Internet Appendix for

"Too Much Liquidity?
Seemingly Excess Cash for Innovative Firms"

Zhaozhao He* and Stephen Ciccone**

* zhaozhao.he@unh.edu.
** Stephen.ciccone@unh.edu
Both authors are with Peter T. Paul College of Business and Economics, University of New Hampshire, 10 Garrison Avenue, Durham, New Hampshire 03824.

1. Additional Tests and Robustness Checks

1.1. System GMM Estimation

To alleviate the concern that the observed results might be contaminated by measurement errors in q or *Cash Flow,* we implement an alternative estimation method—the system GMM estimator. Since *MKTBK* is an imperfect proxy for q, investment opportunities may explain many of the cash and cash flow effects.[1] Moreover, the measurement error of traditional *Cash Flow* grows because an increased amount of R&D investments are expensed over time. We thus follow Brown and Petersen (2009) to calculate *Gross Cash Flow*, which adds back R&D to conventional *Cash Flow*. We then implement the system GMM estimators using lagged levels of the dependent variable dated $t - 3$ to $t - 4$ as instruments for the differenced equations and lagged differences dated $t - 2$ as instruments for the level equations (Brown, Fazzari and Petersen, 2009; Brown and Petersen, 2009). We divide the sample into four subsample periods: 1971–1979, 1980–1989, 1990–1999, and 2000–2012. Standard errors are robust to heteroskedasticity and clustering by firm.

Table IA.1 shows that the GMM estimate of *Cash* generally exhibits an increasing trend across the specifications. The coefficients of *Cash* are −0.018 (*t*-stat = −0.49), 0.075 (*t*-stat = 2.95), 0.125 (*t*-stat = 4.55), and 0.073 (*t*-stat = 4.93) during the four subperiods, confirming our results. The serial correlation tests indicate that we cannot reject the null hypothesis of no second-order serial correlation for three out of four specifications. The Hansen J tests of overidentification suggest that all instruments are valid. Note that the coefficients of *Gross Cash Flow* are positive and significant, consistent with Brown and Petersen (2009).

1.2. Other Robustness Checks

We report the results of all other robustness tests in Table IA.2. First, to address the concern of the measurement error in cash flows, we follow Lewellen and Lewellen (2014) to calculate *Corrected CF,*

[1] Because changes in financial variables correlate with changes in profits, the financial variables may simply reflect information about the growth opportunities.

which adds back extraordinary items and other operating cash flows to correct for the measurement error in the traditional measure of *Cash Flow* that is due to an increase in noncash special item. Column (1) in Panel A of Table IA.1 reports the results of the time-series test using the *Corrected CF.* The coefficient on *Trend* is 0.004 (t-stat = 4.78), exhibiting a comparable trend as that reported in the paper.[2] We also uses the gross cash flow, which adds back R&D to the standard measure of cash flow, as in Brown and Petersen (2009). We find that results remain virtually the same and thus do not report here. Therefore, the result is not affected by the measurement error of cash flow.

Second, to address the concern of measurement error in q, we employ an instrumental variables approach in addition to the system GMM estimation. The validity of instruments crucially relies on the assumption that serial correlation in the measurement error is small or short-lived (Erickson and Whited, 2012). Instead of applying lagged *MKTBK*, as most of the studies do, we use lagged returns, lagged cash and lagged cash flow as instruments to separate the portion of cash flow that is unrelated to *MKTBK*. Lewellen and Lewellen (2014) argue that the lagged returns seem to be a good candidate based on the idea that the measurement error in *MKTBK* mainly comes from the book value of assets rather than from the market value of assets, and that stock prices tend to reflect fundamental value promptly. Thus, the slope on *Cash* (and cash flow) in the second-stage regression captures the portion of *Cash* (and cash flow) that is unrelated to investment opportunities.[3]

Column (2) in Panel A reports the first-stage IV estimates from the cross-sectional regressions of *MKTBK* on *Cash*, *Corrected CF* and past stock returns. The t-statistics are based on standard errors by incorporating Newey-West correction with three lags. All of the instruments are statistically significant at the 1% level except for the lagged *Corrected CF*. The next column shows the time-series test of the *Cash* coefficient with the predicted *MKTBK* and *Corrected CF* incorporated. The results are virtually unchanged.

---

[2] Precisely, the *Corrected CF* is the sum of Income before extraordinary items (IBC), extraordinary items and discontinued operations (XIDOC), depreciation and amortization (DPC), deferred taxes (TXDC), equity in net loss of unconsolidated subsidiaries (ESUBC), losses from the sale of PP&E (SPPIV) and funds from operations—other (FOPO).

[3] Admittedly, these instruments are not perfect, but they are able to deliver precise and conservative estimates under weak assumption (Lewellen and Lewellen 2014, p.25).

The trend in *Cash* coefficient is still statistically significant at a 1% level (t-stat = 5.67). As a result, the dramatic increase in the cash coefficient is not contaminated by the measurement error in *q*.

Third, the real problem might lie in the way of measuring cash reserves, which contain marketable securities. Duchin, Gilbert, Harford, and Hrdlicka (2014) discover that a fraction of cash reserves is held in risky assets and thus challenge the precautionary motive of holding cash. To investigate whether the composition of cash reserves influences the observed results, we exclude short-term securities (*STI*) from *Cash* and repeat the analyses with error-corrected *MKTBK* and cash flow. As shown in column (4) of Panel A, the time trend is exactly the same as reported in columns (1) and (3) and continues to be significant at a 1% level (t-stat = 3.46). Therefore, this finding is robust to a more precise definition of cash holdings and is not driven by the portion of "risky" reserves.

Fourth, to examine whether the upward trend of *Cash* coefficient varies with different specifications, columns (5)–(6) report the results when including lagged *Cash* in Equation (1), and columns (7)–(9) show the results after controlling for previous two, three, and four periods financing sources. Overall, we find that the coefficient on *Trend* for *Cash* coefficient remains at 0.004 and significant at the 1% level. In addition, R&D firms are also increasingly sensitive to previous period cash holdings, although the positive trend is much smaller than the estimates from contemporaneous cash. The time trend has a coefficient of 0.001 with a t-statistic of 4.87.

Fifth, to address the concern that new firms with higher R&D and cash may have entered the sample, we construct a balanced panel containing firms that exist in every year from 1980 to 2012. This panel contains 526 firms and 17,358 observations in total. As shown in column (1) of Panel B, this subsample continues to produce a positive and statistically significant time trend of 0.002 (t-stat = 1.85).

Sixth, we consider industry composition. Following the investment-cash flow literature, we divide manufacturing firms into three industry groups—durable goods, nondurable goods, and high-tech industries. Doing so helps to maintain homogeneity of sample composition.[4] Columns (2) to (4) in Panel B

[4] *High-tech manufacturing industries* are defined as SIC codes 283, 357, 366, 367, 382, 384 (Brown, Fazzari, and Petersen, 2009). *Durable goods industries* are defined as firms not in the high-tech industries and with two-digits SIC codes between 24 and 25, or

show the statistical tests by industries. All three trends are positive and statistically significant at the 1% level. The coefficients of *Trend* are 0.002 (t-stat = 4.51), 0.002 (t-stat = 6.06), and 0.004 (t-stat = 6.07) for durable goods, non-durable goods, and high-tech industries, respectively. Therefore, the role of cash in determining R&D has dramatically increased for all three manufacturing industries. Overall, we find robust evidence for the increased importance of cash holdings for R&D investments.

2. Excess Cash

2.1. Computing Excess Cash

One of the potential concerns on the computation of excess cash is the endogeneity issue of the *MKTBK* variable. As excess cash is related to firm value, using *MKTBK* as a proxy for investment opportunities to predict normal cash levels will be problematic. We tackle this issue by instrumenting the *MKTBK* with a firm's two years lagged sales growth because Jorgenson's (1963) neoclassical investment model specifies that changes in output (sales growth) are important in determining a firm's investment decisions and because it is difficult to argue that current cash holdings affect past sales growth. A similar method has been applied in Dittmar and Mahrt-Smith (2007), who suggest that firms' past sales growth rates are good proxies for investment opportunities. The second-stage estimation applies the cross-sectional regressions with the instrumented *MKTBK*.

We see that in column (2) of Table IA.3, the IV method generates a set of similar estimates as those in column (1). All of our results are not materially different if we use the IV method. We also estimate a pooled ordinary least squares (OLS) regression controlling for industry and year fixed effects and report the results in columns (3) and (4). The estimated coefficients are largely consistent. Column (5) shows the results of the first-stage IV regression.

2.2. Excess cash, corporate investments and financial constraints

---

between 32 and 38, inclusive. *Non-Durable goods industries* are defined as firms not in the high-tech industries and with two-digits SIC codes between 20 and 23, or between 26 and 31, inclusive (Chen and Chen, 2012).

We examine the relation between excess cash and corporate spending by ranking firm-years into excess cash quartiles based on the previous year's holdings of excess cash. The cash quartiles are generated for every year, and firms are regrouped each year. Table IA.4 reports the averages of selected firm characteristics and expenditures for each quartile. As expected, cash holdings increase monotonically across the quartiles, from an average of 6.88% for firms in the lowest quartile (Q1) to 32.23% for those in the highest quartile (Q4).

Turning to investment, we find that R&D expenditures increase with excess cash, using either R&D-to-sales or R&D-to-assets. Firms in the Q4 of excess cash invest more in R&D than those in the Q1. The differences are significant as reported in the last column. The *t*-statistics for difference-in-means tests (Q1 − Q4) are −22.94 using R&D-to-sales and −23.78 using R&D-to-assets. This might explain why firms with more excess cash earn higher future returns, as found in previous studies.

In contrast, capital expenditures, payouts to shareholders, and acquisition costs do not display any strong patterns across excess cash quartiles, consistent with the findings in Opler, Pinkowitz, Stulz and Williamson (1999), who conclude that the propensity to spend excess cash on investment and acquisitions is small. Moving from Q1 to Q4, capital expenditures drop from 0.059 to 0.053 (*t*-stat of the difference = 10.52), payout ratio increases from 0.019 to 0.023 (*t*-stat of the difference = −9.56), and acquisition costs increase marginally from 0.016 to 0.019 (*t*-stat of the difference = −5.57). Thus, R&D increasing with excess cash is distinct from the cases of all other investments.

We have previously shown a larger increase in the R&D-cash sensitivity for financially constrained R&D firms. If cash allows constrained firms to engage more in R&D, then firms that maintain a greater amount of excess cash might suffer more severe financing constraints. To validate this conjecture, we examine the relation between excess cash and financing constraints using a variety of proxies including firm size, age, dividend payer, the occurrence of negative income, industry cash-flow volatility, SA-index (Hadlock and Pierce, 2010), and WW-index (Whited and Wu, 2006). Both indexes are positively correlated with the degree of financial constraints.

Table IA.4 presents the mean values of the financing constraint proxies for each excess cash quartile. Firms in the top quartile, the so-called cash-rich firms, are subject to the most severe financing frictions. They are the youngest and smallest, have the highest SA-index and WW-index, are the least likely to pay dividends, are the most likely to incur operation losses, and are in industries with high cash-flow volatility. These findings cast doubt on the belief that excess cash is simply free cash flow, at least for R&D-intensive firms. Rather, the evidence on the relation among excess cash, R&D spending, and financing constraints implies that innovative firms are undergoing nontrivial financing frictions, which explains why cash has become increasingly important for them.

Table IA.1

System GMM as an alternative estimation method

This table reports the GMM estimation results examining the time-series properties of cash in determining R&D investment. Each column shows the estimates from a 10-year panel using lagged levels of the dependent variable dated *t – 3* to *t – 4* as instruments for the differenced equations and lagged differences dated *t – 2* as instruments for the level equations. The statistics *AR(2)* test the null hypothesis that there is no second-order serial correlation in the first-differenced residuals. The Hansen J tests the null hypothesis that instruments are exogenous. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles. The *t*-statistics in parentheses are based on robust standard errors clustered by firm. *, **, and*** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2)	(3)	(4)
	[1971,1979]	[1980,1989]	[1990,1999]	[2000,2012]
Dependent Variable	R&D	R&D	R&D	R&D
Cash	–0.018	0.075***	0.125***	0.073***
	(–0.493)	(2.950)	(4.551)	(4.932)
Issue	0.237**	–0.004	–0.019	0.063***
	(2.132)	(–0.116)	(–0.615)	(2.615)
Issue_{t-1}	0.014	–0.020	–0.067**	–0.010
	(0.160)	(–1.186)	(–2.096)	(–0.361)
Debt	–0.054	–0.033	0.037	–0.014
	(–1.379)	(–1.148)	(1.308)	(–0.681)
Debt_{t-1}	0.006	0.010	–0.013	0.017
	(0.195)	(0.568)	(–0.678)	(0.890)
Other Funds	0.039	–0.095	0.105**	0.019
	(0.531)	(–1.437)	(2.549)	(1.001)
Other Funds_{t-1}	–0.005	–0.031	–0.029	0.020
	(–0.152)	(–0.838)	(–0.771)	(0.723)
Gross Cash Flow	0.141***	0.002	0.104**	0.093**
	(3.707)	(0.051)	(2.019)	(2.425)
$\text{Gross Cash Flow}_{t-1}$	–0.040	–0.029	–0.136***	–0.055**
	(–1.591)	(–1.124)	(–2.679)	(–2.040)
MKTBK	0.001	0.000	0.004	–0.005***
	(0.734)	(0.138)	(1.641)	(–2.692)
$\text{R\&D}_{t-1}$	0.397*	0.629***	0.699***	0.588***
	(1.905)	(6.013)	(10.513)	(11.405)
$\text{R\&D}_{t-2}$	0.175	0.046	–0.019	0.018
	(1.328)	(0.770)	(–0.663)	(0.873)
N	8,528	14,090	16,531	19,079
AR (2) test (*p*-value)	0.06	0.18	0.80	0.16
Hansen J test (*p*-value)	0.68	0.06	0.14	0.27

Table IA.2
Other Robustness Tests

This table reports the time-series tests of the trends in *Cash* estimates from the cross-sectional yearly regression of *R&D* on *Cash*, current and lagged *Issue, Debt, Cash Flow,* and *Other Funds,* and lagged *MKTBK.* In Panel A, column (1) reports the results using error-corrected CF. Column (2) reports the first-stage estimates from cross-sectional yearly regressions of *MKTBK* on *Cash, Corrected CF* and Annual Returns, incorporating a Newey-West correction with three lags. Column (3) reports the statistical tests with the predicted $MKTBK_{t-1}$ and *Corrected CF*. Column (4) shows the results when excluding the short-term securities (*STI*) from *Cash*. Columns (5) – (6) report results when including Cash $_{t-1}$. Columns (7) – (9) show results when controlling for previous *t-s* periods financing, where s takes values from 2 to 4. In Panel B, column (1) shows results estimated from a balanced sample containing all firms that have survived since 1980. Columns (2) to (4) show statistical tests by durable goods, non-durable goods, and high-tech industries. All models of time trend correct for autocorrelation by incorporating a Yule-Walker correction with four lags. The sample starts with all non-financial and non-utility U.S. industrial firms with non-negative values for cash, sales and common equity, and with lagged assets greater than $1 million dollars in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles. The t-statistics are in italic. *, **, and*** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Panel A: Errors in Variables or Specifications

	Corrected CF	IV For *MKTBK*		Excluding STI	Including Cash $_{t-1}$		Controlling for previous *t-s* periods financing		
	(1)	(2) First-stage	(3) Second-stage	(4)	(5)	(6)	(7) s=2	(8) s=2,3	(9) s=2,3,4
Dependent Variable	*Cash* Coeff.	*MKTBK*	*Cash* Coeff.	*Cash* Coeff.	*Cash* Coeff.	$Cash_{t-1}$ Coeff.	*Cash* Coeff.	*Cash* Coeff.	*Cash* Coeff.
Trend	**0.004** ***		**0.004** ***	**0.004** ***	**0.004** ***	**0.001** ***	**0.004** ***	**0.004** ***	**0.004*****
t-stat	***4.78***		***5.67***	***3.46***	***5.28***	***4.87***	***4.32***	***4.10***	***3.96***
Intercept	-0.012	1.309 ***	-0.017	0.018	-0.017	-0.004	-0.016	-0.015	-0.010
t-stat	*-0.51*	11.74	*-1.14*	*0.68*	*-0.97*	*-0.50*	*-0.65*	*-0.58*	*-0.41*
Lag 1	0.675 ***		0.330 *	0.215	0.276	0.103	0.773 ***	0.729 ***	0.741***
t-stat	*3.99*		*1.84*	*1.26*	1.63	0.62	4.51	4.20	4.19
Lag 2	-0.184		0.078	-0.016	-0.077	-0.115	-0.259	-0.178	-0.180
t-stat	*-0.93*		*0.14*	*-0.09*	-0.46	-0.69	-1.23	-0.85	-0.83
Lag 3	0.318		0.073	0.135	0.272	0.168	0.303	0.279	0.210
t-stat	*1.61*		*0.39*	*0.81*	1.61	1.01	1.44	1.33	0.97
Lag 4	-0.039		-0.041	0.352	0.058	-0.110	-0.039	-0.068	0.006
t-stat	*-0.23*		*0.24*	*1.52*	0.34	-0.65	-0.23	-0.39	0.04
Cash		1.087 ***							
t-stat		9.13							
Cash $_{t-1}$		0.607 ***							
t-stat		11.18							
Corrected CF		0.838 ***							
t-stat		*4.50*							
Corrected CF $_{t-1}$		-0.599 **							
t-stat		*-2.22*							
Annual Return $_{t-1}$		0.163 ***							
t-stat		*4.25*							
Annual Return $_{t-2}$		0.135 ***							
t-stat		*4.72*							
Annual Return $_{t-3}$		0.130 ***							
t-stat		*5.01*							
Annual Return $_{t-4}$		0.116 ***							
t-stat		*6.35*							
Year	41	38	38	38	38	38	37	36	35
Durbin-Watson	1.91		1.98	1.87	1.90	1.94	1.96	1.97	1.94

Panel B: Sample Composition

	(1) Survivors		(2) Durable		(3) Non-Durable		(4) High-tech	
Dependent Variable	*Cash* Coeff.		Cash Coeff.		Cash Coeff.		Cash Coeff.	
Trend	**0.002**	*	**0.002**	***	**0.002**	***	**0.004**	***
t-stat	***1.85***		***4.51***		***6.06***		***6.07***	
Intercept	0.020		-0.006		-0.005		0.005	
t-stat	*0.84*		*-0.55*		*-0.52*		*0.41*	
Lag 1	0.841	***	0.179		0.031		0.142	
t-stat	*4.57*		*1.02*		*0.17*		*0.81*	
Lag 2	-0.345		0.016		-0.333	*	0.069	
t-stat	*-1.53*		*0.09*		*-1.76*		*0.40*	
Lag 3	0.423	*	0.179		-0.220		0.237	
t-stat	*1.80*		*1.02*		*-1.50*		*1.38*	
Lag 4	-0.270		-0.073		0.096		-0.111	
t-stat	*-1.43*		*-0.42*		*0.52*		*-0.64*	
Year	33		38		38		38	
Durbin-Watson	2.00		1.97		1.96		1.96	

Table IA.3
Predicting the normal level of cash

This table presents regression results used to compute excess cash. The dependent variable is the natural logarithm of the ratio of cash and marketable securities over total assets net of cash. Columns (1) and (2) are estimated from cross-sectional yearly regressions. Columns (3) and (4) are estimated from a pooled OLS regression controlling for industry and year fixed effects. Columns (2) and (4) are estimated using an instrumental variables (IV) approach with past sales growth as an instrument for *MKTBK*. The results of the first-stage IV regression, where *MKTBK* is the dependent variable, are reported in column (5). *Cash Flow* is calculated as net income plus depreciation and amortization. *NWC* is working capital net of cash and marketable securities. *Industry Cash Flow Volatility* is the mean of standard deviations of cash flow over assets over 10 year at the two-digit SIC code level. *MKTBK* is calculated as total assets minus book value of common equity plus the market value of common equity, divided by book value of assets. *R&D* is the ratio of R&D expenditures over total assets. *Capex* is the ratio of capital expenditures over total assets. *Leverage* is total debt over total assets. *Dividend Payer* is a variable set to one if the firm paid a dividend in the year, and zero otherwise. The sample comprises all publicly traded nonfinancial and nonutility U.S. industrial firms with lagged assets greater than $1 million in Compustat from 1971 to 2012. All variables are defined in Appendix A. All variables are winsorized at the 1% and 99% levels. Industries are defined at the two-digit SIC codes level. In columns (1) and (2), the *t*-statistics in parentheses are the time-series means of the cross-sectional regressions adjusting for first-order autocorrelation. In columns (3)–(5), the *t*-statistics in parentheses are based on robust standard errors clustered by firm. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

	(1)	(2) IV	(3)	(4) IV	(5) First-stage
	Cross-sectional regression		Fixed-effects regression		
Dependent Variable	*ln(Cash)*	*ln(Cash)*	*ln(Cash)*	*ln(Cash)*	*MKTBK*
Size	–0.042***	–0.044***	–0.033***	–0.035***	–0.038***
	(–5.414)	(–5.683)	(–5.903)	(–5.860)	(–7.356)
Cash Flow	0.500***	0.628***	0.059**	0.085***	–0.056
	(3.044)	(3.079)	(2.222)	(2.720)	(–1.054)
NWC	–0.935***	–0.981***	–0.913***	–0.952***	–0.264***
	(–15.250)	(–14.808)	(–20.960)	(–19.930)	(–5.599)
R&D	1.687***	1.966***	1.566***	1.564***	1.308***
	(15.361)	(12.743)	(33.980)	(22.303)	(15.647)
Industry Cash Flow Volatility	1.613***	1.818***	1.393***	1.311***	0.216
	(6.440)	(7.198)	(9.498)	(8.595)	(1.329)
Sales Growth $_{t-2}$					0.062***
					(12.383)
MKTBK	0.109***	0.099*	0.104***	0.124***	
	(16.866)	(1.877)	(21.682)	(3.704)	
Capex	0.137	0.140	1.044***	1.079***	
	(0.550)	(0.568)	(12.920)	(11.662)	
Dividend Payer	–0.214***	–0.200***	–0.199***	–0.191***	
	(–5.569)	(–5.595)	(–9.440)	(–8.481)	
Leverage	–3.328***	–3.417***	–3.540***	–3.622***	
	(–24.245)	(–24.332)	(–69.812)	(–67.652)	
Constant	–1.889***	–1.723***	–1.490***	–1.359***	1.864***
	(–41.171)	(–39.113)	(–34.967)	(–31.050)	(42.890)
Industry fixed effects	No	No	Yes	Yes	Yes
Year fixed effects	No	No	Yes	Yes	Yes
N	137,430	123,738	137,430	123,738	123,738
Average R^2/Adjusted R^2	0.3434	0.3280	0.4032	0.3894	0.1902

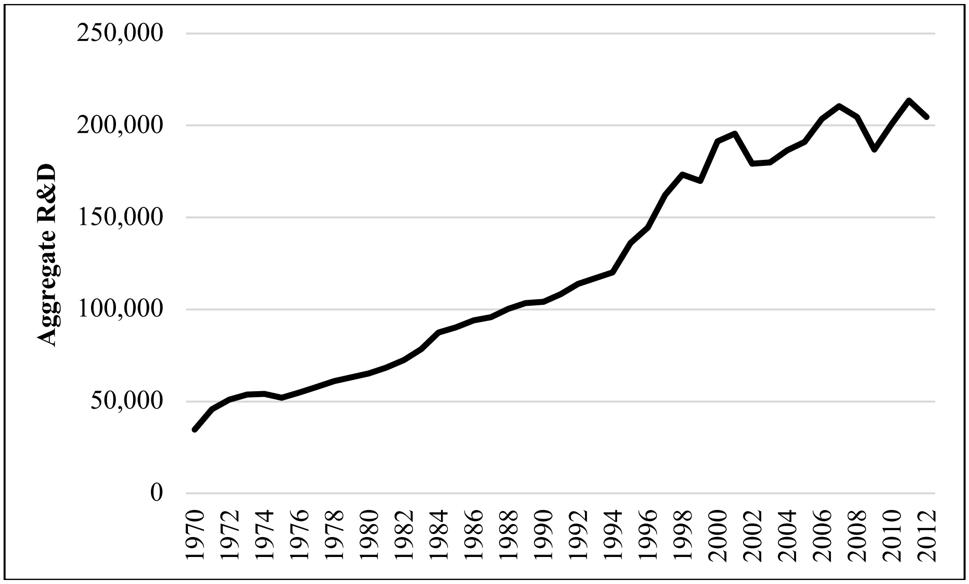
Table IA.4
Firm characteristics by excess cash quartiles

This table reports firm characteristics and spending patterns of R&D firms in each excess cash quartile. The sample consists of a total of 81,559 firm-year observations from 1971 to 2012. *R&D firms* are defined as firms with nonzero R&D capitalized stock in any given period. Firm-years are ranked into quartiles based on the previous year's holdings of excess cash, which is the residual estimated from yearly cross-sectional cash holding regressions in which the dependent variable is the natural log of *Cash* divided by assets net of *Cash*. Quartiles for excess cash are generated every year, and firms are regrouped each year. *Cash-to-assets* is cash plus marketable securities divided by total assets. *R&D-to-sales* is the ratio of R&D expenditure to sales. *R&D-to-assets* is the ratio of R&D expenditure to total assets. *Capex* is the ratio of capital expenditures to total assets. *Payout Ratio* is the sum of dividend and stock repurchase, scaled by total assets. *Acquisition Cost* is the ratio of acquisition costs to total assets. *Age* is the number of years the firm is listed on Compustat. *Size* is the natural logarithm of book assets that is converted to 2005 dollars. *SA index* is a financing constraint index developed by Hadlock and Pierce (2010). *WW index* is a financing constraint index based on Whited and Wu (2006). *Dividend Payer* is an indicator variable equal to one if a common dividend is positive, zero otherwise. *Negative Income* is an indicator variable that equals one if a firm suffers operating loss in any given year. *Industry cash flow volatility* is the mean standard deviation of cash flow over assets for a 10 year rolling window. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles. The *t*-statistics and *p*-values for difference-in-means tests between the first and fourth quartiles of excess cash are reported in the last column.

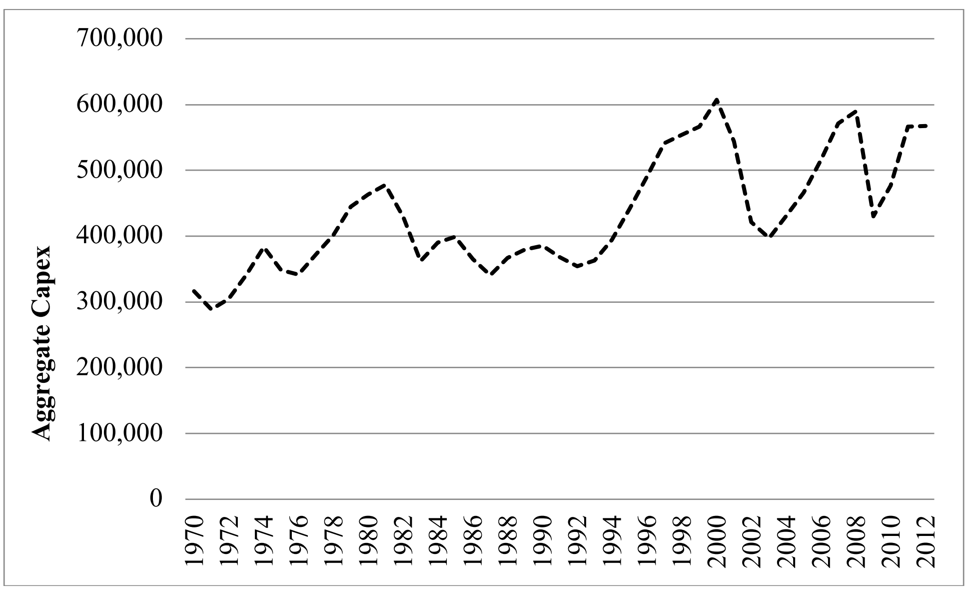
	Quartiles of previous year excess cash				
Variable	First quartile	Second quartile	Third quartile	Fourth quartile	*t*-statistic *p-value*
Excess Cash	–0.9739	–0.1099	0.4637	1.2528	–190.99
					0.0001
Cash-to-assets	0.0688	0.1299	0.2079	0.3223	–130.73
					0.0001
R&D-to-sales	0.1121	0.1782	0.2448	0.2753	–22.94
					0.0001
R&D-to-assets	0.0522	0.0671	0.0783	0.0748	–23.78
					0.0001
Capex	0.0590	0.0575	0.0575	0.0531	10.52
					0.0001
Payout Ratio	0.0188	0.0228	0.0236	0.0230	–9.56
					0.0001
Acquisition Cost	0.0162	0.0167	0.0170	0.0193	–5.57
					0.0001
Age	18.8299	18.7484	16.7952	14.3954	35.16
					0.0001
Size	5.2568	5.4882	5.3358	4.9256	15.78
					0.0001
SA-index	–3.2016	–3.2631	–3.1762	–2.9922	–24.07
					0.0001
WW-index	–0.2391	–0.2498	–0.2389	–0.2130	–20.59
					0.0001
Dividend Payer	0.4005	0.4229	0.3808	0.3091	–17.07
					0.0001
Negative Income	0.2925	0.2933	0.3082	0.3752	–17.07
					0.0001
Industry Cash Flow Volatility	0.1049	0.1093	0.1125	0.1122	–9.82
					0.0001

Figure IA.1
R&D and capital expenditures over 1970–2012

Panels A and B illustrate the aggregate R&D and Capital expenditures since 1970. The dollar values are CPI-adjusted in millions of 2005 dollars. The sample includes all nonfinancial and nonutility U.S. industrial firms with Compustat coverage during the period of 1970–2012. Firms must have non-negative values for cash, sales, and common equity before entering the sample. Firm-years with lagged book assets smaller than $1 million are excluded to mitigate outliers.



Panel A. Aggregate R&D expenditures (in millions of 2005 dollars)



Panel B. Aggregate capital expenditures (in millions of 2005 dollars)

Figure IA.2
Trends of R&D sensitivities to financing sources

This figure plots the parameter estimates of financing variables from the cross-sectional yearly regression of *R&D* on *Cash*, current and lagged *Issue*, *Debt*, *Cash Flow*, and *Other Funds*, and lagged *MKTBK*. The sample consists of R&D firms from 1971 to 2012, a total of 81,559 firm-year observations. *R&D firms* are defined as firms with nonzero R&D capitalized stock in any given period. *Cash* is cash and cash equivalents. *Issue* is the cash proceeds gathered from equity issuance from the statement of cash flow. *Debt* is the cash proceeds from long-term debt issuance. *Cash Flow* is calculated as net income plus depreciation and amortization. *Other Funds* include sale of assets and investment and sources of other funds. All variables are deflated by lagged total assets. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles.

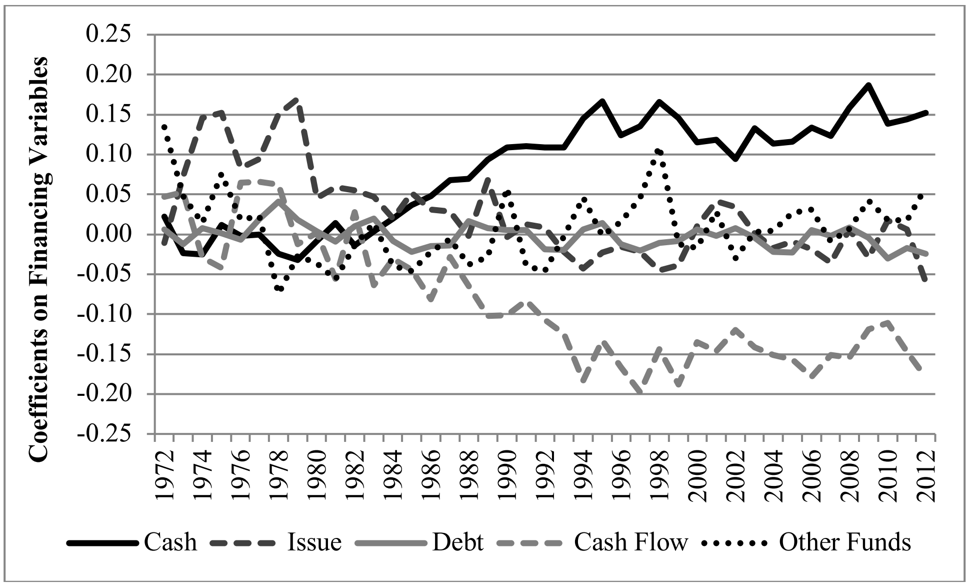


Figure IA.3
Trend of R&D smoothing with cash reserves

This figure plots the parameter estimate of change in cash holdings (Δ *Cash*) from the cross-sectional yearly regression of *R&D* on Δ *Cash*, current and lagged *Issue*, *Debt*, *Cash Flow*, and *Other Funds*, and lagged *MKTBK*. The sample consists of R&D firms from 1971 to 2012, a total of 81,559 firm-year observations. *R&D firms* are defined as firms with nonzero R&D capitalized stock in any given period. Δ *Cash* is the change in cash and cash equivalents from previous year. *Issue* is the cash proceeds gathered from equity issuance from the statement of cash flow. *Debt* is the cash proceeds from long-term debt issuance. *Cash Flow* is calculated as net income plus depreciation and amortization. *Other Funds* include sale of assets and investment and sources of other funds. All variables are deflated by lagged total assets. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles.

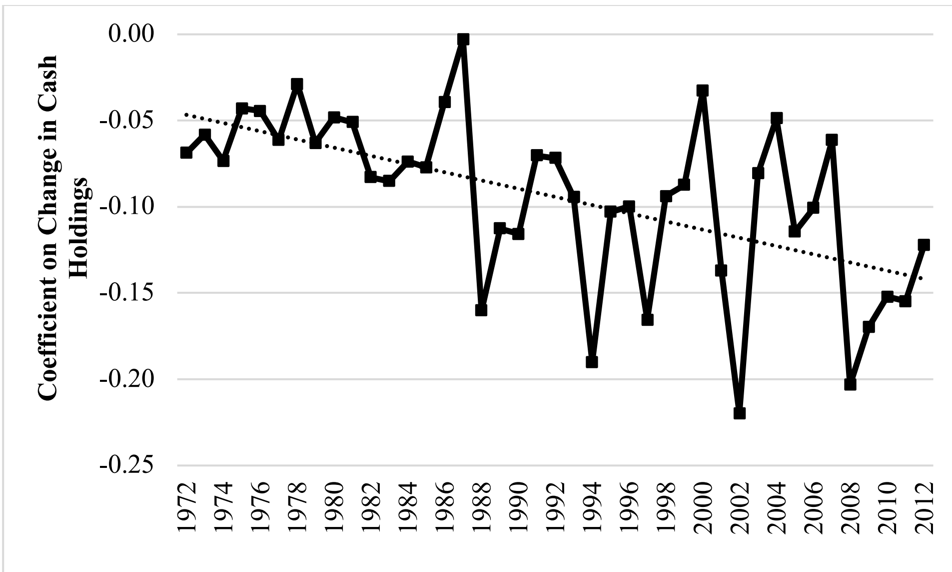


Figure IA.4
The number of firms with positive or negative excess cash every year over 1971–2012

This figure plots the number of firms with positive or negative excess cash every year over 1971–2012. *Excess Cash* is the residual estimated from yearly cross-sectional cash holding regressions in which the dependent variable is the natural log of *Cash* divided by assets net of *Cash*. *Cash* is the cash and cash equivalents. *R&D firms* are defined as firms with nonzero R&D capitalized stock in any given period. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles.

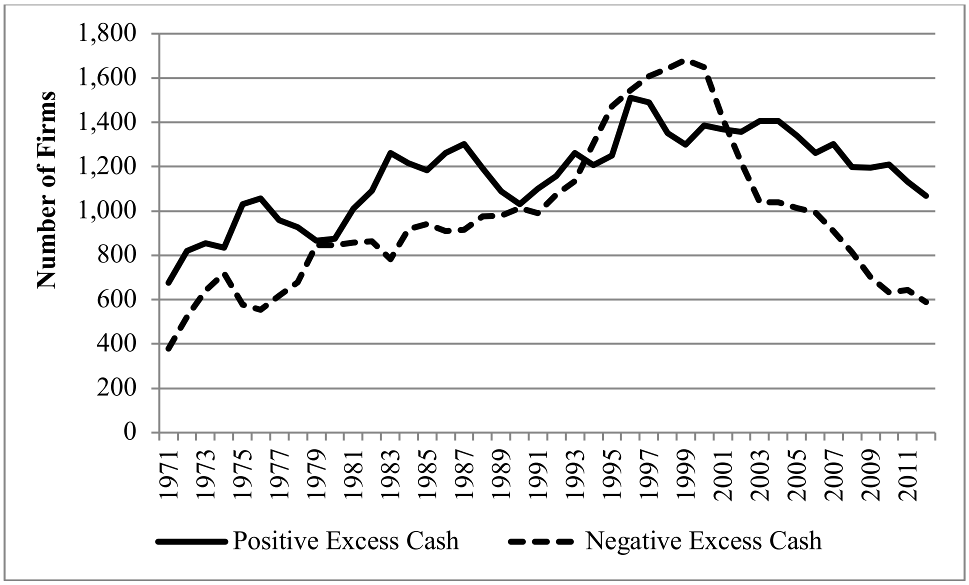
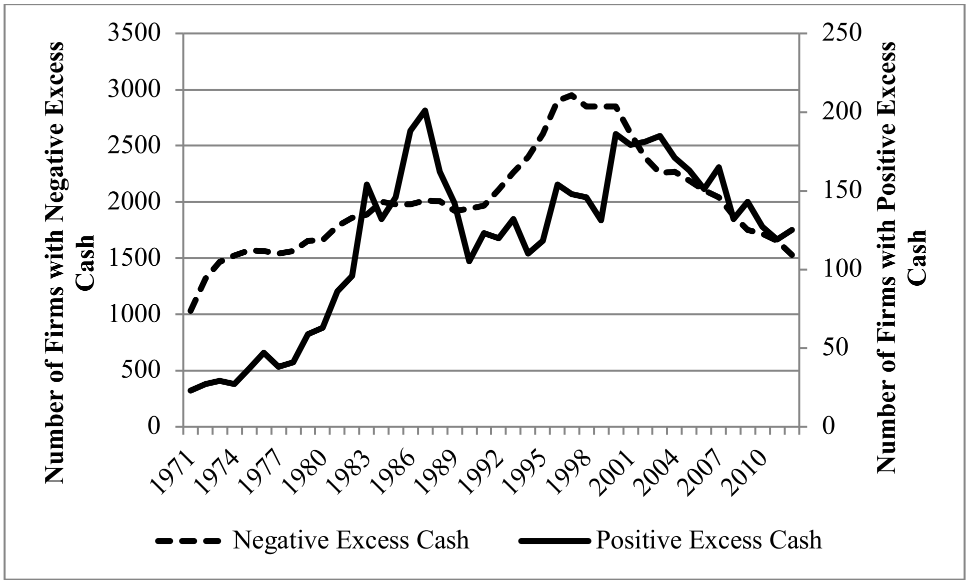


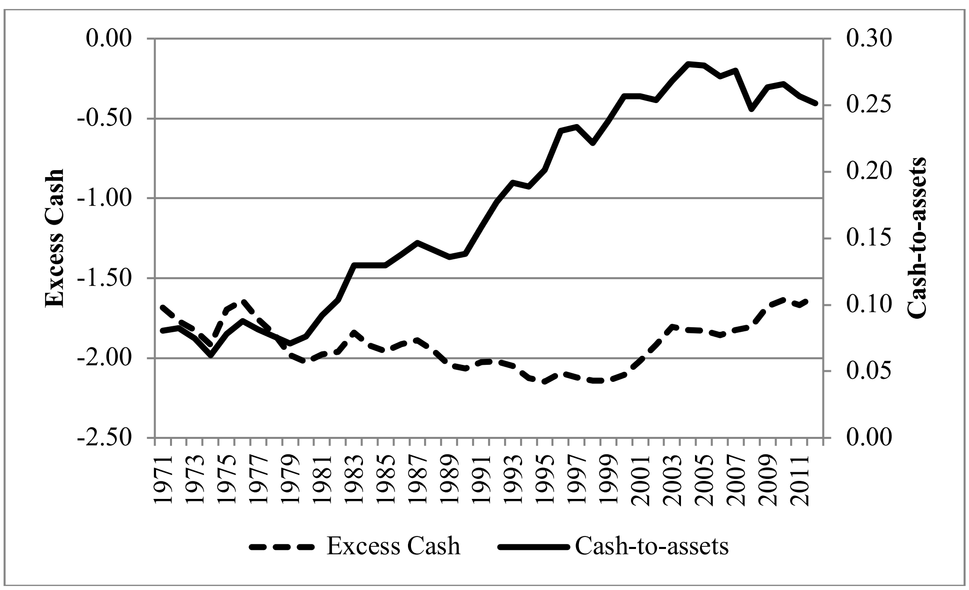
Figure IA.5

Excess cash of R&D firms using a pooled OLS regression

Panel A plots the number of R&D firms with positive or negative excess cash over the sample period. Panel B plots the yearly averages of *Excess Cash* and *Cash* for R&D firms with negative excess cash in a given year. *Excess Cash* is the residual estimated from a pooled OLS regression in which the dependent variable is the natural log of *Cash* divided by assets net of *Cash*. *MKTBK* is instrumented with firm's sales growth rate in year *t – 2*. The model controls for industry fixed effects and year fixed effects. Industry is defined at the two-digit SIC code level. *Cash* is the cash and cash equivalents. *R&D firms* are defined as firms with nonzero R&D capitalized stock in any given period. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. All variables are winsorized at the 1% and 99% percentiles.



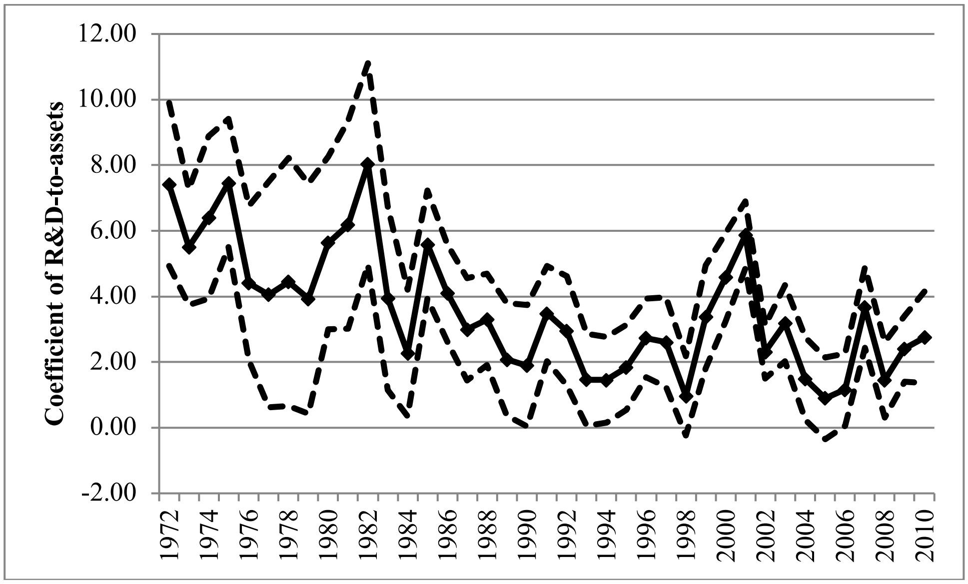
Panel A. Number of R&D firms with positive or negative excess cash



Panel B. Excess cash of R&D firms with negative excess cash

Figure IA.6
Falling market valuation of R&D

This figure shows the coefficient of *R&D*-to-assets from the cross-sectional yearly regression of Fama and French (1998) valuation model, wherein the dependent variable is *MKTBK*. Sample consists of all R&D firms from 1971 to 2012, a total of 81,559 firm-year observations. *R&D firms* are defined as firms with nonzero R&D capitalized stock in any given period. The sample starts with all nonfinancial and nonutility U.S. industrial firms with non-negative values for cash, sales, and common equity, and with lagged assets greater than $1 million in the Compustat database from 1971 to 2012. Dash lines indicate the 95% confidence intervals. Standard errors are heteroskedasticity consistent.



Variable definitions

SA index: The index is based on Hadlock and Pierce (2010) and calculated as {−0.737* Size} + {0.043*$Size^2$} – {0.040 × Age}, Size equals the log of inflation-adjusted book assets, and Age is the number of years the firm is listed with a nonmissing stock price on Compustat. In calculating this index, Size is winsorized at (the log of) $4.5 billion, and Age is winsorized at 37 years.
WW index: The index is based on Whited and Wu (2006) and calculated as {–0.091Cash Flow} – {0.062DividendPayer} + {0.021 Long-Term Debt} – {0.044Size} + {0.102Industry Growth} – {0.035Sales Growth}, where Long-Term Debt is long-term debt over assets, Industry Growth is the four-digit SIC industry sales growth rate, and Sales Growth is the firm sales growth rate. Other variables are defined above.
Corrected CF: Income before extraordinary items (IBC) plus extraordinary items and discontinued operations (XIDOC) plus depreciation and amortization (DPC) plus deferred taxes (TXDC) plus equity in net loss of unconsolidated subsidiaries (ESUBC) plus losses from the sale of PP&E (SPPIV) and funds from operations—other (FOPO) (Lewellen and Lewellen, 2014).
Negative Income: An indicator that equals one if a firm suffers operating loss (NI) in any given year.