow sensitivity of cash than private firms.

I test this prediction in Table 9 by comparing the within-firm cash flow sensitivity of cash of public and private firms. Following Almeida et al. (2004), the dependent variable in my analysis is the change in cash holdings. My main independent variable of interest is cash flow over assets, but I also control for investment opportunities (which I measure as industry return on capital, i.e., market-to-book) and firm size. I include firm and year fixed effects in all my specifications.

The first three columns in Table 9 report the results of estimating the cash flow sensitivity of cash for public firms only. In column 1, I use all public firms in my baseline sample of firms with average annual real sales of at least $5 million. In columns 2 and 3, I follow Almeida et al. (2004) and restrict my sample of public firms to those in the bottom and top three deciles of the size distribution, respectively (as elsewhere in the paper, I use sales to measure size). Consistent with their results, I find that small public firms exhibit a positive and significant cash flow sensitivity of cash, while large public firms exhibit no significant sensitivity. These findings reflect the notion that small public firms, which tend to be young, less well known, and with little reputation, are more financially constrained than large public firms.

In columns 4 through 6, I replicate the same analysis using the private firms in my baseline sample. The cash flow sensitivity of cash of private firms is positive and significant, but its size is less than a third of the estimated sensitivity for small public firms. Interestingly, the sensitivity of private firms seems to be largely independent of their size.[44]

In order to directly compare the cash flow sensitivity of cash of public and private firms, in columns 7 through 9 I combine data on public and private firms. I allow their cash flow sensitivity to differ by interacting cash flow with a dummy variable that equals one if a firm is public. In column 7, I use all public and private firms in my baseline sample, while in columns 8 and 9 I match public and private firms

[44] Following Almeida et al. (2004), I report standard R^2's. This implies that the R^2's for private firms cannot be directly compared to those for public firms, as the average panel length for private firms is shorter than for public firms.

similarly as in Section 3.3.1 (i.e., in column 8, I use local linear propensity score matching; in column 9, I use nearest-neighbor matching on size and industry). Matching allows me to compare the cash flow sensitivity of cash of public and private firms of similar size and other characteristics. Thus, if not for their different listing status, the matched firms should face similar financial constraints.

My results are consistent with the prediction that disclosure costs impose financial constraints on public firms. Indeed, the cash flow sensitivity of cash of public firms is significantly higher than that of private firms, and the difference is greatest when public and private firms are matched. The results show that the largest private firms, those of similar size as public firms, exhibit no significant cash flow sensitivity of cash.

Overall, these results are consistent with the notion that public firms of small and medium size face important financial constraints, due to the combination of two factors: They do not have an established reputation and so they need to persuade investors of their value, but they face a 'two-audiences' problem when communicating with them. On the other hand, when private firms of similar size need to raise capital, they can selectively communicate to investors all the necessary information to accurately value the firms; as a result, they do not need to hoard cash for a 'rainy day'.

5. Conclusions

My aim in this paper is to analyze why most medium and large-size firms in the U.S. choose to remain private, despite the fact that stock market-listed firms enjoy access to a deep pool of relatively cheap capital. I show that an important cost of being public is that listed firms are subject to the SEC's Regulation Fair Disclosure (Regulation FD), which prohibits the selective disclosure of material information by public firms. This induces a 'two-audiences' problem for public firms: while disclosing information reduces information asymmetry among investors and lowers a firm's cost of capital, the information may be exploited by product-market competitors.

My model predicts that, in order to limit the disclosure costs associated with raising capital, public firms, but not private ones, optimally hoard cash. I am able to empirically test this prediction using a rich new dataset on private U.S. firms (and standard data on public firms), and I indeed find that public firms hold more cash than private firms. This is a new stylized fact that is robust to several ways of addressing the endogeneity of the going-public decision, including matching, exploiting within-firm variation, and instrumental variables. The difference between the cash holdings of public and private firms is large, exceeding ten percentage points, and as predicted by my model, it increases in industry disclosure costs, returns on capital, and cash flow volatility.

However, holding cash is costly, and thus public firms cannot hold enough cash to completely eliminate disclosure costs. My model shows that, as a result, the value of being public decreases in industry disclosure costs and information asymmetry. On the other hand, firms in industries that require a large scale to operate efficiently benefit the most from having access to the low-cost capital available through the public equity markets. I test these predictions at the industry level, by examining the relative distribution of public firms across industries in the U.S. economy, and at the firm level. In both cases, I find strong support for my predictions.

My analysis shows that, by banning selective communications between firms and investors, regulations such as the SEC's Regulation FD impose disclosure costs on public firms. However, this finding should not be interpreted as implying that eliminating these regulations would be welfare enhancing. Indeed, Regulation FD helps create a level playing field for investors, and this is arguably an important reason why stock market investors are willing to supply their savings at relatively low cost. Determining the socially optimal degree of investor protection in the financial markets is an exciting avenue for future research.

References

Acharya, V.V., H. Almeida, and M. Campello, 2007, Is cash negative debt? A hedging perspective on corporate financial policies, Journal of Financial Intermediation 16, 515-554.

Acharya, V.V., S.A. Davydenko, and I.A. Strebulaev, 2008, Cash Holdings and Credit Risk, Unpublished working paper, New York University.

Ali, A., S. Klasa, and E. Yeung, 2010, Industry concentration and corporate disclosure policy, Unpublished working paper, University of Texas at Dallas.

Almeida, H., Campello, M., and M.S. Weisbach, 2004, The Cash Flow Sensitivity of Cash, Journal of Finance 49, 1777-1804.

Asker, J., J. Farre-Mensa, and A. Ljungqvist, 2010a, Does the Stock Market Harm Investment Incentives?, Unpublished working paper, New York University.

Asker, J., J. Farre-Mensa, and A. Ljungqvist, 2010b, What do Private Firms Look Like?, Unpublished working paper, New York University.

Asker, J., and A. Ljungqvist, 2010, Competition and the Structure of Vertical Relationships in Capital Markets, Journal of Political Economy 118, 599-647.

Bae, K.H., R. Stulz, and H. Tan, 2008, Do local analysts know more? A cross-country study of the performance of local analysts and foreign analysts, Journal of Financial Economics 88, 581-606.

Bamber, L.S., and Y.S. Cheon, 1998, Discretionary Management Earnings Forecast Disclosures: Antecedents and Outcomes Associated with Forecast Venue and Forecast Specificity Choices, Journal of Accounting Research 36, 167-190.

Bates, T.W., K.M. Kahle, and R.M. Stulz, 2009, Why Do U.S. Firms Hold So Much More Cash than They Used To?, Journal of Finance 64, 1985-2021.

Baum, C., M. Schaffer and S. Stillman, 2007, Enhanced routines for instrumental variables/GMM estimation and testing, Stata Journal 7, 465-506.

Bhattacharya, S., and G. Chiesa, 1995, Proprietary information, financial intermediation, and research incentives, Journal of Financial Intermediation 4, 328-357.

Bhattacharya, S., and J. Ritter, 1983, Innovation and Communication: Signalling with Partial Disclosure, Review of Economic Studies 50, 331-346.

Boot, A.W.A., R. Gopalan, and A.V. Thakor, 2006, The Entrepreneur's Choice between Private and Public Ownership, Journal of Finance 61, 803-836.

Botosan, C., 1997, Disclosure Level and the Cost of Equity Capital, Accounting Review 72, 323-349.

Brau, J.C., and S.E. Fawcett, 2006, Initial Public Offerings: An Analysis of Theory and Practice, Journal of Finance 61, 399-436.

Brav, O., 2009, Access to Capital, Capital Structure, and the Funding of the Firm, Journal of Finance 64, 263-308.

Caliendo, M., and S. Kopeinig, 2008, Some practical guidance for the implementation of propensity score matching, Journal of Economic Surveys 22, 31-72.

Campbell, T.S., 1979, Optimal Investment Financing Decisions and the Value of Confidentiality, Journal of Financial and Quantitative Analysis 14, 913-924.

Chemmanur, T.J., and J. He, 2009, IPO Waves, Product Market Competition, and the Going Public Decision: Theory and Evidence, Unpublished working paper.

Chemmanur, T.J., S. He, and D.K. Nandy, 2010, The Going-Decision and the Product Market, Review of Financial Studies 23, 1855-1908.

Chod, J., and E. Lyandres, 2010, Strategic IPOs and Product Market Competition, Journal of Financial Economics, forthcoming.

Coval, J., and T. Moskowitz, 1999, Home Bias at Home: Local Equity Preference in Domestic Portfolios, Journal of Finance 54, 2045-2073.

Easley, D., and M. O'Hara, 2004, Information and the Cost of Capital, Journal of Finance 59, 1553-1583.

Fan, J., 1992a, Design Adaptive Nonparametric Regression, Journal of the American Statistical Association 87, 998-1004.

Fan, J., 1992b, Local Linear Regression Smoothers and their Minimax Efficiencies, The Annals of Statistics 21, 196-216.

Foley, C.F., J.C. Hartzell, S. Titman, and G. Twite, 2007, Why do firms hold so much cash? A tax-based explanation, Journal of Financial Economics 86, 579–607.

Gabaix, X., 2010, The Granular Origins of Aggregate Fluctuations, Econometrica, forthcoming.

Gao, H., J. Harford, and K. Li, 2010, Determinants of Corporate Cash Policy: A Comparison of Private and Public Firms, Unpublished working paper, University of Washington.

Gibson, S., A. Safieddine, and R. Sonti, 2004, Smart investments by smart money: Evidence from seasoned equity offerings, Journal of Financial Economics 72, 581-604.

Gomes, A., and G. Phillips, 2007, Why Do Public Firms Issue Private and Public Securities?, Unpublished working paper, Washington University in St. Louis.

Gorton, G., and G. Pennachi, 1990, Financial Intermediaries and Liquidity Creation, Journal of Finance 45, 49-71.

Harford, J., S.A. Mansi, and W.F. Maxwell, 2008, Corporate governance and firm cash holdings in the US, Journal of Financial Economics 87, 535-555.

Harris, M.S., 1998, The Association between Competition and Managers' Business Segment Reporting Decisions, Journal of Accounting Research 36, 111-128.

Heckman, J., H. Ichimura, and P. Todd, 1997, Matching as an Econometric Evaluation Estimator: Evidence from Evaluation a Job Training Programme, Review of Economic Studies 64, 605-654.

Heckman, J., H. Ichimura, and P. Todd, 1998, Matching as an Econometric Evaluation Estimator, Review of Economic Studies 65, 261-294.

Hsu, H-C., A.V. Reed, and J. Rocholl, 2010, The New Game in Town: Competitive Effects of IPOs, Journal of Finance, 65, 495-528.

Huberman, G., 2001, Familiarity Breeds Investment, Review of Financial Studies 14, 659-680.

Imbens, G.W., and J.M. Wooldridge, 2009, Recent developments in the econometrics of program evaluation, Journal of Economic Literature 47, 5-86.

Ivkovic, Z., and S. Weisbenner, 2005, Local Does as Local Is: Information Content of the Geography of Individual Investors' Common Stock Investments, Journal of Finance 60, 267-306.

Jensen, M., and W.H. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, Journal of Financial Economics 3, 305-360.

Kleibergen, F., and R. Paap, 2006, Generalized reduced rank tests using the singular value decomposition, Journal of Econometrics 127, 97–126.

Korteweg, A.G., 2007, The Costs of Financial Distress Across Industries, Unpublished working paper, Stanford University.

Krishnaswami, S., and V. Subramaniam, 1999, Information asymmetry, valuation, and the corporate spin-off decision, Journal of Financial Economics 53, 73-112.

Lechner, M., 2002, Some practical issues in the evaluation of heterogenous labour market programmes by matching methods, Journal of the Royal Statistical Society A 165, 59-82.

Levin, R.C., A.K. Klevorich, R.R. Nelson, and S.G. Winter, 1987, Appropriating the Returns from Industrial Research and Development, Brookings Papers on Economic Activity 1987, 783-820.

Loughran, T., 2008, The impact of firm location on equity issuance, Financial Management 37, 1-21.

Maksimovic, V., and P. Pichler, 2001, Technological Innovation and Initial Public Offerings, Review of Financial Studies 14, 459-494.

Malloy, C., 2005, The Geography of Equity Analysis, Journal of Finance 60, 719-755.

Michaely, R., and M.R. Roberts, 2006, Corporate Dividend Policies: Lessons from Private Firms, Unpublished working paper, Cornell University.

Opler, T., L. Pinkowitz, R. Stulz, and R. Williamson, 1999, The determinants and implications of corporate cash holdings, Journal of Financial Economics 52, 3-46.

Pagano, M., F. Panetta, and L. Zingales, 1998, Why Do Companies Go Public? An Empirical Analysis, Journal of Finance 53, 27-64.

Pinedo, A.T., and Tanenbaum, J.R., 2009, Frequently Asked Questions About PIPEs. Available at http://www.sec.gov/info/smallbus/gbfor25_2006/pinedo_tanenbaum_pipefaq.pdf.

Rosenbaum, P., and D.B. Rubin, 1983, The Central Role of the Propensity Score in Observational Studies for Causal Effects, Biometrika 70, 41-55.

Salinger, M., R.E. Caves, and S. Peltzman, 1990, The Concentration-Margins Relationship Reconsidered, Brookings Papers on Economic Activity, Microeconomics 1990, 287-335.

Saunders, A., and S. Steffen, 2010, The Costs of Being Private: Evidence from the Loan Market, Unpublished working paper, New York University.

Silber, W.L., 1991, Discounts on Restricted Stock: The Impact of Illiquidity on Stock Prices, Financial Analysts Journal, 1991, 60-64.

Spiegel, M.I., and H. Tookes, 2009, Dynamic Competition, Innovation and Strategic Financing, Unpublished working paper, Yale University.

Staiger, D. and J.H. Stock, 1997, Instrumental Variables Regression with Weak Instruments, Econometrica 65, 557 – 586.

Sutton, J., 1991, Sunk Costs and Market Structure (Cambridge, MIT Press).

Wooldridge, J.M., 2001, Econometric Analysis of Cross Section and Panel Data (Cambridge, MIT Press).

Yosha, O., 1995, Information disclosure costs and the choice of financing source, Journal of Financial Intermediation 4, 3-20.

Appendix A

Proof of Lemma 1.

Let s_0^{Pub} denote the number of shares that a public firm needs to issue in period 0 to raise $E_0 = k + ck$.

If the firm carries out an SEO in period 1, then it has to issue $s_1^{SEO}(\rho)$ shares, where $s_1^{SEO}(\rho)$ satisfies

$$\frac{s_1^{SEO}(\rho)}{1 + s_0^{Pub} + s_1^{SEO}(\rho)}(1-\tau)(1-\gamma)zk = (\rho - c)k.$$

Note that I normalize the initial number of shares of the firm to one, and assume that this one share is retained by the entrepreneur. Therefore, it follows that

$$s_1^{SEO}(\rho) = \frac{(\rho - c)\left(1 + s_0^{Pub}\right)}{(1-\tau)(1-\gamma)z - (\rho - c)}.$$

In this case, in period 1 the entrepreneur knows that she will receive the following dividend from the firm

$$\frac{1}{1 + s_0^{Pub} + s_1^{SEO}(\rho)}(1-\tau)(1-\gamma)zk = \frac{1}{1 + s_0^{Pub}}\left((1-\tau)(1-\gamma)z - (\rho - c)\right)k. \tag{A1}$$

On the other hand, suppose that the firm carries out a PIPE in period 1 and that investors agree to fund it. Then it has to issue $s_1^{PIPE}(\rho)$ shares, where $s_1^{PIPE}(\rho)$ satisfies

$$\frac{s_1^{PIPE}(\rho)}{1 + s_0^{Pub} + s_1^{PIPE}(\rho)}(1-\mu)(1-\tau)zk = (\rho - c)k.$$

Hence, we have that

$$s_1^{PIPE}(\rho) = \frac{(\rho - c)\left(1 + s_0^{Pub}\right)}{(1-\mu)(1-\tau)z - (\rho - c)}.$$

In this case, the dividend that the entrepreneur will receive from the firm is

$$\frac{1}{1 + s_0^{Pub} + s_1^{PIPE}(\rho)}(1-\tau)zk = \frac{1}{1 + s_0^{Pub}}\left((1-\tau)z - \frac{1}{1-\mu}(\rho - c)\right)k. \tag{A2}$$

Therefore, an SEO is preferable to a PIPE if and only if

$$(1-\tau)(1-\gamma)z-(\rho-c)\geq(1-\tau)z-\frac{1}{1-\mu}(\rho-c)\Leftrightarrow\rho-c\geq\frac{1-\mu}{\mu}\gamma(1-\tau)z=S.$$

Finally, note that if $L^{Pub}=\rho$ and the firm carries out an SEO, it follows from equation (A1) that the dividend that the entrepreneur will receive from the firm is $ck/\left(1+s_0^{Pub}\right)$. Therefore, for $\rho>L^{Pub}$, it is optimal to liquidate the firm in period 1, thus allowing shareholders to recover the firm's precautionary cash holdings *ck* during the liquidation process.

Regarding **Footnote 20** in the main body of the text, note that Lemma 1 shows that the firm will choose to raise capital through a PIPE if and only if $\rho-c<\frac{1-\mu}{\mu}\gamma(1-\tau)z$. In such case, given the assumption that $\gamma<\mu$, it follows that

$$(\rho-c)k<\frac{1-\mu}{\mu}\gamma(1-\tau)zk<(1-\mu)(1-\tau)zk.$$

That is, the maximum dividend that PIPE investors can expect to receive from a firm (provided that they are offered enough shares) is greater than the maximum amount of capital a firm will be willing to raise through a PIPE. ■

Proof of Proposition 1.

Let $V^{Public}(c)$ denote the entrepreneur's period 0 valuation of a public firm that follows the optimal financing policy in period 1 (described in Lemma 1), assuming that the firm raises *ck* units of precautionary cash in period 0. Assuming that period 1 PIPE investors are willing to fund the firm regardless of *c*, it follows from equations (A1) and (A2) that $V^{Public}(c)$ can be written as

$$V^{Public}(c)=\frac{1}{1+s_0^{Pub}}k\left\{\begin{array}{l}(1-\tau)\int_0^c(z+c-\rho)f(\rho)d\rho+\int_c^{c+S}\left((1-\tau)z-\frac{1}{1-\mu}(\rho-c)\right)f(\rho)d\rho+\\+\int_{c+S}^{L^{Pub}}\left((1-\tau)(1-\gamma)z-(\rho-c)\right)f(\rho)d\rho+\int_{L^{Pub}}^{\infty}c\,f(\rho)d\rho\end{array}\right\}\doteq$$

$$\doteq\frac{1}{1+s_0^{Pub}}R^{Public}(c).$$

Note that s_0^{Pub} is characterized by

$$\frac{s_0^{Pub}}{1+s_0^{Pub}} R^{Public}\left(c\right) = k + ck \; .$$

And thus we have that

$$V^{Public}\left(c\right) = \frac{1}{1+s_0^{Pub}} R^{Public}\left(c\right) = \left(1 - \frac{s_0^{Pub}}{1+s_0^{Pub}}\right) R^{Public}\left(c\right) = R^{Public}\left(c\right) - \left(k+ck\right). \tag{A3}$$

In such case, the value of c that maximizes $V^{Public}\left(c\right)$ needs to verify the following expression:

$$\frac{\partial V^{Public}\left(c\right)}{\partial c} = \left(1-\tau\right)\int_0^c k\, f\left(\rho\right) d\rho + \left(1-\tau\right) zkf\left(c\right) - \left(1-\tau\right) zkf\left(c\right) +$$

$$+\int_c^{c+S} \frac{1}{1-\mu} k\, f\left(\rho\right) d\rho + \int_{c+S}^{L^{Pub}} k\, f\left(\rho\right) d\rho + \int_{L^{Pub}}^{\infty} k\, f\left(\rho\right) d\rho - k +$$

$$+\left(\left(1-\tau\right) z - \frac{1}{\mu}\gamma\left(1-\tau\right) z\right) kf\left(c+S\right) - \left(\left(1-\tau\right)\left(1-\gamma\right) z - \frac{1-\mu}{\mu}\gamma\left(1-\tau\right) z\right) kf\left(c+S\right) =$$

$$= \left(1-\tau\right)\int_0^c k\, f\left(\rho\right) d\rho + \frac{1}{1-\mu}\int_c^{c+S} k\, f\left(\rho\right) d\rho + \int_{c+S}^{L^{Pub}} k\, f\left(\rho\right) d\rho + \int_{L^{Pub}}^{\infty} k\, f\left(\rho\right) d\rho - k =$$

$$= -\tau\int_0^c k\, f\left(\rho\right) d\rho + \frac{\mu}{1-\mu}\int_c^{c+S} k\, f\left(\rho\right) d\rho = 0 \Leftrightarrow \tau\int_0^c f\left(\rho\right) d\rho = \frac{\mu}{1-\mu}\int_c^{c+S} f\left(\rho\right) d\rho$$

This condition is necessary and sufficient, given that $f\left(\rho\right)$ is non-increasing and thus $V^{Public}\left(c\right)$ is concave. Indeed, note that

$$\frac{\partial^2 V^{Public}\left(c\right)}{\partial c^2} = -\tau k\, f\left(c\right) + \frac{\mu}{1-\mu} k\left(f\left(c+S\right) - f\left(c\right)\right) < 0 \; .$$

In the model's Perfect Bayesian Equilibrium, period 1 investors' beliefs are such that if a firm has raised $c \neq \hat{c}$ in period 0 and is attempting to raise capital in period 1 through a PIPE, they believe that the firm is a failed firm with probability one, and therefore refuse to fund it. Therefore, if $c \neq \hat{c}$, a public firm's only financing option in period 1 is an SEO. In the extended version of the model, I show that the entrepreneur's equilibrium valuation of a public firm, $V_E^{Public}\left(c\right)$, is such that $V_E^{Public}\left(\hat{c}\right) = V^{Public}\left(\hat{c}\right)$. In addition, I show that the fact that $c = \hat{c}$ maximizes $V^{Public}\left(c\right)$ implies that it also maximizes $V_E^{Public}\left(c\right)$.

On the other hand, the entrepreneur's valuation of a firm that stays private is

$$V^{Private}(c)=\frac{1}{1+s_0^{Pri}}\left\{\begin{array}{l}\int_0^c(1-\tau)\left(\left(1-\lambda\delta_1-\lambda\delta_2(1-\tau)zk\right)z+c-\rho\right)k\ f(\rho)d\rho+\\+\int_c^{L^{Pri}}\left(\left(1-\lambda\delta_1-\lambda\delta_2(1-\tau)zk\right)(1-\tau)z+c-\rho\right)k\ f(\rho)d\rho+\\+\int_{L^{Pri}}^{\infty}ck\ f(\rho)d\rho\end{array}\right\}\doteq$$

$$\doteq\frac{1}{1+s_0^{Pri}}R^{Private}(c),$$

where I use that

$$s_1^{Pri}(\rho)=\frac{(\rho-c)\left(1+s_0^{Pri}\right)}{(1-\tau)\left(1-\lambda\delta_1-\lambda\delta_2(1-\tau)zk\right)z-(\rho-c)}.$$

In this case, s_0^{Pri} is characterized by

$$\frac{s_0^{Pri}}{1+s_0^{Pri}}R^{Private}(c)=k+ck.$$

This implies that

$$V^{Private}(c)=\frac{1}{1+s_0^{Pri}}R^{Private}(c)=\left(1-\frac{s_0^{Pri}}{1+s_0^{Pri}}\right)R^{Private}(c)=R^{Private}(c)-(k+ck). \quad \text{(A4)}$$

Therefore, if $c>0$, it follows that

$$\frac{\partial V^{Private}(c)}{\partial c}=\int_0^c(1-\tau)k\ f(\rho)d\rho+(1-\tau)\left(1-\lambda\delta_1-\lambda\delta_2(1-\tau)zk\right)zkf(c)-$$

$$-\left(1-\lambda\delta_1-\lambda\delta_2(1-\tau)zK\right)(1-\tau)zkf(c)+\int_c^{L^{Pri}}k\,f(\rho)d\rho+\int_{L^{Pri}}^{\infty}k\,f(\rho)d\rho-k=$$

$$=k\left((1-\tau)\int_0^c f(\rho)d\rho+\int_c^{L^{Pri}}f(\rho)d\rho+\int_{L^{Pri}}^{\infty}f(\rho)d\rho-1\right)<0.$$

This proves that it is optimal for a private firm to hold no precautionary cash as long as $\tau>0$. ■

Proof of Proposition 2.

The optimal precautionary cash policy for a public firm is implicitly defined by equation (1), which can be written as follows

$$F \doteq \tau\int_0^c f(\rho)\,d\rho - \frac{\mu}{1-\mu}\int_c^{c+S} f(\rho)\,d\rho = 0.$$

Applying the implicit function theorem, we have that

$$\frac{\partial \hat{c}}{\partial x} = -\left[\left(\frac{\partial F}{\partial c}\right)^{-1}\frac{\partial F}{\partial x}\right]\Big|_{c=\hat{c}}, \text{ where } x \in \{\gamma, z\}.$$

Note that

$$\frac{\partial F}{\partial c}\Big|_{c=\hat{c}} = \tau f(\hat{c}) - \frac{\mu}{1-\mu}\left(f(\hat{c}+S) - f(\hat{c})\right) > 0,$$

given that $f(\rho)$ is non-increasing.

Thus, Proposition 2 follows from the following results:

a) $\frac{\partial F}{\partial \gamma}\Big|_{c=\hat{c}} = -f(\hat{c}+S)(1-\tau)z < 0$

b) $\frac{\partial F}{\partial z}\Big|_{c=\hat{c}} = -f(\hat{c}+S)\gamma(1-\tau) < 0$

c) Under the assumptions that ρ follows a normal distribution $N(0,\sigma^2)$ truncated at 0 and $z = \tilde{z} + E[\rho]$, equation (4) can be written as:

$$F = \sqrt{\frac{2}{\pi\sigma^2}}\left(\tau\int_0^c e^{-\frac{\rho^2}{2\sigma^2}}\,d\rho - \frac{\mu}{1-\mu}\int_c^{c+S} e^{-\frac{\rho^2}{2\sigma^2}}\,d\rho\right) = 0, \text{ where } S = \frac{1-\mu}{\mu}\gamma(1-\tau)\left(\tilde{z} + E[\rho]\right)$$

Therefore, we have that

$$\frac{\partial F}{\partial \sigma^2}\Big|_{c=\hat{c}} = \sqrt{\frac{2}{\pi\sigma^2}}\left(\tau\int_0^{\hat{c}} e^{-\frac{\rho^2}{2\sigma^2}}\frac{\rho^2}{2\sigma^4}\,d\rho - \frac{\mu}{1-\mu}\left(\int_{\hat{c}}^{\hat{c}+S} e^{-\frac{\rho^2}{2\sigma^2}}\frac{\rho^2}{2\sigma^4}\,d\rho + e^{-\frac{(\hat{c}+S)^2}{2\sigma^2}}\frac{1-\mu}{\mu}\gamma(1-\tau)\sqrt{\frac{1}{2\pi\sigma^2}}\right)\right) <$$

$$< \sqrt{\frac{2}{\pi\sigma^2}}\left(\tau\int_0^{\hat{c}} e^{-\frac{\rho^2}{2\sigma^2}}\frac{\rho^2}{2\sigma^4}\,d\rho - \frac{\mu}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} e^{-\frac{\rho^2}{2\sigma^2}}\frac{\rho^2}{2\sigma^4}\,d\rho\right) <$$

$$< \sqrt{\frac{2}{\pi\sigma^2}}\frac{\hat{c}^2}{2\sigma^4}\left(\tau\int_0^{\hat{c}} e^{-\frac{\rho^2}{2\sigma^2}}\,d\rho - \frac{\mu}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} e^{-\frac{\rho^2}{2\sigma^2}}\,d\rho\right) = 0$$

■

Proof of Proposition 3.

Substituting the optimal precautionary cash policy $\hat{c}$ into equation (A3), it follows that the entrepreneur's return on investment of a firm that goes public in period 0 is

$$r^{Public} = \frac{V^{Public}(\hat{c})}{k} = (1-\tau)z\left(\int_0^{\hat{c}+S} f(\rho)d\rho + (1-\gamma)\int_{\hat{c}+S}^{L^{Pub}} f(\rho)d\rho\right) +$$

$$+\hat{c}\left((1-\tau)\int_0^{\hat{c}} f(\rho)d\rho + \frac{1}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} f(\rho)d\rho + \int_{\hat{c}+S}^{\infty} f(\rho)d\rho\right) -$$

$$-\left((1-\tau)\int_0^{\hat{c}} \rho f(\rho)d\rho + \frac{1}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} \rho f(\rho)d\rho + \int_{\hat{c}+S}^{L^{Pub}} \rho f(\rho)d\rho\right) - (1+\hat{c}) =$$

$$= (1-\tau)z\left(\int_0^{\hat{c}+S} f(\rho)d\rho + (1-\gamma)\int_{\hat{c}+S}^{L^{Pub}} f(\rho)d\rho\right) +$$

$$+\left(\tau\int_0^{\hat{c}} \rho f(\rho)d\rho - \frac{\mu}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} \rho f(\rho)d\rho\right) - \int_0^{L^{Pub}} \rho f(\rho)d\rho - 1.$$

Analogously, substituting the optimal zero precautionary cash policy into equation (A4), it follows that the entrepreneur's value of a firm that stays private is

$$r^{Private} = \frac{V^{Private}(\hat{c})}{k} = \int_0^{L^{Pri}} \left(\left(1-\lambda\delta_1 - \lambda\delta_2(1-\tau)zk\right)(1-\tau)z - \rho\right) f(\rho)d\rho - 1.$$

Therefore, the difference reads

$$r^{Public} - r^{Private} = (1-\tau)z\left(\int_0^{\hat{c}+S} f(\rho)\,d\rho + (1-\gamma)\int_{\hat{c}+S}^{L^{Pub}} f(\rho)\,d\rho\right) +$$

$$+\left(\tau\int_0^{\hat{c}} \rho f(\rho)\,d\rho - \frac{\mu}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} \rho f(\rho)\,d\rho\right) - \int_0^{L^{Pub}} \rho f(\rho)\,d\rho -$$

$$-\int_0^{L^{Pri}} \Big(\big(1-\lambda\delta_1 - \lambda\delta_2(1-\tau)zk\big)(1-\tau)z - \rho\Big) f(\rho)\,d\rho =$$

$$= (1-\tau)z\left(\big(\lambda\delta_1 + \lambda\delta_2(1-\tau)zk\big)\int_0^{\hat{c}+S} f(\rho)\,d\rho + \big(\big(\lambda\delta_1 + \lambda\delta_2(1-\tau)zk\big) - \gamma\big)\int_{\hat{c}+S}^{\min\{L^{Pub},L^{Pri}\}} f(\rho)\,d\rho\right) +$$

$$+\left(\tau\int_0^{\hat{c}} \rho f(\rho)\,d\rho - \frac{\mu}{1-\mu}\int_{\hat{c}}^{\hat{c}+S} \rho f(\rho)\,d\rho\right) +$$

$$+\left(\begin{array}{l} \int_{\min\{L^{Pub},L^{Pri}\}}^{L^{Pub}} \big((1-\tau)z(1-\gamma) - \rho\big) f(\rho)\,d\rho\, 1_{\{L^{Pub}\geq L^{Pri}\}} - \\ -\int_{\min\{L^{Pub},L^{Pri}\}}^{L^{Pri}} \Big(\big(1-\lambda\delta_1 - \lambda\delta_2(1-\tau)zk\big)(1-\tau)z - \rho\Big) f(\rho)\,d\rho\, 1_{\{L^{Pub}<L^{Pri}\}} \end{array}\right).$$

The three parts of Proposition 3 follow from the previous expression:

a) Given that $r^{Private}$ does not depend on γ, it suffices to show that r^{Public} is decreasing in γ:

$$\frac{\partial\, r^{Public}}{\partial\gamma} = (1-\tau)z\left(\begin{array}{l} f(\hat{c}+S)\dfrac{\partial(\hat{c}+S)}{\partial\gamma} - \int_{\hat{c}+S}^{L^{Pub}} f(\rho)\,d\rho + \\ +(1-\gamma)\left(f(L^{Pub})\dfrac{\partial L^{Pub}}{\partial\gamma} - f(\hat{c}+S)\dfrac{\partial(\hat{c}+S)}{\partial\gamma}\right) \end{array}\right) +$$

$$+\tau\hat{c}f(\hat{c})\frac{\partial\hat{c}}{\partial\gamma} - \frac{\mu}{1-\mu}\left((\hat{c}+S)f(\hat{c}+S)\frac{\partial(\hat{c}+S)}{\partial\gamma} - \hat{c}f(\hat{c})\frac{\partial\hat{c}}{\partial\gamma}\right) - L^{Pub}f(L^{Pub})\frac{\partial L^{Pub}}{\partial\gamma} =$$

$$= -(1-\tau)z\int_{\hat{c}+S}^{L^{Pub}} f(\rho)\,d\rho + k\big((1-\tau)z(1-\gamma) - L^{Pub}\big) f(L^{Pub})\frac{\partial L^{Pub}}{\partial\gamma} +$$

$$+\left((1-\tau)z\gamma - \frac{\mu}{1-\mu}(\hat{c}+S)\right) f(\hat{c}+S)\frac{\partial(\hat{c}+S)}{\partial\gamma} + \left(\tau + \frac{\mu}{1-\mu}\right)\hat{c}f(\hat{c})\frac{\partial\hat{c}}{\partial\gamma} =$$

$$= \hat{c}\left(\left(\tau + \frac{\mu}{1-\mu}\right)f(\hat{c}) - \frac{\mu}{1-\mu}f(\hat{c}+S)\right)\left(\tau f(\hat{c}) - \frac{\mu}{1-\mu}\big(f(\hat{c}+S) - f(\hat{c})\big)\right)^{-1} f(\hat{c}+S)(1-\tau)z -$$

$$-(1-\tau)z\int_{\hat{c}+S}^{L^{Pub}} f(\rho)\,d\rho - \hat{c}f(\hat{c}+S)(1-\tau)z = -(1-\tau)z\int_{\hat{c}+S}^{L^{Pub}} f(\rho)\,d\rho < 0$$

b) As before, given that $r^{Private}$ does not depend on μ, it suffices to show that r^{Public} is decreasing in μ. Note first that, defining *F* as in the proof of Proposition 2, it follows that

$$\frac{\partial F}{\partial \mu}|_{c=\hat{c}} = -\frac{1}{(1-\mu)^2}\int_{\hat{c}}^{\hat{c}+S} f(\rho)\,d\rho + \frac{1}{\mu(1-\mu)} f(\hat{c}+S)\gamma(1-\tau)z\,.$$

Therefore, we have that

$$\frac{\partial\, r^{Public}}{\partial \mu} = (1-\tau)z\left(f(\hat{c}+S)\frac{\partial(\hat{c}+S)}{\partial \mu} - (1-\gamma) f(\hat{c}+S)\frac{\partial(\hat{c}+S)}{\partial \mu}\right) +$$

$$+\tau\hat{c} f(\hat{c})\frac{\partial \hat{c}}{\partial \mu} - \frac{1}{(1-\mu)^2}\int_{\hat{c}}^{\hat{c}+S}\rho f(\rho)\,d\rho - \frac{\mu}{1-\mu}\left((\hat{c}+S) f(\hat{c}+S)\frac{\partial(\hat{c}+S)}{\partial \mu} - \hat{c} f(\hat{c})\frac{\partial \hat{c}}{\partial \mu}\right) =$$

$$= -\frac{1}{(1-\mu)^2}\int_{\hat{c}}^{\hat{c}+S}\rho f(\rho)\,d\rho + \left((1-\tau)z\gamma f(\hat{c}+S) - \frac{\mu}{1-\mu}(\hat{c}+S) f(\hat{c}+S)\right)\frac{\partial S}{\partial \mu} -$$

$$+\hat{c}\left(\tau f(\hat{c}) - \frac{\mu}{1-\mu}\big(f(\hat{c}+S) - f(\hat{c})\big)\right)\frac{\partial\,\hat{c}}{\partial \mu} =$$

$$= -\frac{1}{(1-\mu)^2}\int_{\hat{c}}^{\hat{c}+S}\rho f(\rho)\,d\rho + \frac{\mu}{1-\mu}\hat{c} f(\hat{c}+S)\frac{1}{\mu^2}\gamma(1-\tau)z +$$

$$-\hat{c}\left(-\frac{1}{(1-\mu)^2}\int_{\hat{c}}^{\hat{c}+S} f(\rho)\,d\rho + \frac{1}{\mu(1-\mu)} f(\hat{c}+S)\gamma(1-\tau)z\right) =$$

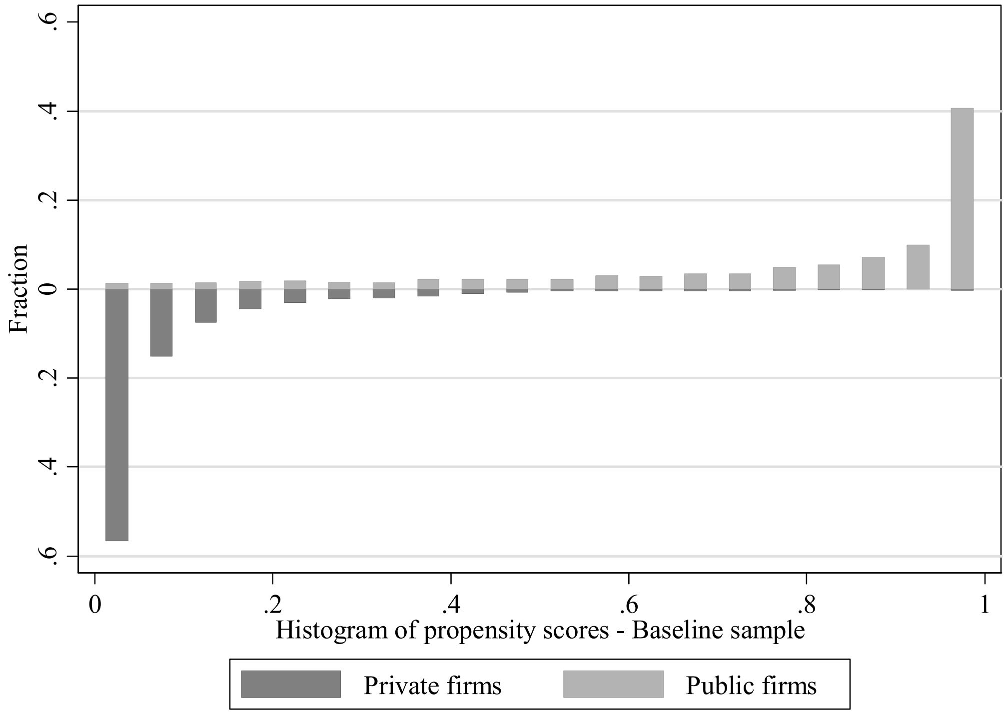
$$= -\frac{1}{(1-\mu)^2}\left(\int_{\hat{c}}^{\hat{c}+S}\rho f(\rho)\,d\rho - \hat{c}\int_{\hat{c}}^{\hat{c}+S} f(\rho)\,d\rho\right) < -\frac{\hat{c}}{(1-\mu)^2}\left(\int_{\hat{c}}^{\hat{c}+S} f(\rho)\,d\rho - \int_{\hat{c}}^{\hat{c}+S} f(\rho)\,d\rho\right) = 0.$$

c) In this case, given that r^{Public} does not depend on *k*, it suffices to show that $r^{Private}$ is decreasing in *k*:

$$\frac{\partial\, r^{Private}}{\partial k} = -\lambda\delta_2\left((1-\tau)z\right)^2\int_0^{L^{Pri}} f(\rho)\,d\rho < 0$$

■

Figure 3. Propensity Scores of Public and Private Firms: Baseline and Matched Samples.
The top graph shows the distribution of propensity scores for private firms (below the abscissa) and public firms (above the abscissa) in my baseline sample of firms with average sales of at least $5 million (specifically, the figure focuses on the sample for which industry concentration is available as a measure of disclosure costs). The propensity scores are based on log sales, cash flow, sales growth, and industry. Similarly, the bottom graph shows the distribution of propensity scores after matching public and private firms using local linear propensity score matching. The width of the histogram bins is set to 0.05.



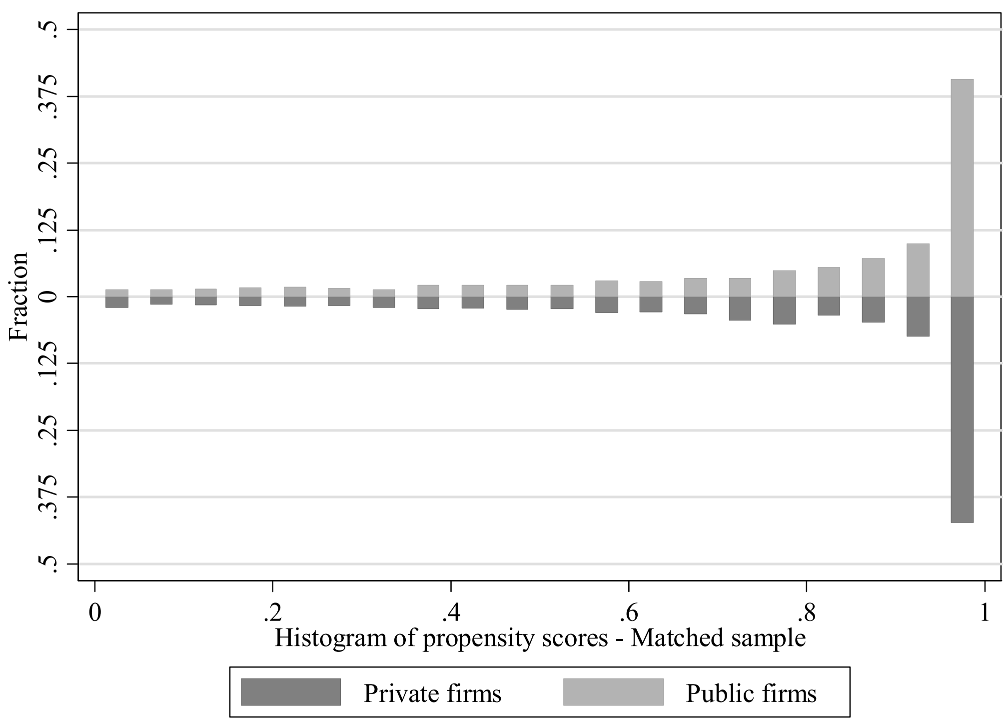
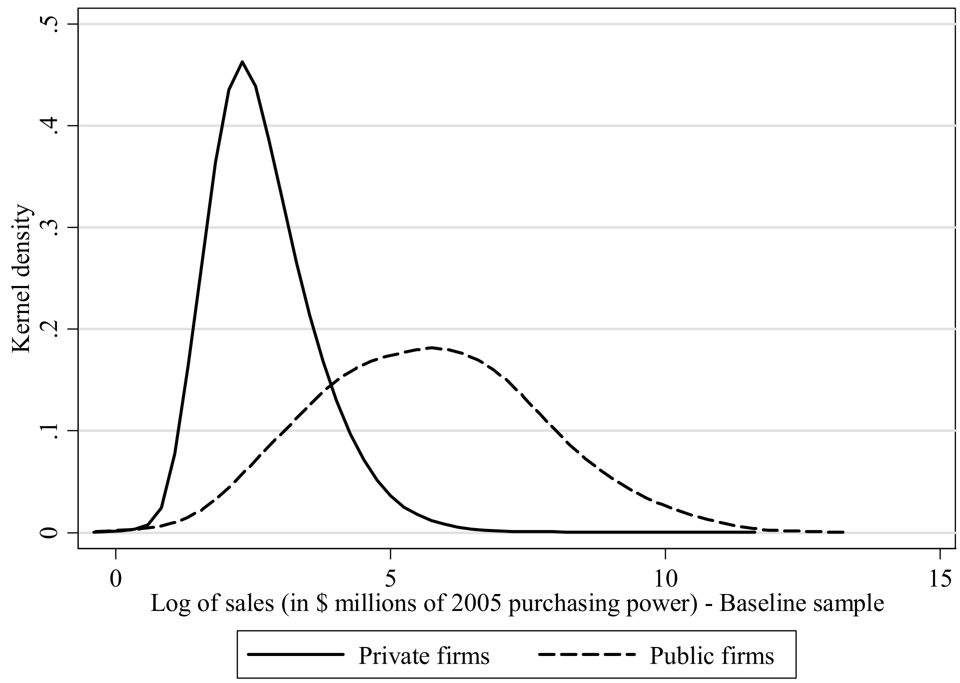


Figure 4. Size Distribution of Public and Private firms: Baseline and Matched Samples.
The top graph shows the size distribution of the public and private firms in my baseline sample of firms with average sales of at least $5 million. The bottom graph shows the size distribution of the public and private firms after matching public and private firms by industry and sales using nearest-neighbor matching (specifically, the graph focuses on the sample for which industry concentration is available as a measure of disclosure costs). The graphs present, for each set of firms, Epanechnikov kernel densities of the natural logarithm of sales in $ millions of 2005 purchasing power. The width of the kernel density window around each point is set to 0.4.



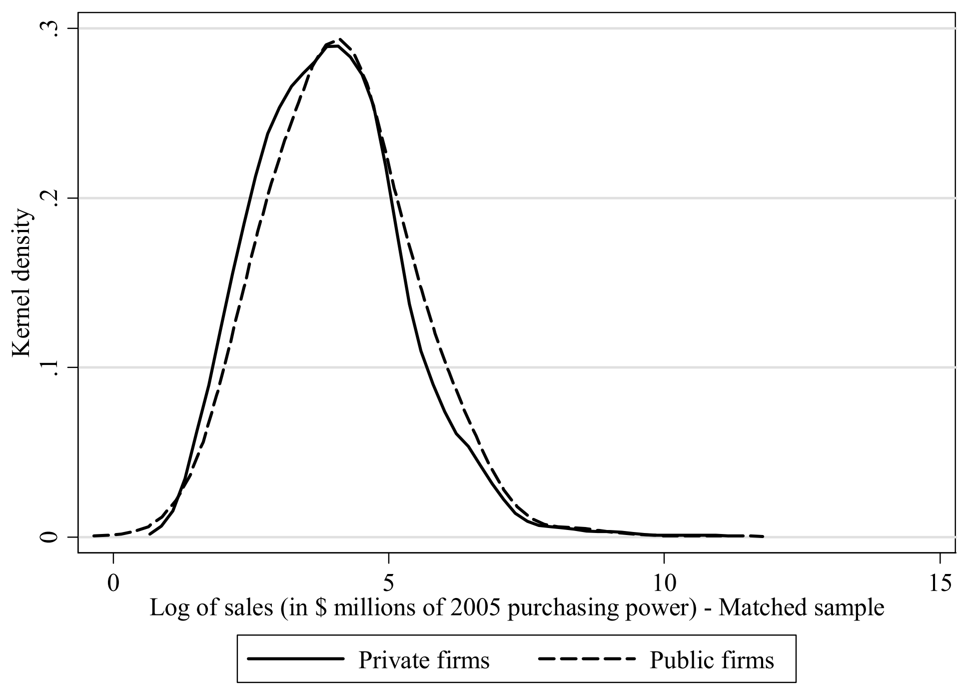


Table 1. Descriptive Statistics.

This table presents descriptive statistics for the public and private firms in my baseline sample of U.S. firms with average annual real sales of at least $5 million during the sample period (2002-2008). The main data source for public firms is the CRSP-Compustat database, while for private firms it is Sageworks. The table reports means, medians, and standard deviations of the key variables used in my empirical analysis. Throughout, I winsorize cash over assets and leverage so that they are between zero and one, while all other variables are 1% winsorized (winsorization follows Bates et al. (2009)). Note that the actual sample size used when computing these descriptive statistics depends on data availability, and thus it is not the same for all variables. For variable definitions, see the main text.

	Public firms			Private firms		
	mean	median	std. dev.	mean	median	std. dev.
Industry disclosure cost						
Patenting intensity (absolute value)	0.778	0.656	0.678	0.375	0.000	0.573
Information leakage (absolute value)	0.774	0.693	0.696	0.665	0.693	0.694
Industry concentration (*C4*)	0.317	0.326	0.159	0.214	0.177	0.159
Industry concentration (HHI)	0.050	0.050	0.040	0.030	0.017	0.037
Industry information asymmetry						
Analyst forecast error	0.007	0.003	0.033	0.011	0.003	0.039
Analyst forecast dispersion	0.030	0.020	0.046	0.041	0.020	0.075
Industry minimum efficient scale						
Median log prop., plant, and equip.	2.441	2.090	1.627	1.014	0.776	0.915
Other industry characteristics						
Industry return on capital	1.664	1.523	0.666	1.349	1.206	0.641
Industry cash flow volatility	0.089	0.095	0.038	0.058	0.048	0.029
Industry R&D	0.252	0.027	0.613	0.029	0.000	0.168
Industry acquisitions	0.026	0.021	0.023	0.032	0.022	0.038
Cash holdings						
Cash / assets	0.219	0.131	0.228	0.115	0.057	0.147
Cash / sales	0.369	0.125	0.571	0.052	0.020	0.110
Cash / net assets	0.483	0.151	0.794	0.182	0.060	0.371
Firm characteristics						
Total assets ($m)	1,510.978	312.662	2,912.927	19.841	5.092	283.877
Total sales ($m)	1,411.775	282.138	2,688.003	33.655	12.266	266.640
Cash flow	0.017	0.065	0.199	-0.013	0.048	0.506
Negative cash flow dummy	0.243	0.000	0.429	0.353	0.000	0.478
Sales growth	0.152	0.085	0.440	0.150	0.078	0.421
Investment	0.028	0.020	0.101	0.036	0.016	0.103
Working capital	0.280	0.254	0.246	0.238	0.229	0.298
Leverage (book)	0.203	0.150	0.219	0.238	0.142	0.262
Dividend dummy	0.270	0.000	0.444	0.126	0.000	0.332
No. observations	21,767			52,815		
No. firms	4,322			18,476		

Table 2. The effect of industry characteristics on a firm's listing status.

This table examines the effect that industry characteristics have on a firm's public or private status, estimating equation (3) in the text using a probit model. Each regression includes an intercept and year effects (not reported for brevity). Heteroskedasticity-robust standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Dependent variable: Public = 1, Private = 0 (Probit model)			
	(1)	(2)	(3)	(4)
Industry disclosure cost				
Patenting intensity	-0.321***	-0.301***		
	0.020	*0.019*		
Information leakage			-0.128***	-0.195***
			0.023	*0.024*
Industry information asymmetry				
Analyst forecast error	-0.875***		-1.906***	
	0.250		*0.293*	
Analyst forecast dispersion		-2.176***		-3.057***
		0.405		*0.492*
Industry minimum efficient scale				
Median log property, plant, and equipment	0.437***	0.437***	0.419***	0.417***
	0.010	*0.010*	*0.015*	*0.014*
Industry R&D	0.775***	0.763***		
	0.048	*0.046*		
Pseudo R^2	27.6%	27.3%	20.1%	20.8%
Wald test: all coefficients = 0 (χ^2)	9,226.7***	8,883.9***	3,868.6***	3,984.4***
No. observations	74,582	71,442	35,132	34,950
No. firms	22,798	22,111	10,160	10,135

Table 3. The effect of industry characteristics on the share of public firms in an industry.

This table examines the effect that industry characteristics have on the share of public firms in an industry, estimating equation (4) in the text by OLS. Each observation is a five-digit NAICS industry. I exclude financial services, utilities, and the government sector. All regressions include an intercept (not reported for brevity). Heteroskedasticity-robust standard errors are shown in italics underneath the coefficient estimates. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

Dependent variable: log (fraction of public …)	**sales**	**firms w. 100+ employees**	**firms w. 500+ employees**	**sales**	**firms w. 100+ employees**	**firms w. 500+ employees**
	(1)	(2)	(3)	(4)	(5)	(6)
Industry disclosure cost						
Patenting intensity	-0.740***	-0.250***	-0.173***			
	0.104	*0.069*	*0.065*			
Information leakage				-0.205	-0.284	-0.305*
				0.261	*0.196*	*0.184*
Industry information asymmetry						
Analyst forecast error	-2.884***	-1.518	-1.687	-3.028***	-2.777***	-2.688***
	0.834	*1.209*	*1.087*	*1.044*	*0.535*	*0.546*
Industry minimum efficient scale						
Median log property, plant, and equipment	0.266***	0.264***	0.187***	0.204***	0.263***	0.186***
	0.035	*0.026*	*0.024*	*0.069*	*0.042*	*0.039*
Industry R&D	0.382**	0.866***	0.769***			
	0.156	*0.154*	*0.147*			
Adjusted R^2	38.6%	31.9%	21.5%	9.4%	23.1%	15.8%
Wald test: all coefficients = 0 (F)	42.1***	40.8***	26.2***	7.8***	29.4***	21.5***
No. observations	287	301	301	105	110	110

Table 4. Comparing the cash policies of public and private firms: Baseline results.
This table compares the cash policies of public and private firms, estimating equation (5) in the text. Each regression includes an intercept and year effects (not reported for brevity). Heteroskedasticity-robust standard errors clustered at the firm level are shown in italics beside the coefficient estimates. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Dependent variable: cash / assets							
	(1)		(2)		(3)		(4)	
	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*
Public	0.125***	*0.006*	-0.010	*0.010*	0.133***	*0.007*	0.018	*0.013*
Industry disclosure cost								
industry concentration	0.017**	*0.007*	-0.023***	*0.008*				
information leakage					0.004	*0.003*	-0.007***	*0.003*
… x public			0.093***	*0.016*			0.031***	*0.006*
Industry return on capital	0.017***	*0.002*	0.009***	*0.002*	0.032***	*0.004*	0.017***	*0.004*
… x public			0.031***	*0.004*			0.038***	*0.007*
Industry cash flow volatility	0.800***	*0.042*	0.464***	*0.048*	0.518***	*0.058*	0.204***	*0.060*
… x public			0.642***	*0.088*			0.704***	*0.115*
Cash flow	0.002	*0.004*	0.012***	*0.004*	-0.008	*0.005*	0.014**	*0.006*
… x public			-0.177***	*0.015*			-0.221***	*0.018*
Industry R&D	0.078***	*0.006*	-0.002	*0.008*	0.076***	*0.007*	0.001	*0.010*
… x public			0.057***	*0.010*			0.056***	*0.013*
Size	-0.019***	*0.001*	-0.014***	*0.001*	-0.019***	*0.001*	-0.012***	*0.001*
Working capital (net of cash)	-0.167***	*0.006*	-0.156***	*0.005*	-0.171***	*0.008*	-0.157***	*0.007*
Investment	-0.107***	*0.009*	-0.078***	*0.008*	-0.150***	*0.012*	-0.107***	*0.011*
Leverage	-0.189***	*0.005*	-0.193***	*0.005*	-0.199***	*0.008*	-0.202***	*0.008*
Dividend dummy	-0.023***	*0.003*	-0.015***	*0.003*	-0.029***	*0.004*	-0.019***	*0.004*
Industry acquisitions	0.007	*0.021*	0.004	*0.021*	-0.233***	*0.048*	-0.137***	*0.047*
Industry information asymmetry	-0.040**	*0.018*	-0.049***	*0.018*	0.040	*0.025*	-0.006	*0.022*
Adjusted R^2	33.7%		36.1%		36.3%		39.4%	
Wald test: all coeff. = 0 (F)	244.8***		217.8***		148.0***		140.3***	
No. observations	48,273		48,273		26,509		26,509	
No. firms	17,065		17,065		8,719		8,719	

Table 5. Cash comparison: Alternative specifications.

In this table, I analyze the robustness of the results presented in Table 4 to several alternative specifications. In Panel A, I add new control variables to those included in columns 2 and 4 of Table 4: a negative cash flow dummy (columns 1 and 2) and industry minimum efficient scale (columns 3 and 4). In Panel B, I include the same control variables as in columns 2 and 4 of Table 4, but I either restrict the sample (columns 1 through 4) or define some variable differently (columns 5 through 10). Specifically, in column 5 I use the Herfindahl-Hirschman index (HHI) for the 50 largest firms to measure industry concentration; given that the HHI is only available for manufacturing firms, in column 6 I use *C4*, *C8*, *C20*, and *C50* (available for all industries) to compute a simplified version of the HHI: $\text{HHI cons.} = 4\left(C4/4\right)^2 + 4\left(\left(C8-C4\right)/4\right)^2 + 12\left(\left(C20-C8\right)/12\right)^2 + 30\left(\left(C50-C20\right)/30\right)^2$. For brevity, in Panel B I only report the estimates that directly test the predictions of my model. The last row in both panels reports the estimated public dummy in the same regression shown in the respective column but without any interaction effects (the estimates of the other variables are not reported to conserve space). Each regression includes an intercept and year effects (also not reported for brevity). Heteroskedasticity-robust standard errors clustered at the firm level are shown in italics either beside (Panel A) or below (Panel B) the coefficient estimates. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

Panel A.

	Dependent variable: cash / assets							
	(1)		(2)		(3)		(4)	
	coeff.	*s.e*	coeff.	*s.e*	coeff.	*s.e*	coeff.	*s.e*
Public	-0.033***	*0.010*	-0.009	*0.013*	-0.010	*0.010*	0.016	*0.013*
Industry disclosure cost								
industry concentration	-0.024***	*0.008*			-0.024***	*0.008*		
information leakage			-0.007***	*0.003*			-0.007***	*0.003*
… x public	0.089***	*0.016*	0.032***	*0.006*	0.091***	*0.016*	0.032***	*0.006*
Industry return on capital	0.009***	*0.002*	0.016***	*0.004*	0.009***	*0.002*	0.017***	*0.004*
… x public	0.030***	*0.004*	0.037***	*0.007*	0.031***	*0.004*	0.038***	*0.007*
Industry cash flow volatility	0.465***	*0.048*	0.214***	*0.060*	0.464***	*0.048*	0.199***	*0.060*
… x public	0.605***	*0.087*	0.640***	*0.114*	0.645***	*0.088*	0.710***	*0.116*
Cash flow	0.007	*0.005*	0.008	*0.008*	0.011***	*0.004*	0.014**	*0.006*
… x public	-0.068***	*0.016*	-0.101***	*0.020*	-0.177***	*0.015*	-0.221***	*0.018*
Negative cash flow dummy	-0.011***	*0.003*	-0.011**	*0.005*				
… x public	0.086***	*0.008*	0.098***	*0.010*				
Industry R&D	-0.002	*0.008*	0.001	*0.010*	-0.002	*0.008*	0.001	*0.010*
… x public	0.056***	*0.010*	0.055***	*0.012*	0.057***	*0.010*	0.056***	*0.013*
Size	-0.013***	*0.001*	-0.011***	*0.001*	-0.015***	*0.001*	-0.013***	*0.001*
Working capital (net of cash)	-0.154***	*0.005*	-0.154***	*0.007*	-0.155***	*0.005*	-0.156***	*0.008*
Investment	-0.077***	*0.008*	-0.102***	*0.011*	-0.079***	*0.008*	-0.108***	*0.011*
Leverage	-0.192***	*0.005*	-0.202***	*0.007*	-0.193***	*0.005*	-0.203***	*0.008*
Dividend dummy	-0.011***	*0.003*	-0.015***	*0.004*	-0.015***	*0.003*	-0.019***	*0.004*
Industry acquisitions	0.009	*0.021*	-0.121***	*0.047*	0.005	*0.021*	-0.132***	*0.046*
Industry inform. asymmetry	-0.056***	*0.018*	-0.008	*0.022*	-0.049***	*0.018*	-0.003	*0.022*
Industry minimum efficient scale					0.001	*0.001*	0.001	*0.001*
Adjusted R^2	36.7%		40.2%		36.1%		39.4%	
Wald test: all coeff. = 0 (F)	207.1***		135.0***		209.2***		134.8***	
No. observations	48,273		26,509		48,273		26,509	
No. firms	17,065		8,719		17,065		8,719	
Public dummy estimate in model w. no interaction effects	0.123***	*0.005*	0.130***	*0.007*	0.125***	*0.006*	0.134***	*0.007*

Panel B.

Dependent variable:	cash / assets						log (cash / net assets)		log (cash / sales)	
	Years 2007 and 2008 excluded		Only C Corps		Alternative industry concentration measure		Baseline sample and controls		Baseline sample and controls	
					HHI	HHI cons.				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Public	-0.003	0.015	-0.004	0.014	-0.022	0.005	0.098	0.354**	-0.144	-0.043
	0.011	*0.015*	*0.011*	*0.015*	*0.014*	*0.009*	*0.108*	*0.161*	*0.111*	*0.164*
Industry disclosure cost										
industry concentration	-0.015*		-0.007		0.038	-0.088**	-0.320***		-0.133	
	0.008		*0.011*		*0.051*	*0.038*	*0.117*		*0.114*	
information leakage		-0.005*		-0.006				-0.181***		-0.180***
		0.003		*0.004*				*0.040*		*0.039*
… x public	0.086***	0.026***	0.072***	0.030***	0.164*	0.287***	0.942***	0.458***	0.501***	0.282***
	0.017	*0.006*	*0.018*	*0.006*	*0.096*	*0.078*	*0.180*	*0.065*	*0.179*	*0.064*
Industry return on capital	0.012***	0.022***	0.011***	0.018***	0.009***	0.009***	0.124***	0.222***	0.155***	0.305***
	0.002	*0.005*	*0.003*	*0.006*	*0.003*	*0.002*	*0.028*	*0.061*	*0.027*	*0.061*
… x public	0.028***	0.034***	0.027***	0.037***	0.035***	0.030***	0.279***	0.343***	0.342***	0.372***
	0.005	*0.009*	*0.005*	*0.008*	*0.006*	*0.004*	*0.044*	*0.077*	*0.046*	*0.077*
Industry cash flow volatility	0.474***	0.194***	0.417***	0.195**	0.145**	0.455***	3.449***	2.213**	2.740***	0.958
	0.054	*0.068*	*0.071*	*0.089*	*0.070*	*0.048*	*0.670*	*0.998*	*0.640*	*0.984*
… x public	0.665***	0.736***	0.610***	0.644***	0.938***	0.689***	5.947***	6.008***	9.405***	10.943***
	0.096	*0.131*	*0.101*	*0.131*	*0.134*	*0.088*	*0.948*	*1.414*	*0.946*	*1.427*
	All control variables included in my baseline cash analysis (Table 4, cols. 2 and 4) are also included but not reported for brevity									
Adjusted R^2	36.7%	39.4%	41.9%	43.8%	44.0%	36.0%	24.1%	27.1%	34.5%	37.8%
Wald test: all coeff. = 0 (F)	200.1***	123.9***	195.8***	128.7***	135.1***	217.2***	279.4***	176.9***	406.5***	245.9***
No. observations	33,805	18,629	30,934	18,181	21,278	48,273	46,895	25,907	46,909	25,913
No. firms	14,305	7,341	9,201	5,021	6,698	17,065	` 16,661	8,553	16,665	8,557
Public dummy estimate in model w/o interaction effects	0.126***	0.130***	0.115***	0.120***	0.128***	0.125***	1.295***	1.204***	1.354***	1.385***
	0.006	*0.008*	*0.006*	*0.008*	*0.008*	*0.006*	*0.055*	*0.073*	*0.058*	*0.078*

Table 6. Cash comparison: Matching.

This table compares the cash policies of public and private firms (equation (5) in the main text), matching public and private firms. Panel A presents results using local linear propensity score matching, with the propensity scores based on size, cash flow, sales growth, and industry (5-digit NAICS). Each regression includes an intercept and year effects (not reported for brevity). Standard errors are bootstrapped to account for matching-induced variance (500 replications) and are shown in italics beside the coefficient estimates. The matching algorithm in Panel B follows Asker, Farre-Mensa, and Ljungqvist (2010a): It matches on size and industry (four-digit NAICS) using nearest-neighbor matching with resampling. In panel B, industry concentration is measured at the four-digit NAICS level, the same level at which industries are matched. Each regression includes an intercept and year effects (not reported for brevity). Heteroskedasticity-robust standard errors clustered at the firm level are shown in italics beside the coefficient estimates. In both Panels A and B, I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

Panel A. Local linear propensity score matching

	Dependent variable: cash / assets							
	(1)		(2)		(3)		(4)	
	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*
Public	0.113***	*0.013*	-0.081***	*0.028*	0.112***	*0.016*	-0.097***	*0.036*
Industry disclosure cost								
industry concentration	0.045**	*0.020*	-0.040	*0.050*				
information leakage					0.012	*0.008*	-0.006	*0.017*
… x public			0.097*	*0.053*			0.028	*0.018*
Industry return on capital	0.031***	*0.006*	-0.010	*0.015*	0.058***	*0.011*	0.024	*0.021*
… x public			0.052***	*0.016*			0.044**	*0.020*
Industry cash flow volatility	0.924***	*0.107*	0.631**	*0.279*	1.092***	*0.172*	0.222	*0.352*
… x public			0.385	*0.290*			1.097***	*0.372*
Cash flow	-0.056***	*0.012*	-0.004	*0.011*	-0.067***	*0.017*	-0.033*	*0.017*
… x public			-0.153***	*0.020*			-0.141***	*0.031*
Industry R&D	0.032***	*0.009*	-0.018	*0.015*	0.033***	*0.012*	-0.019	*0.018*
… x public			0.088***	*0.017*			0.087***	*0.022*
Size	-0.003	*0.003*	-0.005	*0.003*	0.012***	*0.004*	0.011***	*0.004*
Working capital (net of cash)	-0.244***	*0.018*	-0.224***	*0.017*	-0.256***	*0.025*	-0.223***	*0.022*
Investment	-0.175***	*0.032*	-0.116***	*0.024*	-0.207***	*0.038*	-0.156***	*0.027*
Leverage	-0.287***	*0.023*	-0.287***	*0.020*	-0.331***	*0.024*	-0.322***	*0.020*
Dividend dummy	-0.053***	*0.008*	-0.027***	*0.009*	-0.036***	*0.013*	-0.001	*0.013*
Industry acquisitions	-0.141	*0.086*	-0.204***	*0.075*	-0.523***	*0.167*	-0.516***	*0.153*
Industry information asymmetry	-0.108*	*0.057*	-0.072	*0.053*	-0.061	*0.039*	-0.004	*0.040*
Adjusted R^2	38.5%		44.0%		41.0%		47.1%	
Wald test: all coeff. = 0 (F)	108.3***		134.1***		75.3***		104.5***	
No. observations	34,212		34,212		14,848		14,848	
No. firms	11,166		11,166		4,662		4,662	

Panel B. Nearest-neighbor matching on industry and size

	Dependent variable: cash / assets							
	(1)		(2)		(3)		(4)	
	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*
Public	0.121***	*0.011*	-0.073**	*0.029*	0.131***	*0.013*	-0.037	*0.027*
Industry disclosure cost								
industry concentration	-0.003	*0.029*	-0.075	*0.047*				
information leakage					0.016*	*0.010*	-0.002	*0.013*
… x public			0.150***	*0.058*			0.040***	*0.015*
Industry return on capital	0.023***	*0.007*	0.001	*0.010*	0.048***	*0.012*	0.023	*0.015*
… x public			0.035***	*0.011*			0.038**	*0.015*
Industry cash flow volatility	0.988***	*0.169*	0.668**	*0.280*	0.964***	*0.181*	0.588***	*0.218*
… x public			0.660**	*0.323*			0.874***	*0.289*
Cash flow	-0.090***	*0.016*	-0.020	*0.018*	-0.118***	*0.013*	-0.025	*0.016*
… x public			-0.118***	*0.024*			-0.139***	*0.027*
Industry R&D	0.033***	*0.011*	-0.018	*0.014*	0.028**	*0.012*	-0.031**	*0.014*
… x public			0.092***	*0.018*			0.106***	*0.018*
Size	0.008**	*0.004*	0.005	*0.004*	0.014***	*0.004*	0.010***	*0.004*
Working capital (net of cash)	-0.223***	*0.025*	-0.204***	*0.023*	-0.212***	*0.024*	-0.205***	*0.023*
Investment	-0.156***	*0.043*	-0.105***	*0.025*	-0.166***	*0.042*	-0.128***	*0.026*
Leverage	-0.294***	*0.021*	-0.298***	*0.018*	-0.285***	*0.026*	-0.289***	*0.020*
Dividend dummy	-0.020	*0.020*	-0.002	*0.021*	0.007	*0.018*	0.033*	*0.018*
Industry acquisitions	-0.200*	*0.111*	-0.232**	*0.106*	-0.597***	*0.226*	-0.622***	*0.209*
Industry information asymmetry	0.094	*0.164*	0.078	*0.161*	-0.179*	*0.096*	-0.267***	*0.089*
Adjusted R^2	39.3%		44.8%		43.2%		49.2%	
Wald test: all coeff. = 0 (F)	77.7***		99.6***		61.3***		84.8***	
No. observations	13,782		13,782		9,484		9,484	
No. firms	3,198		3,198		1,964		1,964	

Table 7. Cash comparison: IPO sample.

This table examines changes in cash holdings around the IPO for a particular sample of 98 IPO firms: firms that go public for the sole purpose of allowing existing shareholders to cash out. This ensures that the IPO proceeds do not mechanically increase the firms' cash holdings. To be part of the sample, the IPO needs to have taken place between 1990 and 2007. In columns 1 through 3, I include firm and year fixed effects (not reported for brevity). Heteroskedasticity-robust standard errors are shown in italics underneath the coefficient estimates. In columns 4 and 5, I match each IPO firm to up to five already public firms (matching is on industry and size, and it is done the first year that the IPO firm is in the sample). Regressions 4 and 5 include an intercept and year effects (also not reported for brevity); heteroskedasticity-robust standard errors clustered at the firm level are shown in italics underneath the coefficient estimates. In columns 3 through 5, additional control variables include size, sales growth, cash flow, investment, and industry cash flow volatility. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Dependent variable: cash / assets				
	Own difference			Comparison with matched controls	
	(1)	(2)	(3)	(4)	(5)
Pre-IPO (all years)				-0.070***	
				0.012	
Year >3 Pre-IPO					-0.072***
					0.019
Year 3 Pre-IPO					-0.086***
					0.013
Year 2 Pre-IPO					-0.070***
					0.013
Year 1 Pre-IPO (last year before IPO)					-0.054***
					0.015
Post-IPO (all years)	0.045***		0.040***	0.000	
	0.008		*0.009*	*0.017*	
Year 1 Post-IPO (first year after IPO)		0.050***			-0.009
		0.011			*0.016*
Year 2 Post-IPO		0.038***			-0.013
		0.011			*0.016*
Year 3 Post-IPO		0.043***			0.001
		0.012			*0.018*
Year >3 Post-IPO		0.028**			0.006
		0.014			*0.020*
Additional control variables	No	No	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	No	No
Year fixed effects	Yes	Yes	Yes	Yes	Yes
R^2 (1-3: within / 4-6: adjusted)	15.6%	15.8%	18.1%	16.5%	16.4%
No. observations	1,080	1,080	1,080	5,344	5,344
No. firms	98	98	98	463	463

Table 8. Cash comparison: Instrumental variables.

This table compares the cash policies of public and private firms (equation (5) in the main text), instrumenting a firm's public or private status. As explained in the main text, I use a firm's location and its industry minimum efficient scale as source of identification. Each regression includes an intercept and year effects (not reported for brevity) and is estimated by GMM. Heteroskedasticity-robust standard errors clustered at the firm level are shown in italics beside the coefficient estimates. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Dependent variable: cash / assets							
	(1)		(2)		(3)		(4)	
	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*	coeff.	*s.e.*
Public	0.450***	*0.034*	0.265***	*0.035*	0.378***	*0.038*	0.316***	*0.046*
Industry disclosure cost								
industry concentration	0.001	*0.010*	-0.091***	*0.012*				
information leakage					0.002	*0.003*	-0.009**	*0.003*
… x public			0.235***	*0.024*			0.028***	*0.007*
Industry return on capital	0.003	*0.002*	0.001	*0.002*	0.007	*0.006*	0.002	*0.006*
… x public			0.016***	*0.005*			0.013	*0.009*
Industry cash flow volatility	0.202**	*0.079*	-0.127	*0.088*	0.055	*0.093*	-0.413***	*0.118*
… x public			0.837***	*0.109*			1.000***	*0.146*
Cash flow	0.016***	*0.004*	0.023***	*0.004*	0.008	*0.006*	0.032***	*0.007*
… x public			-0.104***	*0.036*			-0.185***	*0.038*
Industry R&D	0.063***	*0.006*	-0.127***	*0.036*	0.075***	*0.007*	-0.152***	*0.049*
… x public			0.187***	*0.037*			0.221***	*0.051*
Size	-0.065***	*0.005*	-0.060***	*0.005*	-0.053***	*0.005*	-0.054***	*0.006*
Working capital (net of cash)	-0.144***	*0.007*	-0.132***	*0.006*	-0.156***	*0.009*	-0.140***	*0.009*
Investment	-0.043***	*0.012*	-0.022*	*0.012*	-0.115***	*0.015*	-0.066***	*0.015*
Leverage	-0.139***	*0.008*	-0.145***	*0.008*	-0.155***	*0.010*	-0.149***	*0.011*
Dividend dummy	-0.022***	*0.004*	-0.015***	*0.004*	-0.021***	*0.005*	-0.008*	*0.005*
Industry acquisitions	0.067***	*0.024*	0.045**	*0.023*	-0.147***	*0.054*	-0.050	*0.055*
Industry information asymmetry	-0.051**	*0.024*	-0.062***	*0.024*	0.098***	*0.032*	0.080**	*0.032*
Underident. test: K-P *rk* LM stat.	299.0		242.3		198.8		151.6	
Weak ident. test: K-P *rk* Wald *F* stat.	315.6		42.0		223.9		27.1	
Adjusted R^2	13.9%		17.1%		25.7%		22.5%	
Wald test: all coeff. = 0 (*F*)	218.1***		186.7***		140.8***		125.1***	
No. observations	48,252		48,252		26,485		26,485	
No. firms	17,060		17,060		8,713		8,713	

Table 9. Cash flow sensitivity of cash.

This table compares the cash flow sensitivity of cash of public and private firms, following Almeida, Campello, and Weisbach (2004). The dependent variable is the change in cash holdings. The main independent variable of interest is cash flow over assets, but I also control for investment opportunities (which I measure as industry return on capital, i.e., market-to-book) and firm size. Each regression includes a firm-specific intercept and year effects (not reported for brevity). Heteroskedasticity-robust standard errors are shown in italics underneath the coefficient estimates. I use ***, **, and * to denote significance at the 1%, 5%, and 10% level (two-sided), respectively.

	Dependent variable: Δ(cash / assets)								
	Public firms			Private firms			Public and private firms		
	All firms	Small firms (bottom three deciles)	Large firms (top three deciles)	All firms	Small firms (bottom three deciles)	Large firms (top three deciles)	All firms	Local linear propen. score matching	Nearest-neighbor matching
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Cash flow	0.068***	0.084***	-0.004	0.023***	0.015*	0.023***	0.025***	0.004	-0.003
	0.014	*0.019*	*0.030*	*0.005*	*0.008*	*0.008*	*0.005*	*0.017*	*0.011*
Cash flow x public							0.035**	0.073***	0.100***
							0.015	*0.023*	*0.022*
Industry return on capital	0.007***	0.011	0.004	0.003	0.000	0.005	0.007***	0.012***	0.013**
	0.003	*0.008*	*0.003*	*0.002*	*0.005*	*0.004*	*0.002*	*0.004*	*0.006*
Size	-0.011***	-0.011**	-0.005	0.014***	0.019**	0.005	-0.003	-0.014***	-0.012***
	0.003	*0.005*	*0.007*	*0.004*	*0.008*	*0.007*	*0.003*	*0.005*	*0.004*
R^2	15.4%	19.2%	11.1%	34.7%	34.2%	39.3%	27.7%	15.6%	26.5%
Wald test: all coeff. = 0 (*F*)	12.6***	6.0***	7.4***	6.9***	2.0**	2.2**	9.6***	5.1***	10.6***
No. observations	21,857	5,645	7,668	42,674	12,788	12,335	64,531	30,077	13,730
No. firms	4,436	1,330	1,330	18,892	5,667	5,667	23,328	7,813	2,801