THE OHIO STATE UNIVERSITY
FISHER COLLEGE OF BUSINESS

Fisher College of Business
Working Paper Series

Charles A. Dice Center for Research in Financial Economics

Firm-level Irreversibility

Hang Bai,
University of Connecticut

Erica X. N. Li,
Cheung Kong Graduate School of Business

Chen Xue,
University of Cincinnati

Lu Zhang,
The Ohio State University and NBER

Dice Center WP 2023-26
Fisher College of Business WP 2023-03-026

Firm-level Irreversibility

Hang Bai*
University of Connecticut

Erica X. N. Li†
Cheung Kong Graduate School of Business

Chen Xue‡
University of Cincinnati

Lu Zhang§
The Ohio State University and NBER

June 2023¶

Abstract

Contradicting Cooper and Haltiwanger (2006), Clementi and Palazzo (CP, 2019) report a largely symmetric investment rate distribution in Compustat, with a large fraction of negative investment rates, 18.2%, and conclude "no sign of irreversibility (p. 289)." CP's analysis is flawed. A data error on depreciation rates understates gross investment and shifts the whole gross investment rate distribution leftward. Nonstandard sample screens on age and acquisitions further curb its right tail, which is then truncated at 0.2. Fixing these problems restores the heavily asymmetric investment rate distribution with a fat right tail. The fraction of negative investment rates is small, only 4.9%–6.2%.

*School of Business, University of Connecticut, 2100 Hillside Road, Unit 1041F, Storrs, CT 06269. Tel: (510) 725-8868. E-mail: hang.bai@uconn.edu.

†Cheung Kong Graduate School of Business, 1 East Chang An Avenue, Oriental Plaza, Beijing 100738, China. Tel: 86-8537-8102. E-mail: xnli@ckgsb.edn.cn.

‡Lindner College of Business, University of Cincinnati, 2338 Lindner Hall, Cincinnati OH 45221. Tel: (513) 556-7078. E-mail: xuecx@ucmail.uc.edu.

§Fisher College of Business, The Ohio State University, 2100 Neil Avenue, Columbus OH 43210; and NBER. Tel: (614) 292-8644. E-mail: zhanglu@fisher.osu.edu.

¶We thank Stefan Nagel (the Editor), an anonymous Associate Editor, and three anonymous referees for extensive comments that have helped improve our work. We also thank Berardino Palazzo for sharing the codes and data used to generate Clementi and Palazzo's (2019) Figure 1 and Table 1. This study supersedes our prior work "Does costly reversibility matter for U.S. public firms?" circulated as NBER working paper #26372. In our companion paper titled "Asymmetric investment rates" (Bai, Li, Xue, and Zhang 2022, NBER working paper #29957), we have constructed firm-specific current-cost capital stocks for the entire Compustat universe. All remaining errors are our own.

1 Introduction

Initiated by Arrow (1968), a prominent theoretical literature on costly reversibility has long been established in the real options framework (Bernanke 1983; McDonald and Siegel 1986; Dixit and Pindyck 1994) and the neoclassical q-theory of investment (Abel 1983; Abel and Eberly 1994, 1996; Abel et al. 1996). The basic insight is that firms face higher costs in cutting than in expanding capital stocks, reducing negative investment and raising the hurdle for positive investment.

The available evidence on costly reversibility is mostly at the plant level. In a balanced panel with 7,000 large manufacturing plants from 1972 to 1988, Cooper and Haltiwanger (2006) document a heavily right-skewed investment rate distribution (Panel A in Figure 1), with a fraction of 10.4% for negative investment rates (below −1%) and 81.5% for positive investment rates (above 1%). Cooper and Haltiwanger write: "This striking asymmetry between positive and negative investment is an important feature of the data that our analysis seeks to match (p. 614)." Their structural estimation establishes a causal mechanism that links this asymmetry to costly reversibility.[1]

In a recent *Journal of Finance* article, CP claim to refute the Cooper-Haltiwanger plant-level conclusion at the firm level in Compustat. CP "start by documenting investment behavior among publicly traded U.S. firms (p. 282)" in an exercise that is "akin to that conducted by Cooper and Haltiwanger (2006) on manufacturing plants (p. 282)." CP's Figure 1 (cut-and-pasted as Panel B in Figure 1) shows a largely *symmetric* quarterly gross investment rate distribution in their 1978–2016 sample in Compustat, with a large fraction, 18.2%, of negative investment rates.

CP claim that upon "being hit by adverse profitability shocks, large public firms have ample latitude to divest their least productive assets (p. 281)." "[E]ach quarter on average 18.2% of firms record negative gross investment. We take the latter as strong evidence against the assumption of ir-

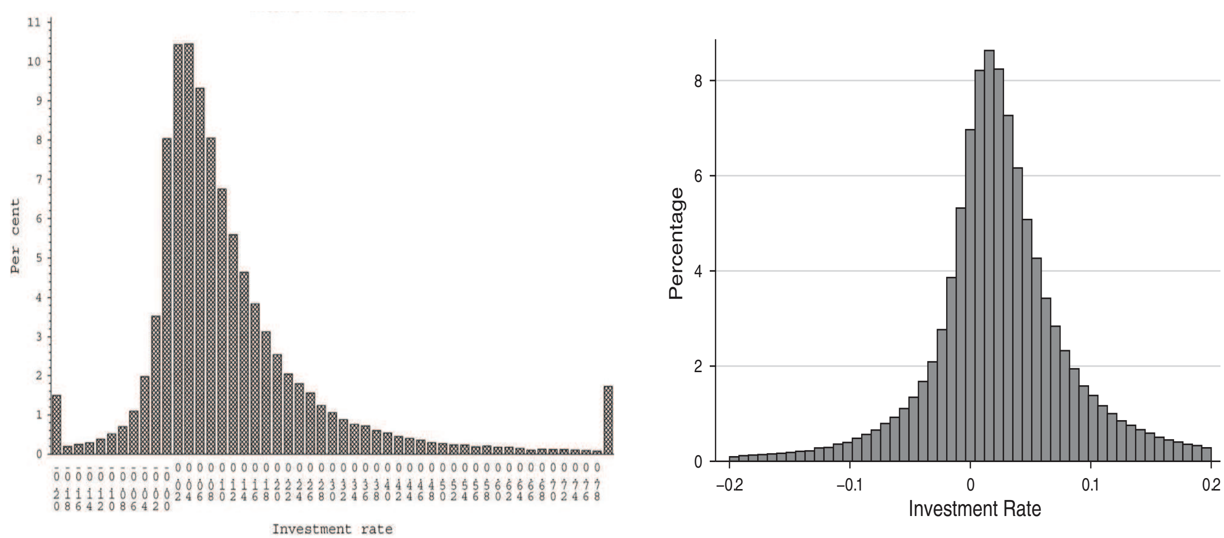
[1]Other plant-level studies on lumpy and irreversible investment include Caballero, Engel, and Haltiwanger (1995), Doms and Dunne (1998), Cooper, Haltiwanger, and Power (1999), Caballero (1999), and Nilsen and Schiantarelli (2003). In our companion paper, Bai et al. (2022) extend the Cooper-Haltiwanger evidence to the firm level in Compustat. We show that the firm-level current-cost investment rate distribution is heavily right-skewed, with a small fraction of negative investment rates, 5.51%, but a huge fraction of positive investment rates, 91.64%.

Figure 1 : A Tale of Two Gross Investment Rate Distributions

Cooper and Haltiwanger's (2006) Figure 1 shows the plant-level (annual) gross investment rate distribution in 7,000 manufacturing plants in their 1972–1988 balanced panel from Longitudinal Research Database (Panel A). CP's Figure 1 shows the firm-level (quarterly) gross investment rate distribution in their Compustat sample from the first quarter of 1978 (1978Q1) to the fourth quarter of 2016 (2016Q4) (Panel B).

Panel A: Cooper and Haltiwanger's Figure 1

Panel B: CP's Figure 1



reversibility (p. 282)." "[C]apital accumulation at public firms is likely to be very different from that emerging from the analysis of a representative sample of manufacturing establishments (p. 285)." "While comprehensive, our study emphasizes features, such as the volatility and reversibility of investment (p. 285)." The 18.2% estimate means that "plenty of firms downsize, at all times (p. 287)." For U.S. public firms, "investment displays substantial volatility and no sign of irreversibility. In fact, in each quarter a large fraction of firms reduce their deployment of plant, property, and equipment (p. 289)." "[T]he irreversibility assumption has no empirical support (p. 303)."

The Replication Network defines a replication broadly "as any study whose primary purpose is to establish the correctness of a previous study" but accepts "many gradations of replications, stretching from pure reproduction of key finding(s) of a previous study; to checking the robustness of those findings to changes in data, estimation procedure, model specification, etc."[2] More specifi-

[2]https://replicationnetwork.com/why-replications/

cally, we distinguish reproduction, replication, and reanalysis. Reproduction means redoing a prior study in exactly the same way as in Hamermesh's (2007) pure replication. We treat replication as his scientific replication, which means "different sample, different population, and perhaps similar, but not identical model (p. 716)." We take reanalysis to be Clemens's (2017) reanalysis and extension tests, which materially alter the model specifications of the original study and use new data.

In Section 2, we reproduce CP's empirics and identify three design issues and at least five discrepancies in reporting, coding, and data. First, CP measure gross investment rates as the net growth rates of net property, plant, and equipment (PPE) plus depreciation rates. However, instead of accounting depreciation rates embedded in net PPE, CP add back geometric depreciation rates from Bureau of Economic Analysis (BEA). Because the BEA rates are lower, CP underestimate gross investment flows and shift the whole investment rate distribution leftward. This data error gives rise to a high fraction of negative investment rates, 18.2%. Fixing this error reduces it to only 6.08%. This data error also curbs the right tail of the gross investment rate distribution. The reason is that accounting depreciation rates are much more right-skewed than the BEA rates.

CP further restrict the right tail of the gross investment rate distribution with two nonstandard sample criteria: (i) dropping the first 12 quarters for each firm; and (ii) dropping firm-quarters associated with acquisitions larger than 5% of assets. Removing the former screen raises the investment rate skewness from 2.29 to 3.39, and eliminating the latter screen from 2.29 to 3.46.

We further identify at least five discrepancies between CP's reporting, coding, and data: (i) imposing lifetime exchange code restrictions; (ii) requiring lifetime net PPE; (iii) merging investment rates with returns; and (iv) deflating net PPE. Discrepancy (iv) shifts the entire investment rate distribution further leftward. Adjusting for all four allows us to reproduce CP's Table I closely.

Despite all the issues and discrepancies, the right tail of the investment rate distribution still survives, albeit weakly. In discrepancy (v), CP clean up whatever remains at the right tail by cutting it off at 0.2 to arrive at their largely symmetric Figure 1. After failing to reproduce Figure 1, we have requested and obtained CP's codes and data for their Table I and Figure 1. The Stata codes in question are: "gen hist = inv_rate," "replace hist = . if hist > 0.2," and "replace hist = . if hist < −0.2."[3]

CP's work is messy. Their Figure 1 is based on a different sample from Table I. The key discrepancies between the two samples are: (a) not requiring share code to be 10 or 11; (b) not requiring exchange code to be 1, 2, or 3; (c) not merging with monthly stock returns; and (d) not requiring book-to-market.[4] As a result, their Figure 1 sample is 23% larger than their Table I sample (364,429 vs. 296,226 firm-quarters). In the Figure 1 sample, the investment rate skewness is 2.08, and the fraction of negative investment rates is 19.96% (2.18 and 18.24% in the Table I sample, respectively).

In their Figure 1 sample, the truncated right tail (> 0.2) contains 14,867 firm-quarters (4.08% of the sample), with a mean of 36.21% and a median of 29.72%. The truncated left tail (< -0.2) contains 5,340 firm-quarters (1.47% of the sample), with a mean of −30.88% and a median of −28.98%. Because the truncation cuts more in the right tail, it reduces the investment rate skewness to −0.08. CP do not report the skewness of 2.18 in Table I, but their Figure 1 is weakly *left*-skewed.[5]

In Section 3, we replicate CP's Table I and Figure 1. Within CP's empirical setup, after we fix all their design issues and discrepancies, the investment rate skewness rises from 2.18 to 4.8, and the negative investment fraction drops from 18.24% to 5.89%. After we further adjust for small differences in sampling criteria and treatment for outliers (for more reliable estimates), the skewness is 3.53, and the negative investment fraction 6.17%. With gross PPE (not net PPE) as the deflator,

[3]See our second annotation on p. 9 in the Internet Appendix D.1 on CP's "investment_rate_bea.do."

[4]See our first annotation on p. 9 in the Internet Appendix D.1 on CP's "investment_rate_bea.do."

[5]While leaving the nature of CP's practices to the reader to decipher, we note from John, Loewenstein, and Prelec (2012): "Questionable research practices (QRPs), such as excluding data points on the basis of post hoc criteria, can spuriously increase the likelihood of finding evidence in support of a hypothesis (p. 524)." "QRPs are the steroids of scientific competition, artificially enhancing performance and producing a kind of arms race in which researchers who strictly play by the rules are at a competitive disadvantage (p. 524)." In "Federal Policy on Research Misconduct," the Office of Science and Technology Policy (2000) defines research misconduct as "fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results." Specifically, "[f]alsification is manipulating research materials, equipment, or processes, or changing or omitting data or results such that the research is not accurately represented in the research record." In "Best Practice Guidelines on Research Integrity and Publishing Ethics," Wiley states: "Data fabrication is the intentional misrepresentation of research data by making-up findings, recording, or reporting of results. Data falsification is the manipulation of research materials, equipment, or processes, including omitting and changing data, with the intention of giving a false impression. Changes to images can create misleading results when research data are collected as images. Inappropriate image manipulation is one form of fabrication or falsification that journals can identify" (https://authorservices.wiley.com/ethics-guidelines/index.html).

the investment rate skewness remains at 3.59, and the negative investment fraction falls to 4.94%. All three replications produce a heavily asymmetric investment rate distribution with a fat right tail.

In Section 4, we conduct a reanalysis on the baseline investment model. Estimates from simulated method of moments strongly indicate costly reversibility and operating leverage, which are a good start to explaining the average value premium and investment moments. Section 5 concludes.

A separate Internet Appendix furnishes supplementary results and CP's codes (annotated by us). The discrepancies between CP's reporting, coding, and data are only visible in their codes (not described in their paper). Because the discrepancies are essential for us to closely reproduce their Table I and Figure 1, we opt to include the annotated codes in our Internet Appendix.

2 Reproduction

Reproducing CP's Table I and Figure 1, we identify three design issues in their empirical procedure in Section 2.1 and at least five discrepancies between their reporting and coding in Section 2.2.

2.1 Three Procedural Issues

The column "Table I" in Panel A of Table 1 shows the quarterly investment rate moments in CP's Table I. The column "CP's data" shows the moments that we calculate with their Table I sample. While emphasizing the fraction of negative investment rates of 18.2%, CP do not report the skewness of 2.18, the 5th percentile of -6.85%, the median of 2.46%, or the 95th percentile of 16.81%. These moments show that the investment rate distribution is already right-skewed in their data.

Panel B in Table 1 shows our reproduction based on CP's description of their procedures in their paper. The column denoted "moments" shows that we get close but not exact. Our reproduction sample contains 379,923 firm-quarters. The sample size is 28.25% larger than 296,226 in CP's data. In Panel C, we identify four discrepancies between CP's description and codes. As we detail in Section 2.2, after adjusting for these four discrepancies, we reproduce CP's Table I almost exactly.

Table 1 : Reproduction of CP's Table I on Quarterly Investment Rate Moments, 1978Q1–2016Q4

In Panel A, the column "Table I" reports CP's Table I. The column "CP's data" shows the moments on CP's Table I dataset. Panel B shows our reproduction based on CP's procedure described in their paper. The column "moments" shows our reproduction of the moments. The remaining three columns show comparative statics by changing one aspect of CP's procedure, while keeping all others unchanged: (i) Using accounting depreciation rates (acct. δ); (ii) keeping the first 12 quarterly observations (no age>3); and (iii) not dropping firm-quarters with acquisitions larger than 5% of total assets (with M&A). Panel C shows our reproduction based on CP's description and codes, with four discrepancies not described in their paper: (i) imposing lifetime exchange code restrictions (lifetime ex. code); (ii) requiring lifetime PPE data (lifetime PPE); (iii) requiring stock returns when merging with investment rates (require returns, losing 1978Q1–Q2 and 2016Q3–Q4); and (iv) deflating PPE data (deflate PPE). The last column "all" incorporates all four discrepancies.

	Panel A: CP		Panel B: Reproduction per CP's description			
	Table I	CP's data	moments	acct. δ	no age>3	with M&A
#Firm-quarters	296,218	296,226	379,923	364,234	471,943	391,808
Mean	3.5	3.47	3.92	7.46	5.03	4.54
Standard deviation	9.5	9.54	9.94	11.96	12.62	11.70
Skewness		2.18	2.29	3.37	3.39	3.46
Autocorrelation	26.2	26.22	25.43	33.84	27.08	19.38
Negative investment	18.2	18.24	16.91	6.08	16.98	16.54
Inaction rate	16.5	16.53	14.56	9.14	13.78	14.24
Positive spikes	3.8	3.77	4.22	7.90	6.30	5.16
Negative spikes	1.2	1.18	1.21	0.78	1.24	1.19
5th percentile		−6.85	−6.90	−1.98	−6.99	−6.78
Median		2.46	2.81	4.96	3.04	2.92
95th percentile		16.81	18.00	25.70	23.21	20.44

	Panel C: Reproduction per CP's description and codes				
	lifetime ex. code	lifetime PPE	require returns	deflate PPE	all
#Firm-quarters	346,710	325,746	372,003	379,923	296,185
Mean	3.89	3.92	3.92	3.51	3.52
Standard deviation	9.44	9.86	9.92	9.91	9.44
Skewness	2.16	2.27	2.30	2.31	2.18
Autocorrelation	26.00	25.89	25.53	25.34	26.28
Negative investment	16.27	16.87	16.95	18.98	18.33
Inaction rate	14.52	14.37	14.58	16.44	16.25
Positive spikes	3.96	4.18	4.21	4.07	3.87
Negative spikes	1.10	1.19	1.20	1.24	1.13
5th percentile	−6.55	−6.88	−6.90	−7.27	−6.98
Median	2.85	2.84	2.82	2.41	2.49
95th percentile	17.39	17.91	17.98	17.54	17.05

The remaining three columns in Panel A identify three conceptual issues in CP's procedure and quantify their respective impact on CP's results. These issues are: (i) Combining BEA depreciation rates with accounting net investment rates; (ii) dropping the first 12 quarters for each firm; and (iii) dropping firm-quarters with acquisitions larger than 5% of total assets.

2.1.1 Combining BEA Depreciation Rates with Accounting Net Investment Rates

CP measure capital stock as "item $PPENTQ$, defined as the net value of property, plant, and equipment. Net quarterly investment is the difference between two consecutive values of this variable (p. 286)." CP acknowledge: "Our convention amounts to assuming that *accounting depreciation is an accurate proxy for economic depreciation* (p. 286, our emphasis)."

However, in the same paragraph, CP go on to say: "The gross investment rate is set equal to

$$\frac{PPENTQ_t - PPENTQ_{t-1}}{PPENTQ_{t-1}} + \delta_j$$

where δ_j is *the average depreciation rate of industry j estimated using data from the U.S. Bureau of Economic Analysis (BEA)* (p. 286, our emphasis)." As such, CP mix up the depreciation rates in the same paragraph, giving rise to a flawed gross investment rate measure. Because accounting depreciation rates are embedded in net PPE (as CP acknowledge), one should add back accounting, not BEA's geometric, depreciation rates when calculating gross investment rates.

We are not saying that accounting depreciation rates are better than BEA's geometric depreciation rates as measures of economic depreciation rates. Rather, we are insisting on logical consistency: Capital stocks and depreciation rates in CP's gross investment rates are *not* independent of each other. If one uses net PPE as capital (as CP do), to calculate gross investment rates, one must add back accounting depreciation rates embedded in net PPE via the capital accumulation equation. Financial accountants use accounting depreciation, together with investment flows, to calculate net PPE per the capital accumulation equation. Alternatively, if one insists on using economic depreciation rates, one must reconstruct capital stocks with investment flows that are obtained elsewhere.[6]

Most U.S. firms use straight-line depreciation for financial accounting (Wahlen, Baginski, and Bradshaw 2018, p. 506), whereas BEA estimates geometric depreciation rates (Hulten and Wykoff 1981; Fraumeni 1997). From Figure 2, accounting and geometric depreciation rates are quite different. Panel A shows the average quarterly BEA industry-level depreciation rates assigned to the firm level based on the NAICS or SIC codes.[7] The distribution is largely symmetric, ranging from 0.57% to 4.42% per quarter, with a mean of 2.51% and a median of 2.33%.

Panel B shows the quarterly accounting depreciation rates, measured as the amount of depreciation and amortization (item DPQ) minus the amortization of intangibles (annual item AM divided by four, zero if missing; the quarterly version of item AM is unavailable), scaled by item PPENTQ. The linear interpolation works because of the straight-line depreciation. The accounting rate distribution is much more dispersed, ranging from 0% to more than 35%, and heavily right-skewed. Its mean is 6.09%, and 5th, 50th, and 95th percentiles are 1.48%, 4.27%, and 16.54%, respectively.

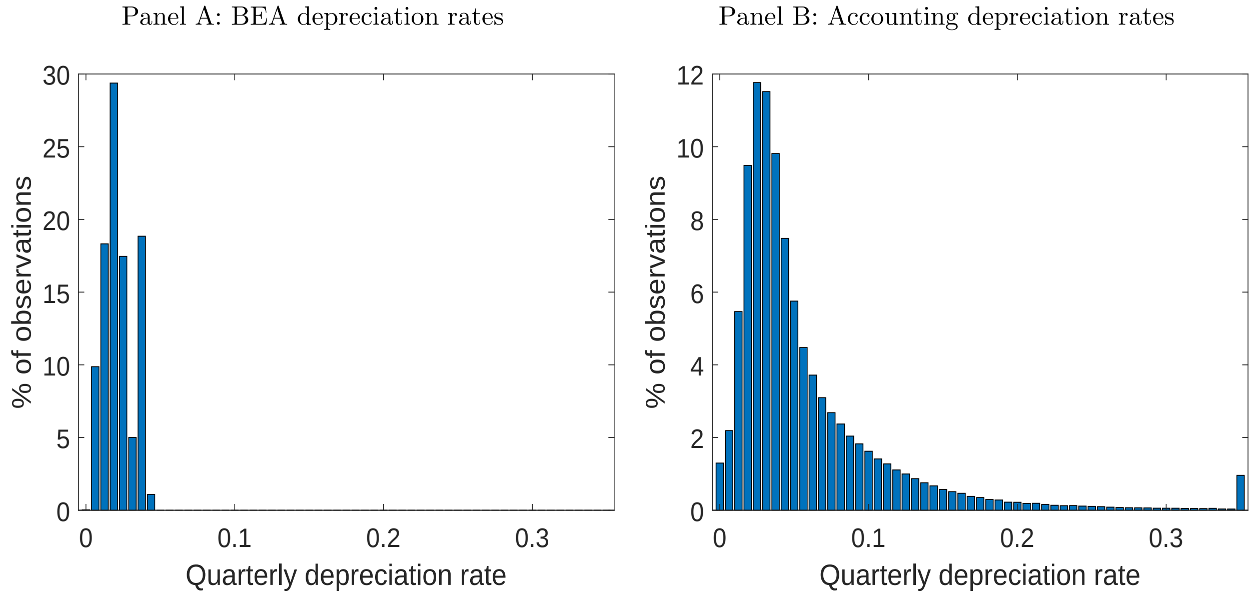
Most important, by adding the lower BEA depreciation rates back to the net growth rates of net PPE, which embeds higher accounting depreciation rates, CP erroneously shift the whole gross investment rate distribution leftward, giving rise to a higher fraction of negative investment rates, 18.2%. This data error paints a misleading picture of investment reversibility in U.S. public firms. Fixing this data error reduces the negative investment fraction to only 6.08% (column "acct. δ"

[6] From the capital accumulation equation, one can take capital as given to measure investment, or take investment as given to measure capital. In Bai et al. (2022), we solve this chicken-or-egg problem by measuring gross investment as the sum of net investment (change in net PPE) and accounting depreciation. Although net PPE and accounting depreciation are suboptimal measures of capital and economic depreciation, respectively, their combination accurately measures gross investment (Bai et al., Section 3.1). We then combine this gross investment with BEA's economic depreciation rates (and price deflators) to construct current-cost capital stocks via the perpetual inventory method.

[7] The data on the BEA's average industry-level depreciation rates used in CP are from Berardino Palazzo. These depreciation rates (constant over time) are for the 2- or 3-digit NAICS industries. We assign firms to the NAICS industries based on their NAICS codes from Compustat (item NAICSH). When NAICS codes are not available (especially prior to June 1985), we use SIC codes (item SICH) and convert them into NAICS codes via the mapping tables from the Census Bureau. Because the conversion from SIC to NAICS codes might not be unique, there exist multiple assigned NAICS industries for a given firm. In these cases, we use the average depreciation rate across the assigned industries. We do not find CP's documentation in their paper to be clear on how they process BEA's raw data to obtain their depreciation rates or how they apply NAICS-based industry depreciation rates to SIC industries. In Bai et al. (2022), we construct our own industry-specific, time-varying economic depreciation rates from BEA. However, for our reproduction purpose, we opt to use CP's depreciation rates data directly to ease comparison.

Figure 2 : The Distributions of Quarterly BEA and Accounting Depreciation Rates, 1978:Q1–2016:Q4

Panel A shows the histogram of BEA depreciation rates assigned to the firm level in our reproduction sample (column "moments" in Table 1). Panel B shows the histogram of accounting depreciation rates in our reproduction sample with accounting depreciation (column "acct. δ" in Table 1).



in Panel B of Table 1). Also, by adding back the largely symmetric BEA depreciation rates, instead of the right-skewed accounting depreciation rates, CP weaken the right-skewness of the gross investment rate distribution. Fixing this data error raises its skewness from 2.18 to 3.37.

2.1.2 Dropping the First 12 Quarters for Each Firm

In column "moments" in Panel B of Table 1, we impose CP's sample criteria: (i) Excluding financial firms, utilities, and unclassified firms; (ii) dropping the first 12 quarters for each firm;[8] (iii) dropping firm-quarters associated with acquisitions larger than 5% of total assets; (iv) discarding firm-quarters in the top and bottom 0.5% of the pooled distribution of quarterly investment rates; and (v) dropping firm-quarters with missing values of investment rates or book-to-market.

[8]CP exclude "companies that have fewer than 12 quarters of data (p. 286)." This sentence might be interpreted differently, but we verify in their codes that CP drop the first 12 quarters for each firm. See our first annotation on p. 8 in the Internet Appendix D.1 on CP's "investment_rate_bea.do."

Criterion (ii) and (iii) are not standard. Column "no age>3" in Panel B shows that removing the age screen (adding back the first 12 quarters for every firm) raises the investment rate skewness from 2.29 (column "moments") to 3.39. As such, the impact of the age screen is to curb the right tail of the investment rate distribution. Intuitively, younger firms (that issue equity) invest more than older firms (Lyandres, Sun, and Zhang 2008).

2.1.3 A Merger&Acquisition (M&A) Screen at 5% of Total Assets

It is stringent to impose criterion (iii) that drops firm-quarters associated with acquisitions larger than 5% of total assets. CP's is the only paper (to our knowledge) that uses the 5% cutoff. In asset pricing, it is customary to keep M&As, which are not random corporate events. Firms with M&As tend to be growth firms, momentum winners, high investment firms, and high profitability firms, and firms without M&As tend to be value firms, momentum losers, low investment firms, and low profitability firms. In corporate finance, it might be informative to separate internal from external growth, but the common cutoff is 15% of assets (Whited 1992), as opposed to CP's 5%.[9]

Column "with M&A" in Panel B of Table 1 shows the impact of removing CP's 5% M&A screen. The investment rate skewness rises from 2.29 (column "moments") to 3.46. (In untabulated results, with the 15% M&A screen, the investment rate skewness is 2.53.) As such, similar to the age screen, the impact of the 5% M&A screen is to curb the right tail of the investment rate distribution.

2.2 Discrepancies

We further document four discrepancies between CP's reporting and coding in order to closely reproduce their Table I and several more discrepancies to fully reproduce their Figure 1. All the discrepancies are implemented in their codes but not described in their paper.

[9]A subtlety arises because the common 15% cutoff is on annual data, but CP's 5% is on quarterly data. However, acquisitions are not smoothly distributed across quarters within a year but are instead lumpy (concentrating within a single quarter). In the 1978–2016 Compustat sample with nonmissing annual acquisitions (item AQC) and quarterly acquisitions (computed from year-to-date item AQCY), for a median firm-year with acquisitions, 98.5% of the annual acquisition amount occurs within one quarter. The 15% annual cutoff excludes 7,637 firm-years. For comparison, the 15% quarterly cutoff excludes 6,511 firm-quarters, while the 5% quarterly cutoff excludes 16,750 firm-quarters. As such, the 5% quarterly cutoff is indeed a more stringent M&A screen than the common 15% annual cutoff.

2.2.1 Closely Reproducing Table I

As noted, our reproduction sample in column "moments" in Panel B is 28.26% larger than CP's Table I data. Panel C identifies four discrepancies to reconcile the differences (column "all").

First, CP require a firm to be listed and traded continuously on one of the three major exchanges (NYSE, Amex, and NASDAQ) throughout its life to be included in their sample.[10] If a firm is ever suspended in trading, its entire history would be excluded from their sample. In column "moments" we keep the firm's available data during its normal trading periods. Also, if a firm starts small, first listed on, say, Philadelphia Stock Exchange, and only later moves to NYSE, such a firm would be excluded from CP's sample altogether. We instead keep the firm's data on NYSE. As shown in column "lifetime ex. code," their screen loses 8.74% of firm-quarters (from 379,923 to 346,710).

Second, CP go through an empty capital accumulation recursion to build net PPE by accumulating past changes in net PPE. This recursion is redundant (current net PPE is directly available in Compustat).[11] However, the recursion requires all historical data to be available, thereby losing 14.26% of firm-quarters from 379,923 to 325,746 in column "lifetime PPE."

Third, CP merge investment rates with monthly CRSP stock returns from 1979 to 2016 with a two-calendar-quarter lag. Specifically, CP merge investment rates from fiscal quarters ending in calendar quarter q with monthly stock returns from calendar quarter $q+2$. Doing so loses investment rates from the first two quarters of 1978 and the last two quarters of 2016, which amount to 2.08% of firm-quarters (column "require returns").[12]

Fourth, CP adjust for inflation by deflating net PPE with an aggregate (nonresidential fixed assets) price deflator from BEA. Doing so does not change the sample size but reduces the mean investment rate from 3.92% to 3.51%.[13] Doing so also shifts the entire investment rate distribution further leftward. As a result, the fraction of negative investment rates increases from 16.91%

[10] See our annotation on p. 1 in the Internet Appendix D.3 on CP's "crsp_cleaning.do."
[11] See our annotation on p. 6 in the Internet Appendix D.1 on CP's "investment_rate_bea.do."
[12] See our annotation on p. 3 in the Internet Appendix D.5 on CP's "inv_rate_moments.do."
[13] See our annotation on p. 5 in the Internet Appendix D.1 on CP's "investment_rate_bea.do."

(column “moments” in Panel B) to 18.98% (column “deflate PPE” in Panel C).

Finally, column “all” in Panel C of Table 1 adjusts for all four discrepancies simultaneously and reproduces column “CP’s data” almost exactly. Our reproduction (based on CP’s description and codes) in column “all” has a sample size of 296,185 firm-quarters, close to 296,226 in CP’s data. As for the investment rate distribution, our mean is 3.52% relative to CP’s 3.47%. Our standard deviation is 9.44% relative to CP’s 9.54%. Our skewness is 2.18, which is identical to CP’s. Our fraction of negative investment rates is 18.33%, which is close to CP’s 18.24%.

2.2.2 The (Missing) Right Tail in CP’s Figure 1

Figure 3 quantifies the impact of the three design issues identified in Section 2.1 on the right tail of the investment rate distribution in CP’s Figure 1. Comparing Panel A (based on column “acct. δ” in Table 1) with CP’s Figure 1 (Panel B in our Figure 1) shows that using BEA depreciation rates has the largest impact on curbing the right tail of the investment rate distribution. From Panels B and C, imposing the 3-year age screen and the 5% M&A screen further curbs the right tail.

Figure 4 attempts to reproduce CP’s Figure 1 on the investment rate distribution. Even after imposing the three design issues and the four discrepancies identified in Section 2.2, the right tail of the investment rate distribution still survives, albeit weakly. CP clean up whatever remains at the right tail by cutting it off at 0.2 to arrive at their largely symmetric Figure 1.

As noted, the investment rate skewness is 2.18 in CP’s Table I data and 2.08 in CP’s Figure 1 data. Neither is reported in their paper. However, the skewness silently communicated in their truncated Figure 1 is weakly negative, -0.08, which, if true, would be a shocking refutation of the prior consensus (Cooper and Haltiwanger 2006). Without being alert of this truncation, a reader might be (mis)led into thinking that CP’s Table I and Figure 1 are from the same data.

This discrepancy (truncating at 0.2) might be questionable (footnote 5). On the one hand, CP position their work as “documenting investment behavior among publicly traded U.S. firms (p. 282)” in an exercise that is “*akin to that conducted by Cooper and Haltiwanger (2006) on manufac-*

Figure 3 : The Impact of BEA Depreciation Rates, the Age Screen, and the M&A Screen on CP's Figure 1 on the Gross Investment Rate Distribution, 1978Q1–2016Q4

Panels A–C are based on our data in columns "acct. δ," "no age>3," and "with M&A" in Table 1, respectively.

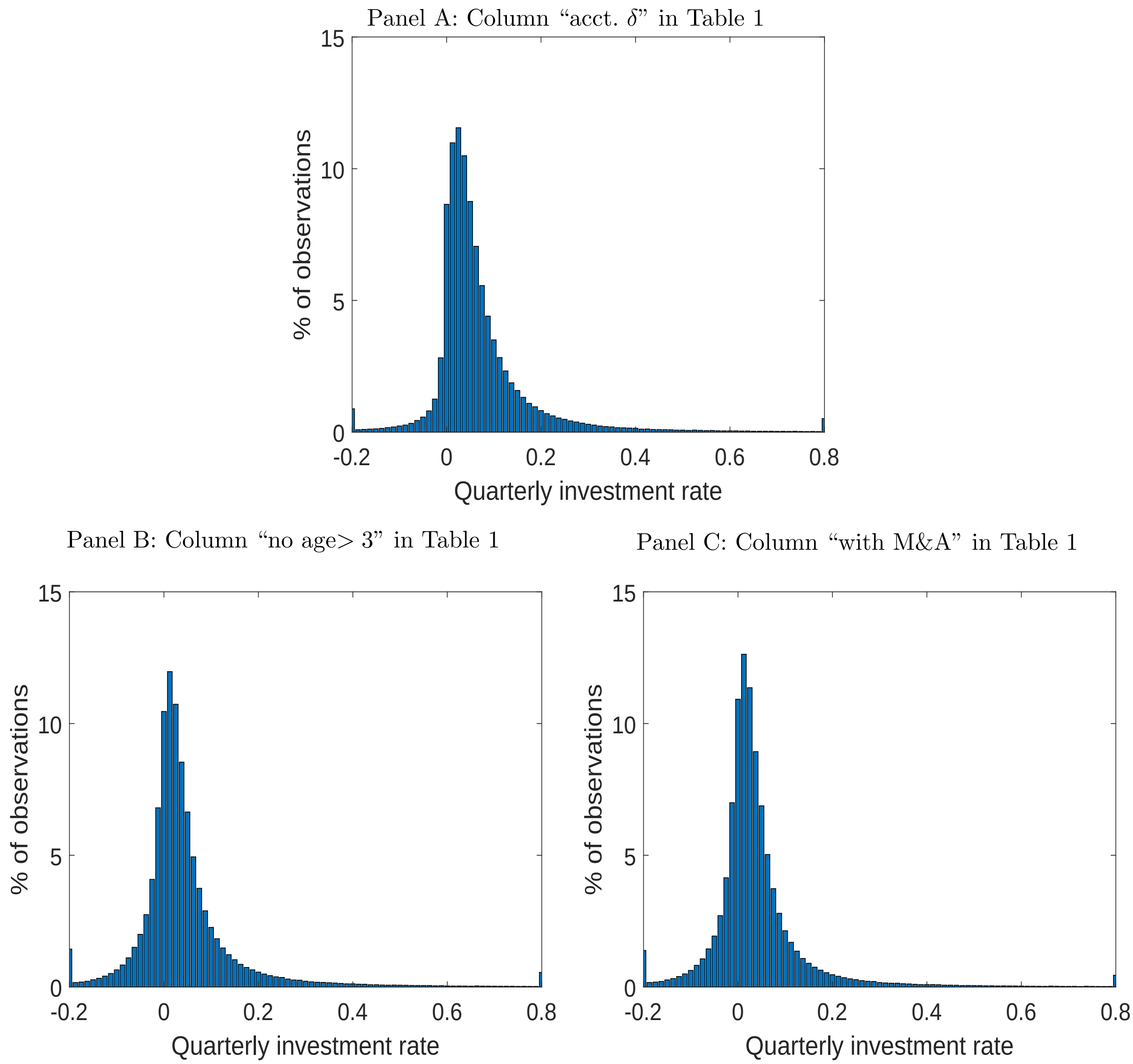
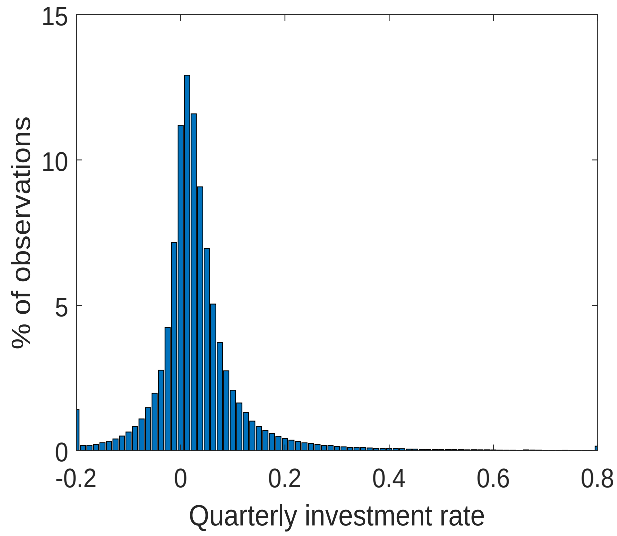


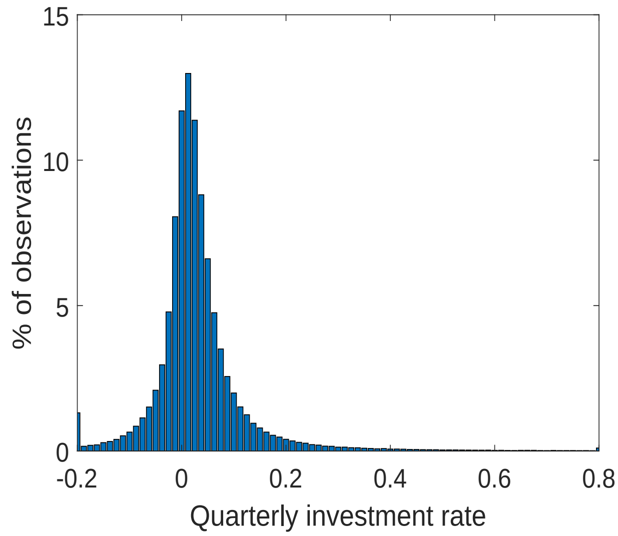
Figure 4 : Reproduction of CP's Figure 1 on the Gross Investment Rate Distribution, 1978Q1–2016Q4

Panels A and B are based on our reproduction samples in columns "moments" and "all" in Table 1, respectively. Panels C and D are based on CP's data in their Table I and Figure 1, respectively.

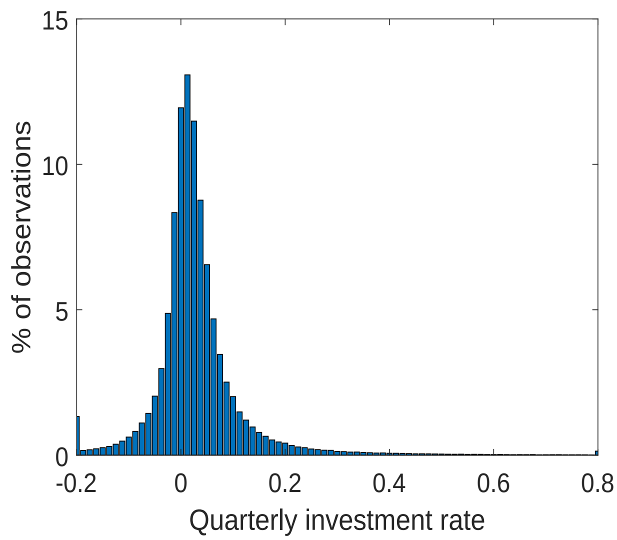
Panel A: Column "moments" in Table 1



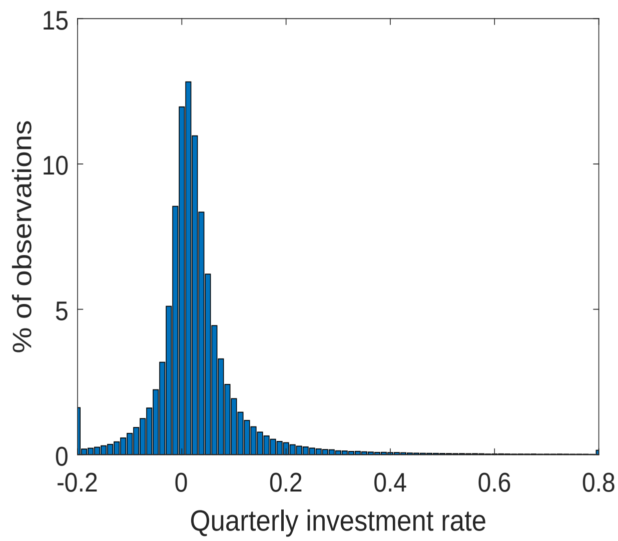
Panel B: Column "all" in Table 1



Panel C: Column "CP's data" in Table 1



Panel D: CP's Figure 1 data without truncation



turing plants (p. 282, our emphasis)." "Studies of the plant-level investment process such as Doms and Dunne (1998) and Cooper and Haltiwanger (2006) provide empirical evidence needed to discipline quantitative studies on the role of cross-sectional heterogeneity in macroeconomic models. In an analogous fashion, in this section we carefully describe investment at U.S. public companies for the purpose of informing modeling choices in the quantitative analysis of production-based asset pricing models (p. 285)." As noted, CP conclude "*no sign of irreversibility* (p. 289, our emphasis)."

On the other hand, Cooper and Haltiwanger (2006) reiterate, like a broken record, that the fat right tail of the investment rate distribution is the "smoking gun" on irreversibility:

"It is transparent that the investment rate distribution is non-normal having a considerable mass around 0, fat tails, and is highly skewed to the right (standard tests for non-normality yield strong evidence of skewness and kurtosis) (p. 614)."

"These properties of the investment distribution illustrates a key feature of the micro-data: investment rates are highly asymmetric (p. 614)."

"This striking asymmetry between positive and negative investment is an important feature of the data that our analysis seeks to match (p. 614)."

"[T]he investment distribution at the micro-level is very asymmetric and has a fat right tail (p. 616)."

"While Table 1 shows some range of inaction, the more robust finding in Figure 1 and Table 1 is that the distribution of investment is skewed and kurtotic with a fat right tail (p. 616)."

"All of the models are able to produce both positive and negative spikes but, naturally, the asymmetry in spike rates is most prominent in the irreversibility specification (p. 620)."

"Further, with the non-convex adjustment and the irreversibility, the model produces both positive and negative investment bursts of the frequency found in the data (p. 623)."

"The average 90th percentile from the LRD is 0.299 and the 10th percentile is given by -0.014. These moments capture the asymmetry and fat right tail of the investment distribution (p. 627)."

"The LRD indicates that plants exhibit periods of inactivity as well as large positive investment bursts but little evidence of negative investment. The resulting distribution of investment rates at the micro-level is highly skewed even though the distribution of shocks is not. A model, which incorporates both convex and non-convex aspects of adjustment, including irreversibility, fits these observations best (p. 629)."

"In the actual micro-data, while we do observe some range of inaction as we report in Table 1, the more robust finding is that the distribution of investment is skewed and kurtotic with a mass around 0 and a fat right tail (p. 630)."

3 Replication

We perform a scientific replication of CP's Table I, by fixing all the procedural issues and discrepancies identified in Section 2. Table 2 shows the details. We start with our reproduction in column "moments" in Table 1 that already adjusts for the discrepancies between CP's reporting and coding.

3.1 Quantifying the Cumulative Impact of the Three Design Issues

First, we fix CP's data error in the gross investment rates by changing BEA depreciation rates to accounting depreciation rates. Column "acct. δ" in Table 1 (the same column in Table 2) already performs this step. As in Tang (2009), our gross investment rate is $(\text{PPENTQ}_{it} - \text{PPENTQ}_{it-1})/\text{PPENTQ}_{it-1} + \delta_{it}$, in which PPENTQ is item PPENTQ, and δ_{it} is the accounting

Table 2 : Replication of CP's Table I on Quarterly Investment Rate Moments, 1978Q1–2016Q4

Column "moments" is the same in Table 1 after adjusting for all the discrepancies between CP's reporting and coding. Column "acct. δ" is the same column in Table 1 that fixes the depreciation rate issue in column "moments." Column "no age>3" removes the 12-quarter age screen from column "acct. δ." Column "with M&A" removes the 5% M&A screen from column "no age>3." Column "others" adjusts for other differences in sampling criteria from column "with M&A." CP drop financials, utilities, and unclassified firms, while we drop financials, firms with negative book equity, firm-quarters with negative or zero assets, net PPE, or sales. CP also require nonmissing book-to-market, while we do not. From column "others," column "1%–99%" changes CP's 0.5%–99.5% truncation to 1%–99% winsorization. Finally, column "/gross PPE" perturbs column "1%–99%" by changing the scaler from net PPE to gross PPE.

	moments	acct. δ	no age> 3	with M&A	others	1%–99%	/gross PPE
#Firm-quarters	379,923	364,234	460,050	475,788	504,692	509,788	378,122
Mean	3.92	7.46	9.09	9.96	9.61	9.89	5.03
Standard deviation	9.94	11.96	15.67	18.13	17.64	18.14	9.48
Skewness	2.29	3.37	4.18	4.80	4.80	3.53	3.59
Autocorrelation	25.43	33.84	35.55	29.69	30.22	31.73	34.52
Negative investment	16.91	6.08	6.00	5.89	5.76	6.17	4.94
Inaction rate	14.56	9.14	8.87	8.65	8.65	8.57	18.08
Positive spikes	4.22	7.90	10.94	12.10	11.58	11.92	4.82
Negative spikes	1.21	0.78	0.83	0.82	0.79	1.02	0.12
5th percentile	−6.90	−1.98	−1.96	−1.86	−1.76	−2.31	−1.10
Median	2.81	4.96	5.39	5.57	5.29	5.29	2.61
95th percentile	18.00	25.70	32.67	36.23	35.20	37.64	19.50

depreciation rate, calculated as the amount of depreciation and amortization (item DPQ) minus the amortization of intangibles (item AM divided by four, zero if missing), scaled by net PPE.[14]

Second, in column "no age> 3" we remove the 12-quarter age screen from column "acct. δ" to its left in Table 2, thereby showing the cumulative impact of fixing the first two design issues in CP. (In contrast, the columns in Table 1 only show the separate impact of each as in comparative statics.) Removing the age screen raises the sample size by 26.31% from 364,234 to 460,050 firm-quarters. The investment rate skewness rises from 3.37 to 4.18, and the 95th percentile from 25.7% to 32.67%. The negative investment fraction drops slightly from 6.08% to 6%.

Third, in column "with M&A" we further remove the 5% M&A screen from column "no age>3" to its left in Table 2, showing the cumulative impact of fixing all three design issues in CP. The sam-

[14] In terms of timing, for 2002Q2, for example, we take period-t stock variables from the balance sheet for 2002Q1 and period-t flow variables from the income or cash flow statement for 2002Q2.

ple size increases further by 3.42% to 475,788 firm-quarters. The investment rate skewness goes up further to 4.8, the 95th percentile to 36.23%, and the negative investment fraction drops to 5.89%.

Next, starting from column "with M&A", column "others" adjusts for small differences between CP's sampling and what we view as more standard practice. CP drop financials, utilities, and unclassified firms, while we drop financials, firms with negative book equity, and firm-quarters with negative or zero assets, net PPE, or sales. CP require nonmissing book-to-market, but we do not. The sample size rises further by 6.07% to 504,692 firm-quarters, largely because we include utilities.

The last step in our replication is to adjust for the treatment of outliers. CP drop firm-quarters in the top and bottom 0.5% of the pooled distribution of gross investment rates. In contrast, we consider winsorization as yielding more reliable investment rate moments. For the pooled firm-quarters of the fiscal quarters ending in a given calendar quarter, we winsorize gross investment rates at the 1%–99% level. Column "1%–99%" details the evidence. The investment rate skewness is 3.53, the 95th percentile 37.64%, and the fraction of negative investment rates 6.17%.

3.2 Scaling Investment Rates with Gross PPE

In our companion paper, we identify an essential tension in investment rate measurement. Within financial accounting, Tang's (2009) investment rate is most conceptually appropriate. Alas, because accounting depreciation rates are much higher than economic depreciation rates, net PPE tends to be much lower than the replacement cost of capital. Consequently, the investment rates scaled by net PPE tend to be much higher than BEA's aggregate estimates. A tension arises because the latter estimates seem more plausible in terms of economic magnitude. Many authors opt to scale instead by gross PPE to bring investment rates more in line with BEA's to alleviate the tension.

Our companion paper explains why. By integrating national accounting in BEA with financial accounting in Compustat, we have constructed firm-specific current-cost capital stocks for the entire Compustat universe. Table 8 in our companion paper shows that gross PPE is much closer to the replacement cost than net PPE. The ratio of the replacement cost divided by gross PPE is on average 0.98, but the average ratio divided by net PPE ratio is 2.11.

Intuitively, gross PPE is a historical-cost measure that ignores both depreciation and capital price inflation. As a proxy for the replacement cost of capital, ignoring depreciation creates an upward bias, while ignoring price inflation creates a downward bias. The two biases largely offset each other in the data. Our companion paper shows that the aggregate economic depreciation rate is 5.71% and the aggregate price inflation rate 4.14%. In contrast, net PPE is a historical-cost measure that uses aggressive accounting depreciation and ignores capital price inflation. Both create downward biases. As such, gross PPE is a better proxy for the replacement cost than net PPE.

In Table 2, the last column denoted "/gross PPE" shows the moments of gross investment scaled by gross PPE. Because of the lower coverage of item PPEGTQ, the sample size drops to 378,122 firm-quarters. The mean investment rate is 5.03% per quarter, and the standard deviation 9.48%. The skewness remains high, 3.59, and the negative investment fraction stays low, 4.94%.[15]

3.3 Replicating CP's Figure 1

To replicate CP's Figure 1 on the gross investment rate distribution, we plot Panels A, C, and D in Figure 5 based on our replication samples in columns "with M&A," "1%–99%," and "/gross PPE" in Table 2, respectively. After fixing all CP's design issues and discrepancies, Panel A shows a heavily asymmetric distribution of gross investment rates with a fat right tail.

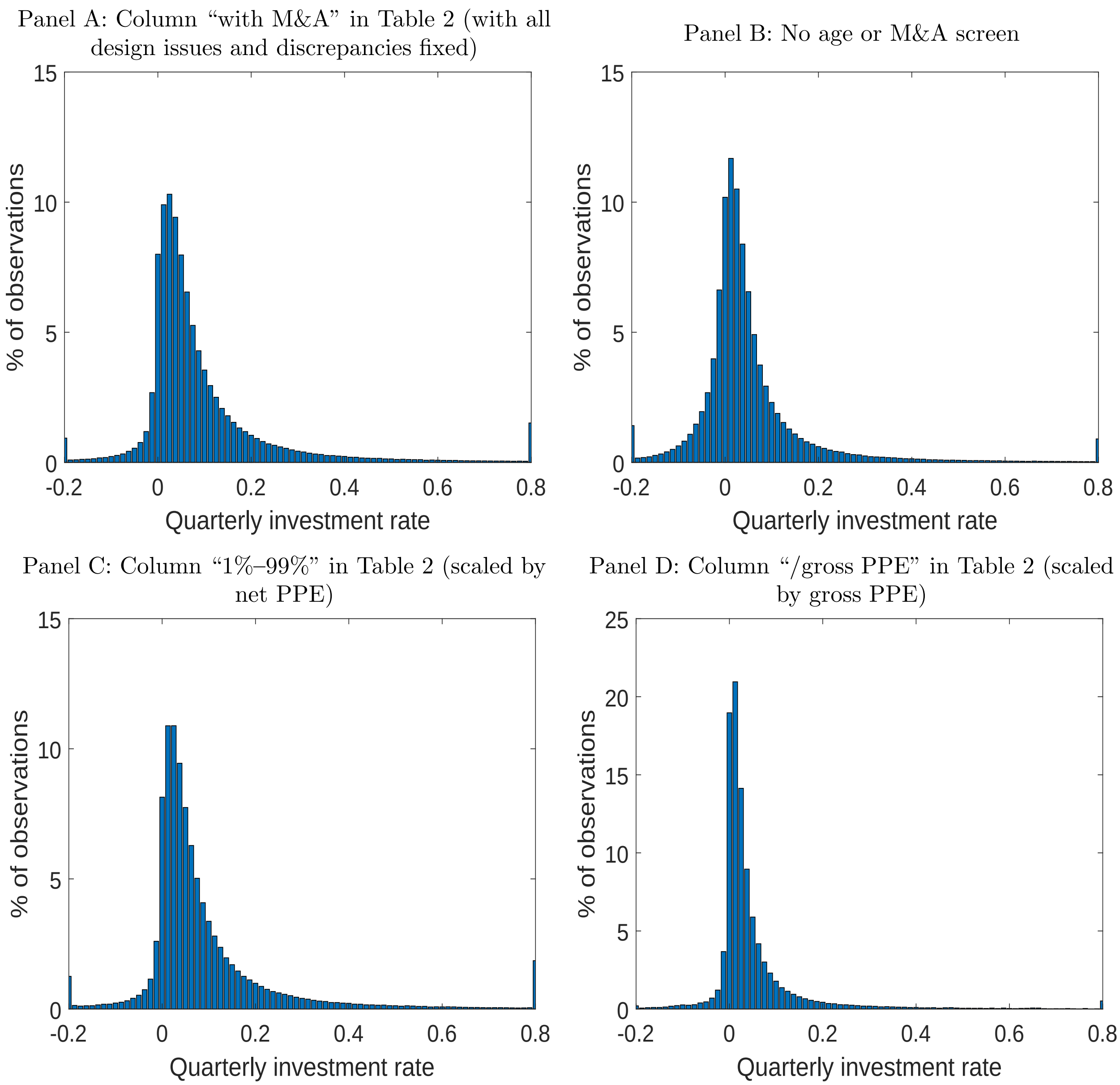
Panel B perturbs on Panel A by only removing the 12-quarter age and 5% M&A screens but still using BEA depreciation rates as in CP.[16] Panel B still shows a long right tail. (This panel in effect combines Panels B and C in Figure 3 that show the impact of each screen separately.) As

[15]In our companion paper, firm-specific current-cost capital stocks are in annual frequency due to the data limitations of the BEA fixed assets account, which provides only annual data. It might be possible to extend our data infrastructure to quarterly data by imputing annual geometric depreciation rates as well as capital and investment price deflators to the quarterly frequency. We leave such an extension to future work. For our replication purpose, we cut the Gordian knot by simply using gross PPE as a proxy for the replacement cost of capital.

[16]Results from this combination are not reported in Table 1, which shows each screen separately, or in Table 2, which shows cumulative impact with the data error on depreciation rates. For completeness, the sample size of this combination is 488,042 firm-quarters. The mean investment rate is 5.8%, standard deviation 14.86%, skewness 4.31, and autocorrelation 21.36%. The fraction of negative investment rates is 16.57%, inactive rates 13.46%, positive spikes 7.35%, and negative spikes 1.21%. Finally, the 5th, 50th, and 95th percentiles are −6.85%, 3.17%, and 26.35%, respectively.

Figure 5 : Replication of CP's Figure 1 on the Gross Investment Rate Distribution, 1978Q1–2016Q4

Panel A is from our replication sample in column "with M&A" in Table 2, in which we cumulatively fix all CP's design issues and discrepancies. Panel B performs a perturbation on Panel A by only removing the 12-quarter age and the 5% M&A screens but still using BEA depreciation rates as in CP. Panel C is from our replication sample in column "1%–99%" and Panel D from column "/gross PPE" in Table 2.



such, CP's data error on depreciation rates alone cannot fully explain the thin (missing) right tail in their Figure 1. Their age and M&A screens also combine to play a role.

Panel C shows the investment rate distribution in our complete replication, with net PPE as the deflator, and Panel D with gross PPE as the deflator. Above all, Panels A, C, and D restore the fat right tail of the gross investment rate distribution to all its glory.

3.4 Robustness

In the Internet Appendix, we perform a battery of robustness checks on our replication (with both net and gross PPE as capital as in the last two columns of Table 2). Without going through the details, we can report robustness from sample-split (in mid 1996), various M&A and age screens, as well as size split on market equity or capital. With net PPE as capital, the investment rate skewness ranges from 3.25 to 4.83, and the fraction of negative investment rates from 3.97% to 6.76%. With gross PPE, the skewness varies from 3.32 to 4.57, and the negative fraction 3.34% to 5.39%. Across all 20 robustness checks, the investment rate distribution is reliably asymmetric with a fat right tail.

We also show robustness across 19 NAICS nonfinancial sectors. With net PPE as capital, the investment rate skewness ranges from 1.81 to 10.83, with 4.61 in the full sample.[17] The fraction of negative investment rates varies from 3.18% to 11.18%, with 6.27% in the full sample. With gross PPE as capital, the skewness goes from 2.23 to 8.86, with 5.08 in the full sample, and the negative fraction from 2.79% to 9.8%, with 5.02% in the full sample. Across all 38 sector-deflator perturbations, the investment rate distribution is robustly asymmetric with a fat right tail.

4 Reanalysis

CP's flawed empirical analysis has permeated their theoretical work, which zooms in on their faulty fraction of negative investment rates, 18.2%, while turning a blind eye to the asymmetry in the investment rate distribution. CP state: "Our calibration strategy sets our study apart from any

[17]Because some sectors can have only a small number of firms in a given quarter, we report panel data moments for sectors (as opposed to the time series averages of cross-sectional moments in the remainder of our study).

other investigation of equity prices in production-based models, as we do not target any feature of the cross-section of returns. Rather, we require the model to be consistent with our evidence on investment and we evaluate its implications for equity returns (p. 292)." Throughout their Tables II, VI and VII, CP force their model to match the 18.2% fraction but not the investment rate skewness.

CP claim: "The investment-based one-factor model does not explain investment (p. 285)," once requiring that "the investment process implied by the asset pricing models under consideration conform closely with this evidence [the 18.2% fraction] (p. 289)." "[A]s long as we require it to be consistent with the cross-sectional evidence on investment, the model simply cannot generate greater dispersion in returns (p. 301)!" "The data strongly suggest that U.S. public firms do adjust to adverse profitability shocks by divesting capital. When capital adjustment costs are parameterized to reflect this feature of the data, states of nature characterized by low aggregate productivity (high marginal utility) see value firms disinvest. This makes them safer, leading to a lower value premium (p. 307)."

Because the defective 18.2% fraction is hardwired into CP's analysis, we view their theoretical work as mostly wrong. As such, we see no need to reproduce or replicate it further.[18] Instead, adopting Clemens's (2017) robustness tests, we perform a reanalysis by asking whether a baseline investment model can explain the value premium and investment dynamics jointly.

In the Internet Appendix, we implement a baseline investment model via simulated method of moments. We estimate four parameters (the upward and downward adjustment cost parameters, the fixed cost of production, and the conditional volatility of firm-specific productivity), while targeting seven data moments (the average value premium, the volatility and skewness of individual stock excess returns, the volatility, skewness, and autocorrelation of investment rates, and the fraction of negative investment rates). The point estimates strongly indicate costly reversibility and operating leverage, which are a good start to explaining the value premium and investment moments.

[18]In addition, CP's model includes many features that are not the most standard specifications, such as labor and wage rate, maintenance investment (with zero adjustment costs), a wedge in the purchasing price of investment higher than maintenance investment, and an exogenous stochastic depreciation rate process, among others. CP do not document how each departure from the prior literature affects their quantitative results. Finally, because CP's article predates the *Journal*'s code sharing policy, we do not have access to their codes for the theoretical analysis.

Our reanalysis embodies what philosopher of science Michael Weisberg (2013, p. 100) calls "minimalist idealization," which is the practice of constructing and studying models that include only the core causal forces. A minimalist model has a special place in science because it can reveal the most important causal powers at the heart of a phenomenon. Adding more details to the model does not improve the explanation but only allows a more thorough characterization of a specific event.

Replicating a theoretical study is more challenging than replicating an empirical study. The target in the latter is a statistic, such as the fraction of negative investment rates, which dwells in what Bhaskar (1975) calls the empirical domain (observable). Alas, the target in the former is a causal mechanism, which lives in the Bhaskarian real domain (unobservable). Also, an empirical pattern can be caused by multiple, unobservable mechanisms, whose relative strength varies over time.

To study a mechanism, a theorist must build a model (a thought experiment) (Mäki 2005). As in a material experiment, one makes (false) assumptions to shield the targeted mechanism from other interfering ones. Following Mäki (2004), we view truth as correspondence with mind-independent reality and nominate a model's targeted mechanism (not its assumptions) as its truth-bearer.

Alas, the economy is not a machine, in which one can study the mechanisms separately and then piece them back together, without ever needing to change parameter values, per Lucas (1980).[19] Rather, the economy is a complex adaptive system, in which one chooses which mechanism to target and which to ignore based on one's perspective (Wimsatt 2007). However, one perspective can be incompatible with others. How should we decide on what is real? What are a model's truth-makers?

Besides evidential truth-making, we adopt the Levins-Wimsatt robustness criterion: "[W]e attempt to treat the same problem with several alternative models, each with different simplifications, but with a common biological assumption. Then, if these models, despite their different assumptions, lead to similar results we have what we can call a robust theorem that is relatively free of the

[19]Browning, Hansen, and Heckman (1999) emphasize the model-dependence of point estimates: "Different microeconomic studies make different assumptions, often implicit, about the economic environments in which agents make their decisions. They condition on different variables and produce parameters with different economic interpretations. A parameter that is valid for a model in one economic environment cannot be uncritically applied to a model embedded in a different economic environment (p. 546)."

details of the model. Hence, our truth is the intersection of independent lies (Levins 1966, p. 423)."

This robustness is broadly aligned with Whewell's consilience in his 1840 "The Philosophy of the Inductive Sciences." Laudan (1971, p. 369) quotes Whewell: "The Consilience of Inductions takes place when an Induction, obtained from one class of facts, coincides with an Induction, obtained from another different class. This Consilience is a test of the truth of the Theory in which it occurs."

We clarify that we only nominate the asymmetry causal mechanism as the truth-bearer, as opposed to the specific assumptions in the original model. Our companion paper provides broad evidential truth-making. For the "robustness theorem," the Internet Appendix lists 28 articles published since 1999 on asset pricing theory, all of which build on the asymmetry mechanism.

For robustness within the neoclassical investment model, since its derivation in Zhang (2005), the asymmetry mechanism has appeared in diverse specifications, often to address different questions. Tuzel (2010) shows that firms with higher real estate holdings earn higher average returns because real estate faces higher disinvestment costs and depreciates more slowly than other capital. Lin and Zhang (2013) embed the asymmetry mechanism to study the covariance versus characteristic tests. Kuehn and Schmid (2014) show that asymmetry helps explain the credit spread puzzle. Bai et al. (2019) embed the mechanism in a disaster model to explain the CAPM failure. Herskovic, Kind, and Kung (2023) shows the asymmetry mechanism at work in a long-run risks model.

For consilience across the real options framework, Carlson, Fisher, and Giammarino (2004) derive the asymmetry mechanism from irreversibility and operating leverage. Carlson, Fisher, and Giammarino (2006, 2010) apply the mechanism to study seasoned equity offerings. Cooper (2006) derives the asymmetry mechanism in a different setup with irreversibility and nonconvex adjustment costs. Gu, Hackbarth, and Johnson (2018) clarify the interaction between costly reversibility and operating leverage when affecting risk and expected returns. As such, per the robustness-consilience criterion, the asymmetry mechanism is squarely in the fabric of our reality.[20]

[20] A two-shock literature has recently developed an alternative mechanism of the value premium, but with a negative price of risk on an investment (or adjustment technology) shock. This negative sign implies, counterfactually,

5 Conclusion

The CP article, which concludes no sign of irreversibility in Compustat, is wrong. Its fatal flaws originate from (i) a data error from the logic inconsistency between their measures of net investment and depreciation rates; (ii) nonstandard sample screens that drop the first 12 quarters for each firm and firm-quarters associated with acquisitions higher than 5% of assets, both curbing the right tail of the investment rate distribution; and (iii) a questionable research practice that cuts off the right tail at 0.2 when plotting the distribution. Fixing CP's design issues and discrepancies, our replications show an investment rate skewness between 3.53 to 4.8 and a fraction of negative investment rates between 4.94% to 6.17%. In accordance with Cooper and Haltiwanger (2006), the firm-level gross investment rate distribution in Compustat is heavily asymmetric with a fat right tail.

high marginal utility in the 1990s, when the second shock is hugely positive. Because this negative price of risk is right in the heart of the alternative mechanism, one must reject it as, to quote Wolfgang Pauli, "not even wrong."

References

Abel, Andrew B., 1983, Optimal investment under uncertainty, *American Economic Review* 73, 228–233.

Abel, Andrew B., Avinash K. Dixit, Janice C. Eberly, and Robert S. Pindyck, 1996, Options, the value of capital, and investment, *Quarterly Journal of Economics* 111, 753–777.

Abel, Andrew B., and Janice c. Eberly, 1994, A unified model of investment under uncertainty, *American Economic Review* 84, 1369–1384.

Abel, Andrew B., and Janice c. Eberly, 1996, Optimal investment with costly reversibility, *Review of Economic Studies* 63, 581–593.

Arrow, Kenneth J., 1968, Optimal capital policy with irreversible investment, in J. N. Wolfe, ed., *Value, Capital, and Growth: Papers in Honour of Sir John Hicks*, Edinburgh: Edinburgh University Press.

Bai, Hang, Erica X. N. Li, Chen Xue, and Lu Zhang, 2022, Asymmetric investment rates, NBER working paper #29957.

Bai, Hang, Kewei Hou, Howard Kung, Erica X. N. Li, and Lu Zhang, 2019, The CAPM strikes back? An equilibrium model with disasters, *Journal of Financial Economics* 131, 269–298.

Bernanke, Ben S., 1983, Irreversibility, uncertainty, and cyclical investment, *Quarterly Journal of Economics* 98, 85–106.

Bhaskar, Roy, 1975, *A Realist Theory of Science*, Routledge.

Browning, Martin, Lars Peter Hansen, and James J. Heckman, 1999, Micro data and general equilibrium models, in John B. Taylor and Michael Woodford, Eds., *Handbook of Macroeconomics* Volumn 1, Elsevier B.V.

Caballero, Ricardo J., 1999, Aggregate investment, in John B. Taylor and Michael Woodford, Eds., *Handbook of Macroeconomics* Volumn 1, Elsevier B.V.

Caballero, Ricardo J., Eduardo M. R. A. Engel, and John C. Haltiwanger, 1995, Plant-level adjustment and aggregate investment dynamics, *Brookings Papers on Economic Activity* 2, 1–39.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2004, Corporate investment and asset price dynamics: Implications for the cross-section of returns, *Journal of Finance* 59, 2577—2603.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2006, Corporate investment and asset price dynamics: Implications for SEO event studies and long-run performance, *Journal of Finance* 61, 1009–1034.

Carlson, Murray, Adlai Fisher, and Ron Giammarino, 2010, SEO risk dynamics, *Review of Financial Studies* 23, 4026–4077.

Clemens, Michael A., 2017, The meaning of failed replications: A review and proposal, *Journal of Economic Surveys* 31, 326–342.

Clementi, Gian Luca, and Berardino Palazzo, 2019, Investment and the cross-section of equity returns, *Journal of Finance* 74, 281–321.

Cooper, Ilan, 2006, Asset pricing implications of nonconvex adjustment costs and irreversibility of investment, *Journal of Finance* 61, 139—170.

Cooper, Russell W., and John C. Haltiwanger, 2006, On the nature of capital adjustment costs, *Review of Economic Studies* 73, 611–633.

Cooper, Russell W., and John C. Haltiwanger, and Laura Power, 1999, Machine replacement and the business cycle: Lumps and bumps, *American Economic Review* 89, 921–946.

Dixit, Avinash K., and Robert S. Pindyck, 1994, *Investment Under Uncertainty* Princeton: Princeton University Press.

Doms, Mark, and Timothy Dunne, 1998, Capital adjustment patterns in manufacturing plants, *Review of Economic Dynamics* 1, 409–429.

Fraumeni, Barbara M., 1997, The measurement of depreciation in the U.S. National Income and Product Accounts, *Survey of Current Business* July, 7–23.

Gu, Lifeng, Dirk Hackbarth, and Tim Johnson, 2018, Inflexibility and stock returns, *Review of Financial Studies* 31, 278–321.

Hamermesh, Daniel S., 2007, Viewpoint: Replication in economics, *Canadian Journal of Economics* 40, 715–733.

Herskovic, Bernard, Thilo Kind, and Howard Kung, 2023, Micro uncertainty and asset prices, *Journal of Financial Economics* 149, 27–51.

Hulten, Charles R., and Frank C. Wykoff, 1981, The estimation of economic depreciation using vintage asset prices, *Journal of Econometrics* 15, 367–396.

John, Leslie K., George Loewenstein, and Drazen Prelec, 2012, Measuring the prevalence of questionable practices with incentives for truth telling, *Psychological Science* 23, 524–532.

Kuehn, Lars-Alexander, and Lukas Schmid, 2014, Investment-based corporate bond pricing, *Journal of Finance* 69, 2741–2776.

Laudan, Larry, 1971, William Whewell on the consilience of inductions, *The Monist* 55, 368–391.

Levins, Richard, 1966, The strategy of model building in population biology, *American Scientist* 54, 421–431.

Lin, Xiaoji, and Lu Zhang, 2013, The investment manifesto, *Journal of Monetary Economics* 60, 351–366.

Lucas, Robert E., Jr., 1980, Methods and problems in business cycle theory, *Journal of Money, Credit, and Banking* 12, 696–715.

Lyandres, Evgeny, Le Sun, and Lu Zhang 2008, The new issues puzzle: Testing the investment-based explanation, *Review of Financial Studies* 21, 2825–2855.

Mäki, Uskali, 2004, Some truths about truth for economists and their critics and clients, in Peter Mooslechner, Helene Schuberth, and Martin Schürz, Eds., *Economic Policy-Making under Uncertainty: The Role of Truth and Accountability in Policy Advice*, Edward Elgar.

Mäki, Uskali, 2005, Models are experiments, experiments are models, *Journal of Economic Methodology* 12, 303–315.

McDonald, Robert L, and Daniel Siegel, 1986, The value of waiting to investment, *Quarterly Journal of Economics* 101, 707–728.

Nilsen, Øividn Anti, and Fabio Schiantarelli, 2003, Zeros and lumps in investment: Empirical evidence on irreversibilities and nonconvexities, *Review of Economics and Statistics* 85, 1021–1037.

Tang, Tony T., 2009, Information asymmetry and firms' credit market access: Evidence from Moody's credit rating format refinement, *Journal of Financial Economics* 93, 325–351.

The Office of Science and Technology Policy, 2000, Federal Policy on Research Misconduct, Executive Office of the President, https://ori.hhs.gov/federal-research-misconduct-policy

Tuzel, Selale, 2010, Corporate real estate holdings and the cross section of stock returns, *Review of Financial Studies* 23, 2268–2302.

Wahlen, James M., Stephen P. Baginski, and Mark T. Bradshaw, 2018, *Financial Reporting, Financial Statement Analysis, and Valuation: A Strategic Perspective*, Cengage Learning, Inc.

Weisberg, Michael, 2013, *Simulation and Similarity: Using Models to Understand the World*, Oxford University Press.

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

Wimsatt, William C., 2007, *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality*, Harvard University Press.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

The Internet Appendix (for Online Publication Only): "Firm-level Irreversibility"

A Robustness

Table S1 shows subsample analysis for the quarterly gross investment rates. Figures S1 and S2 plot the gross investment rate distributions across ten robustness checks with net and gross PPE as capital, respectively. Table S2 shows panel data moments of gross investment rates by NAICS sectors. Figures S3 and S4 plot the gross investment rate distributions across 19 NAICS nonfinancial sectors with net and gross PPE as capital, respectively.

B Reanalysis

The baseline investment model is as follows:

$$V(K_{it}, X_t, Z_{it}) = \max_{\{\chi_{it}\}} \left(\max_{\{I_{it}\}} \Pi_{it} - I_{it} - H_{it} + E_t \left[M_{t+1} V(K_{it+1}, X_{t+1}, Z_{it+1}) \right], 0 \right) \quad \text{(S1)}$$

$$\Pi_{it} = X_t Z_{it} K_{it}^{\alpha} - f \quad \text{(S2)}$$

$$\log(X_{t+1}) \equiv x_{t+1} = \overline{x}(1-\rho_x) + \rho_x x_t + \sigma_x \epsilon_{t+1}^x \quad \text{(S3)}$$

$$\log(Z_{it+1}) \equiv z_{it+1} = \rho_z z_{it} + \sigma_z \epsilon_{it+1}^z \quad \text{(S4)}$$

$$K_{it+1} = I_{it} + (1-\delta) K_{it} \quad \text{(S5)}$$

$$H_{it} = \frac{\theta_{it}}{2} \left(\frac{I_{it}}{K_{it}} - \delta \right)^2 K_{it} \quad \text{with } \theta_{it} \equiv \theta_U \mathbf{1}_{\{I_{it}/K_{it} \geq \delta\}} + \theta_D \mathbf{1}_{\{I_{it}/K_{it} < \delta\}} \quad \text{(S6)}$$

$$M_{t+1} = \beta \exp([\gamma_0 + \gamma_1 (x_t - \overline{x})](x_t - x_{t+1})) \quad \text{(S7)}$$

Firm i maximizes the market value, $V(K_{it}, X_t, Z_{it})$, in equation (S1). Π_{it} is operating profits, $\alpha \in (0,1)$ is the curvature parameter, $\delta \in (0,1)$ is the depreciation rate, $f > 0$ the fixed cost of production, K_{it} capital, X_t the aggregate productivity, Z_{it} firm-specific productivity, with ϵ_{t+1}^x and ϵ_{it+1}^z i.i.d. standard normal shocks uncorrelated for any i, and ϵ_{it+1}^z and ϵ_{jt+1}^z uncorrelated for any $i \neq j$. H_{it} is the asymmetric adjustment cost function, in which $\mathbf{1}_{\{\cdot\}}$ is the indicator that equals one if the event in $\{\cdot\}$ is true and zero otherwise, and $\theta_D > \theta_U > 0$ are constant parameters. M_{t+1} is the stochastic discount factor, in which $\beta \in (0,1)$, $\gamma_0 > 0$, and $\gamma_1 < 0$ are constant parameters.

When the inner maximand in equation (S1) is greater than or equal to zero, firm i stays in the economy. Evaluating the value function at the optimum yields $V_{it} = D_{it} + E_t[M_{t+1}V_{it+1}]$, in which $D_{it} \equiv \Pi_{it} - I_{it} - H_{it}$, and $E_t[M_{t+1}r^S_{it+1}] = 1$, in which $r^S_{it+1} \equiv V_{it+1}/(V_{it} - D_{it})$ is the stock return. When the inner maximand is negative, firm i exits at the beginning of t. We set its stock return over period $t-1$, r^S_{it}, to be a predetermined delisting return, denoted $\widetilde{R}$. The exit firm enters an immediate reorganization process. The current shareholders receive nothing and leave. New shareholders take over the firm's capital to form a new firm. For tractability, we assume that the reorganization process occurs instantaneously. At the beginning of period t, the exit firm is replaced by a new firm with a new firm-specific log productivity of $\bar{z}$, which is its unconditional mean.

B.1 Predetermined Parameters

There are in total 14 parameters, $\{\beta,\ \gamma_0,\ \gamma_1,\ \alpha,\ \overline{x},\ \rho_x,\ \sigma_x,\ \delta,\ \widetilde{R},\ \rho_z,\ \sigma_z,\ f,\ c^+,\ c^-\}$. Following Bloom (2009), we calibrate a set of predetermined parameters but estimate the key parameters via simulated method of moments (SMM).

We work in monthly frequency. We set $\beta = 0.9999$, $\gamma_0 = 18$, and $\gamma_1 = -450$, which yield an average Sharpe ratio of 0.34 per annum, an average interest rate of 0.67%, and an interest rate volatility of 3.31%. We update the King-Rebelo (1999) estimates of persistence and conditional volatility of productivity shocks. We retrieve from FRED the time series of private nonfarm business multifactor productivity constructed by BLS in the annual 1963–2021 sample.[1] We fit a linear time trend to the log productivity and use the residuals to estimate the persistence, ρ_x, and the conditional volatility, σ_x. In the annual frequency, ρ_x is 0.72117 and σ_x 0.01643. The implied quarterly estimates are 0.92153 and 0.009212, and monthly 0.97313 and 0.005463, respectively.[2]

The curvature parameter, α, is 0.7. In the 1963–2021 sample, the average (performance-related) delisting return in CRSP is -24.53%, to which we set $\widetilde{R}$. Bai et al. (2022) estimate the median firm-level economic depreciation rate as 6.86% per annum. We set δ to be 6.86%/12. The persistence, ρ_z,

[1]https://fred.stlouisfed.org/series/MFPNFBS

[2]Heer and Maussner (2009, p. 549–550) offer the formulas that convert the estimates between frequencies.

and conditional volatility, σ_z, of firm-specific productivity, z_{it}, affect the cross-sectional dispersion similarly. We fix $\rho_z = 0.97$ per Imrohoroglu and Tuzel (2014) but estimate σ_z. We set $\overline{x} = -3.98$ to yield a large enough long-term average capital to ensure a stable investment rate distribution.

B.2 SMM Estimation and Tests

We estimate $\mathbf{c} \equiv (\sigma_z, f, \theta_U, \theta_D)$, in which σ_z is the conditional volatility of firm-specific productivity, f the fixed cost of production, θ_U the upward and θ_D downward adjustment cost parameters. We target seven data moments, $\mathbf{\Psi_d}$, four annual investment rate moments (the volatility 37.63%, skewness 3.35, autocorrelation 33.58%, and the fraction of negative investment rates 5.65%) from Bai et al. (2022) (extended through 2021) and three return moments (the average value premium 0.38% per month, and the volatility 55.68% and skewness 1.43 of annual stock excess returns) in the 1963–2021 sample.[3] We solve the model via value function iteration. We simulate $S = 500$ panels of size $(N, T + b)$, in which $N = 3,500$ is the number of firms, $T = 59 \times 12 = 708$ the number of months, and $b = 300$ the burn-in. Following Bloom et al. (2018), we set the diagonal elements of the weighting matrix, $\mathbf{W}$, to $1/(\mathbf{\Psi_d})^2$ and the off-diagonal elements to zero. The SMM estimator thus minimizes the sum of squared percentage deviations of the model moments from the data moments.

Table S3 shows the estimation and tests. Costly reversibility is highly significant. The downward adjustment cost parameter, θ_D, is 102.26 ($t = 30.86$), but the upward adjustment cost parameter, θ_U, is only 0.14 ($t = 1.42$). Operating leverage is also significant. The fixed cost of production, f, is 0.0496 ($t = 3.8$). The conditional volatility of firm-specific productivity is 0.215 ($t = 8.45$).

The baseline model does a good job in matching the moments in terms of economic magnitude. The average value premium is 0.4% per month, which is close to 0.38% in the data. The fraction of negative investment rates is 5.84% in the model, which is close to 5.65% in the data. The stock return volatility and skewness are also close. However, the model understates the investment rate volatility, 23.8% versus 37.6% per annum, but overstates its skewness, 4.78 versus 3.35. Finally, the SMM test is powerful enough to reject the model with the overidentification test.

In untabulated results, the fraction of inactive investment rates is 19.4% in the model. Although

[3]The 0.38% per month estimate ($t = 2.12$) is from Kenneth French's data library in the 1963–2021 sample. The value premium estimate is 0.34% ($t = 1.81$) in the q-data library due to sampling differences. We use French's estimate because it represents a somewhat higher hurdle for the baseline model to match.

substantially higher than 2.88% in the data (Bai et al. 2022), it represents an improvement relative to Cooper and Haltiwanger (2006). Their Table 3 reports this fraction to be between 58.8% to 69% across their model specifications. The better fit originates from our adjustment cost function, which centers the asymmetry around the "normal" investment rate of δ (as opposed to zero).

The remaining data-model gap is likely due to capital heterogeneity. Capital is homogeneous in the model but heterogeneous in the data. Firms face different adjustment costs when constructing a new building versus buying new laptops. Smoothing over heterogeneous capital goods likely yields a lower fraction of inactive investment rates in the data, but this smoothing is absent in the model. For this reason, Cooper and Haltiwanger (2006) do not target the fraction of inactive investment rates in their estimation. We follow the same practice. In addition, we assume the same normal investment rate for all firms, but δ likely varies across firms and over time in the data.

More important, we conduct two comparative statics: (i) symmetric adjustment costs, by lowering θ_D to 0.1385; and (ii) a low fixed cost of production, $f = 0.025$. In each experiment, we only change the parameter in question, while keeping all the other parameters unchanged.

The first experiment has a dramatic impact on the average value premium and the investment rate moments. The fraction of negative investment rates, rises from 5.65% to 41.17%, the investment rate skewness drops from 4.78 to 2.29, and the volatility jumps from 23.8% to 663%. Most important, the value premium falls from 0.4% per month to -0.51%.

The experiment of lowering f from 0.0496 to 0.025 shows the impact of operating leverage in the model. The average value premium falls from 0.4% to 0.05% per month, and the stock return volatility from 45.5% to 29.8% per annum. However, the investment rate moments are less sensitive.

C A Meta-study on the Asymmetry Mechanism

Table S4 lists 28 theoretical articles in asset pricing that have been published since 1999. All feature costly reversibility in their models, in which the asymmetry causal mechanism plays a role.

D CP's Codes, Annotated

At the end of this document, we append CP's codes, annotated by us, in five files. The pages are numbered separately (starting from page 1) within each file for ease of references.

D.1 investment_rate_bea.do

D.2 accounting_data_cleaned.do

D.3 crsp_cleaning.do

D.4 merge_crsp_compustat.do

D.5 inv_rate_moments.do

References

Bai, Hang, Kewei Hou, Howard Kung, Erica X. N. Li, and Lu Zhang, 2019, The CAPM strikes back? An equilibrium model with disasters, *Journal of Financial Economics* 131, 269–298.

Bai, Hang, Erica X. N. Li, Chen Xue, and Lu Zhang, 2022, Asymmetric investment rates, NBER working paper #29957.

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

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen J. Terry, 2018, Really uncertain business cycles, *Econometrica* 86, 1031–1065.

Cooper, Russell W., and John C. Haltiwanger, 2006, On the nature of capital adjustment costs, *Review of Economic Studies* 73, 611–633.

Heer, Burkhard, and Alfred Maussner, 2009, *Dynamic General Equilibrium Modeling: Computational Methods and Applications*, 2nd Edition, Springer.

Imrohoroglu, Ayse, and Selale Tuzel, 2014, Firm-level productivity, risk, and return, *Management Science* 60 (8), 2073–2090.

King, Robert G., and Sergio T. Rebelo, 1999, Resuscitating real business cycles, in John B. Taylor and Michael Woodford, Eds., *Handbook of Macroeconomics* Volume 1B, Elsevier B. V.

Zhang, Lu, 2005, The value premium, *Journal of Finance* 60, 67–103.

Table S1 : Robustness, Quarterly Gross Investment Rates in Compustat, 1978Q1–2016Q4

In Panel A, the gross investment rate is $(\text{PPENTQ}_{it} - \text{PPENTQ}_{it-1})/\text{PPENTQ}_{it-1} + \delta_{it}$, in which PPENTQ is capital (net PPE, item PPENTQ), and δ_{it} is the accounting depreciation rate (the amount of depreciation and amortization [item DPQ] minus the amortization of intangibles [item AM divided by four, zero if missing] scaled by net PPE). In Panel B, the gross investment rate is $(\text{PPENTQ}_{it} - \text{PPENTQ}_{it-1} + \text{DPQ}_{it} - \text{AM}_{it}/4)/\text{PPEGTQ}_{it-1}$, in which PPEGTQ is gross PPE (item PPEGTQ). All moments are in percent, except for the number of firm-quarters (#obs.) and skewness (skew). ρ_1 is the cross-sectional correlation between the investment rates and one-quarter-lagged values. f_- is the fraction of negative investment rates (below −1%), and f_0 the fraction of inactive investment rates (between −1% and 1%). $f^-_{0.2}$ the fraction of negative investment spikes (below −20%), and $f^+_{0.2}$ the fraction of positive investment spikes (above 20%). $(I-\text{CAPX})/K \leq 15\%$ excludes firm-quarters in which the difference between gross investment and capital expenditure (CAPX) is higher than 15% of capital. CAPX is item CAPXQ minus item SPPEQ calculated from year-to-date figures (we set missing SPPEQ to zero). $(I-\text{CAPX})/K \leq 5\%$ excludes firm-quarters in which $I-\text{CAPX}$ is higher than 5% of capital. Age> 3 excludes the first 12 quarters for a given firm, and Age> 5 the first 60 quarters. For size classification, we use the NYSE breakpoints and split all fiscal quarters ending in a calendar quarter into two groups based on the beginning-of-quarter market equity or capital. In all experiments, we continue to winsorize each quarter at the 1%–99% level in the full sample to ensure comparability across subsamples. Except for #obs., all other moments are time series averages of the cross-sectional moments.

	#obs.	mean	std	skew	ρ_1	f_-	f_0	$f^+_{0.2}$	$f^-_{0.2}$	5th	50th	95th
Panel A: Gross investment rates scaled by net PPE												
Full sample	509,788	9.89	18.14	3.53	31.73	6.17	8.57	11.92	1.02	−2.31	5.29	37.64
1978Q1–1996Q2	227,561	9.50	17.97	3.64	27.04	6.22	8.60	11.16	1.08	−2.47	5.02	36.13
1996Q3–2016Q4	282,227	10.25	18.29	3.44	35.90	6.12	8.55	12.60	0.96	−2.16	5.53	39.00
$(I-\text{CAPX})/K \leq 15\%$	482,963	7.67	13.24	3.71	35.10	6.40	8.96	8.28	1.05	−2.53	4.95	26.25
$(I-\text{CAPX})/K \leq 5\%$	448,354	6.96	12.84	3.95	34.92	6.76	9.52	6.90	1.12	−2.91	4.54	24.21
Age>3 years	425,637	8.39	15.78	3.97	28.66	6.25	8.95	9.18	0.98	−2.33	4.85	29.92
Age>5 years	370,480	7.98	15.09	4.14	27.84	6.10	8.88	8.31	0.92	−2.16	4.71	27.91
Small ME	405,379	10.38	19.11	3.32	31.13	6.76	9.56	13.27	1.14	−2.94	5.37	40.95
Big ME	103,625	7.79	12.59	4.68	31.70	3.97	5.04	6.47	0.49	−0.27	5.03	24.30
Small net PPE	415,226	10.99	19.44	3.25	32.04	6.40	8.83	14.02	1.10	−2.57	5.93	42.33
Big net PPE	94,562	5.03	8.82	4.83	13.93	5.22	7.75	2.75	0.60	−1.42	3.73	14.56
Panel B: Gross investment rates scaled by gross PPE												
Full sample	378,122	5.03	9.48	3.59	34.52	4.94	18.08	4.82	0.12	−1.10	2.61	19.50
1978Q1–1996Q2	177,858	5.32	10.25	3.72	30.84	5.15	16.42	5.13	0.19	−1.29	2.73	20.64
1996Q3–2016Q4	200,264	4.77	8.78	3.48	37.75	4.74	19.58	4.54	0.05	−0.93	2.50	18.47
$(I-\text{CAPX})/K \leq 15\%$	359,691	4.02	7.20	3.73	42.38	5.10	18.81	2.84	0.12	−1.21	2.46	14.30
$(I-\text{CAPX})/K \leq 5\%$	335,805	3.69	6.95	3.93	42.20	5.35	19.87	2.57	0.12	−1.37	2.30	13.17
Age>3 years	318,797	4.02	7.67	4.17	27.63	4.90	19.26	2.95	0.10	−1.07	2.37	14.07
Age>5 years	279,275	3.73	7.13	4.41	25.34	4.76	19.66	2.50	0.09	−0.98	2.29	12.70
Small ME	297,353	5.21	9.95	3.40	34.51	5.39	19.88	5.37	0.13	−1.36	2.56	21.12
Big ME	80,204	4.20	6.81	4.37	31.28	3.34	12.35	2.60	0.05	−0.18	2.66	13.60
Small gross PPE	305,230	5.53	10.16	3.32	35.62	5.08	18.61	5.65	0.13	−1.19	2.80	22.05
Big gross PPE	72,892	2.86	5.00	4.57	17.79	4.34	16.93	1.24	0.06	−0.85	2.12	8.35

Table S2 : Panel Data Moments of Quarterly Gross Investment Rates by NAICS Sectors, 1978Q1–2016Q4

In Panel A, the gross investment rate is $(\text{PPENTQ}_{it} - \text{PPENTQ}_{it-1})/\text{PPENTQ}_{it-1} + \delta_{it}$, in which PPENTQ is capital (net PPE, item PPENTQ), and δ_{it} is the accounting depreciation rate, calculated as the amount of depreciation and amortization (item DPQ) minus the amortization of intangibles (item AM divided by four, zero if missing) scaled by net PPE. In Panel B, the gross investment rate is $(\text{PPENTQ}_{it} - \text{PPENTQ}_{it-1} + \text{DPQ}_{it} - \text{AM}_{it}/4)/\text{PPEGTQ}_{it-1}$, in which PPEGTQ is gross PPE (item PPEGTQ) as capital. All moments are in percent, except for the number of firm-quarters (#obs.) and skewness (skew). ρ_1 is the cross-sectional correlation between the investment rates and one-quarter-lagged values. f_- is the fraction of negative investment rates (below -1%), and f_0 the fraction of inactive investment rates (between -1% and 1%). $f^-_{0.2}$ the fraction of negative investment spikes (below -20%), and $f^+_{0.2}$ the fraction of positive investment spikes (above 20%). In all experiments, we continue to winsorize each quarter at the 1%–99% level in the full sample to ensure comparability across sectors. Because some NAICS sectors have a small number of firms in a given quarter, we report panel data moments, as opposed to the time series averages of cross-sectional moments in Table S1.

	#obs.	mean	std	skew	ρ_1	f_-	f_0	$f^+_{0.2}$	$f^-_{0.2}$	5th	50th	95th
	Panel A: Gross investment rates, scaled by net PPE											
Full sample	509,788	10.41	20.38	4.61	34.95	6.27	8.56	12.89	1.09	−2.23	5.31	40.02
Agriculture, forestry, fishing, and hunting	1,666	5.96	17.65	5.75	6.69	9.48	13.21	5.34	1.62	−4.98	3.21	20.91
Mining	27,368	8.05	19.64	4.41	12.36	11.18	12.70	9.83	2.16	−8.79	4.34	33.92
Utilities	21,853	3.44	9.22	10.83	18.10	3.18	9.38	1.59	0.39	0.15	2.44	7.92
Construction	7,125	11.23	21.94	3.91	30.26	8.59	8.69	15.05	1.60	−5.87	6.22	41.37
Nondurable goods	88,110	9.21	19.44	4.83	32.28	5.80	9.67	10.21	0.90	−1.61	4.51	36.83
Durable goods	159,967	9.93	18.31	4.77	27.75	5.86	8.05	12.05	0.93	−1.74	5.56	35.90
Wholesale trade	22,088	11.24	21.81	4.12	25.32	6.58	7.81	14.03	1.23	−2.62	5.69	44.74
Retail trade	31,250	8.84	15.69	5.83	32.17	4.77	6.56	8.60	0.66	−0.80	5.71	27.90
Transportation and warehousing	14,333	7.22	15.76	5.85	18.59	6.78	12.28	7.23	0.82	−2.16	4.00	25.38
Information	48,289	16.98	27.92	3.69	44.99	6.12	4.67	24.88	1.35	−2.57	9.18	65.25
Real estate and rental and leasing	4,583	12.82	27.46	4.17	35.56	10.87	11.63	17.22	1.83	−6.25	6.04	53.27
Professional, scientific, and technical services	26,658	14.75	23.89	3.81	33.49	6.15	5.96	21.15	1.26	−2.36	8.64	55.00
Management of companies and enterprises	106	10.45	16.96	1.81	22.12	6.60	10.38	17.92	0.00	−6.90	4.07	56.00
Administrative and waste management services	13,345	13.63	23.63	3.96	29.66	6.27	5.46	17.68	1.36	−2.55	7.93	50.05
Educational services	2,535	12.15	19.91	4.10	19.16	5.96	4.85	15.07	1.03	−2.07	7.43	43.23
Health care and social assistance	11,887	12.66	22.33	3.70	22.59	6.78	7.14	17.18	1.33	−3.46	6.92	48.96
Arts, entertainment, and recreation	4,875	9.40	22.46	4.79	15.42	7.63	16.64	11.79	1.33	−3.60	3.64	40.88
Accommodation and food services	13,073	6.77	15.62	5.35	22.42	8.54	13.17	6.96	1.25	−3.95	3.85	25.14
Other services, except government	2,954	9.66	20.96	4.37	15.90	8.06	9.72	10.97	1.32	−3.19	4.85	37.59

	#obs.	mean	std	skew	ρ_1	f_-	f_0	$f_{0.2}^+$	$f_{0.2}^-$	5th	50th	95th
Panel B: Gross investment rates, scaled by gross PPE												
Full sample	378,122	5.48	11.42	5.08	38.93	5.02	17.78	5.54	0.13	−1.01	2.62	21.61
Agriculture, forestry, fishing, and hunting	1,185	3.74	10.32	4.94	15.10	8.69	21.01	3.54	0.42	−3.03	1.92	14.40
Mining	24,128	5.09	11.81	4.09	19.15	9.80	20.70	6.06	0.41	−4.88	2.48	23.46
Utilities	20,904	2.27	4.94	8.86	10.46	2.79	19.75	1.01	0.06	0.15	1.68	5.37
Construction	4,629	5.67	11.64	3.93	23.02	7.84	15.62	6.26	0.19	−3.27	2.94	22.83
Nondurable goods	65,213	4.59	9.90	5.26	33.72	4.36	20.51	4.32	0.10	−0.67	2.24	17.82
Durable goods	119,041	4.93	9.91	5.42	32.77	4.38	18.95	4.39	0.11	−0.68	2.55	18.21
Wholesale trade	15,581	6.20	12.63	4.28	30.09	5.62	15.46	6.82	0.13	−1.42	2.94	25.89
Retail trade	21,442	5.44	10.12	5.85	41.50	3.82	13.47	4.33	0.07	−0.36	3.18	18.04
Transportation and warehousing	12,343	4.56	9.49	5.67	27.01	5.66	19.05	3.65	0.09	−1.37	2.54	16.13
Information	32,373	9.38	16.90	3.97	50.66	4.76	11.20	11.93	0.14	−0.85	4.23	38.21
Real estate and rental and leasing	3,592	6.57	13.45	4.46	19.10	9.69	18.40	7.57	0.25	−4.18	3.30	26.52
Professional, scientific, and technical services	18,304	7.53	13.82	4.48	36.68	5.04	12.86	8.53	0.18	−1.05	3.87	29.42
Management of companies and enterprises	79	5.52	10.68	2.23	16.46	7.59	18.99	10.13	0.00	−5.84	2.37	36.57
Administrative and waste management services	9,814	7.32	14.05	4.50	30.88	5.33	11.50	7.91	0.16	−1.23	3.69	28.47
Educational services	1,752	7.15	12.21	3.82	22.51	5.14	11.13	7.82	0.23	−1.10	4.11	25.29
Health care and social assistance	8,704	7.51	13.67	3.67	29.57	6.32	12.07	9.09	0.22	−2.08	3.71	31.24
Arts, entertainment, and recreation	3,882	6.18	13.90	4.30	21.33	6.31	24.45	8.04	0.28	−1.99	2.36	29.60
Accommodation and food services	9,279	4.71	10.37	4.92	31.40	7.04	20.83	4.47	0.23	−2.43	2.44	18.72
Other services, except government	2,074	5.64	12.64	3.99	20.14	7.52	17.65	6.12	0.34	−2.77	2.63	23.46

Table S3 : SMM Estimation and Tests, 1963–2021

We estimate the upward (θ_U) and downward (θ_D) adjustment cost parameters, the fixed cost of production (f), and the conditional volatility of firm-specific productivity shocks (σ_z) to target seven data moments. The moments include four moments of annual gross investment rates (the time series averages of cross-sectional standard deviation, std, skewness, skew, autocorrelation, ρ_1, and the fraction of negative investment rates, f_-), and three return moments (the time series averages of cross-sectional standard deviation, std, and skewness, skew, of annual stock excess returns, as well as the average value premium, $\overline{R}_{\text{H}-\text{L}}$, in monthly percent). Panel A reports the parameter estimates and their t-values. Panel B shows the test of overidentification, including the test statistic, χ^2; the degree of freedom, d.f.; and the p-value of the χ^2 statistic. Panel C shows the data and model moments as well as the t-values of their individual differences.

	Panel A: Parameter estimates				Panel B: The χ^2 test		
	θ_U	θ_D	f	σ_z	χ^2	d.f.	p-value
Estimate	0.1385	102.2631	0.0496	0.2152	103.10	3	0.00
t-value	1.4201	30.8570	3.8001	8.4534			

Panel C: Individual moments and the significance of their model errors

	Gross investment rates				Returns		
	std	skew	ρ_1	f_-	std	skew	$\overline{R}_{\text{H}-\text{L}}$
Data	0.3763	3.3518	0.3358	0.0565	0.5568	1.4325	0.3789
Model	0.2376	4.7773	0.3024	0.0584	0.4545	1.2970	0.4038
t(Diff)	4.00	−8.39	1.89	−6.90	2.12	1.73	−3.21

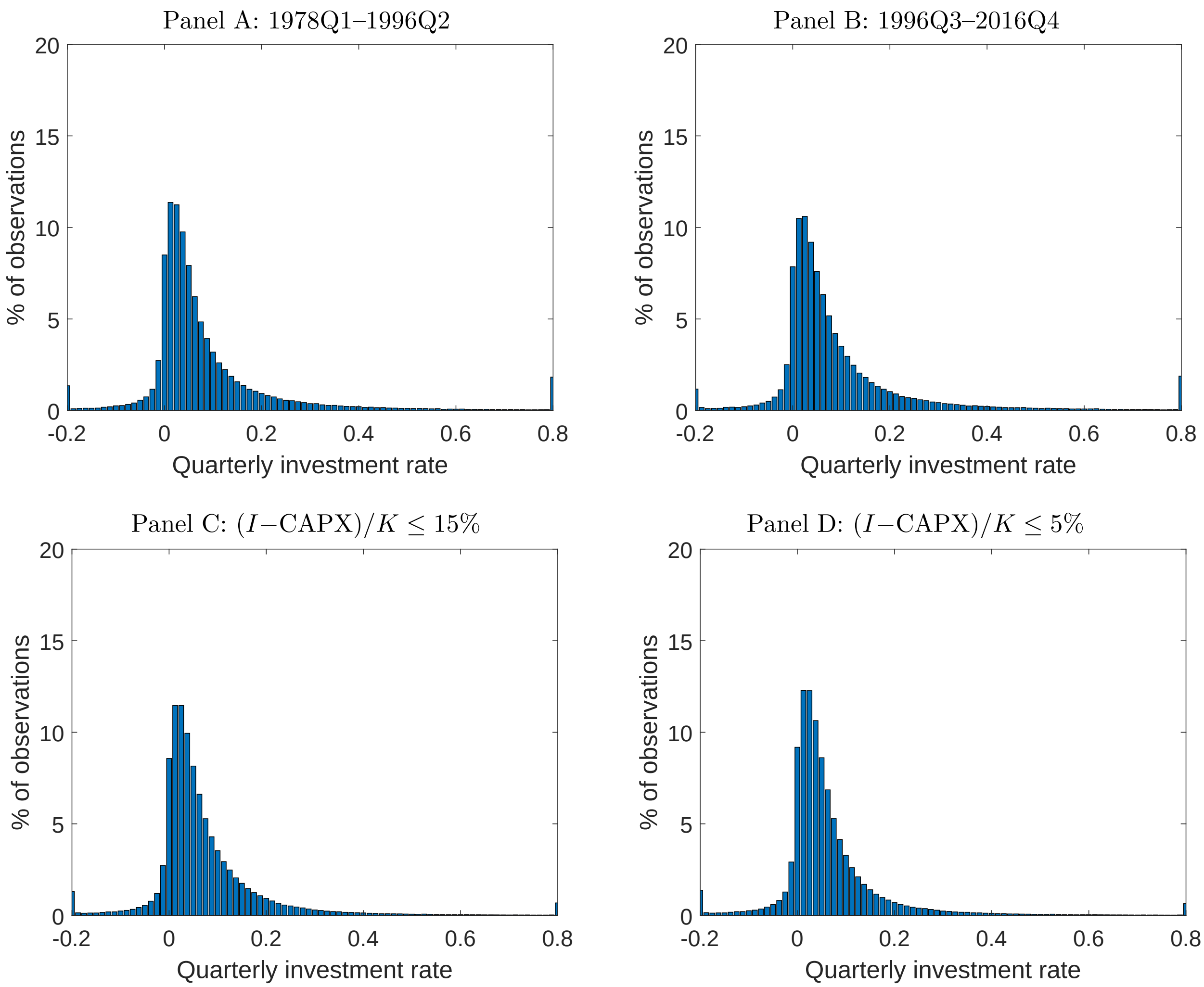
Table S4 : A Meta-study of the Asymmetry Mechanism

This table lists 28 theoretical publications in asset pricing since 1999 with costly reversibility in their models.

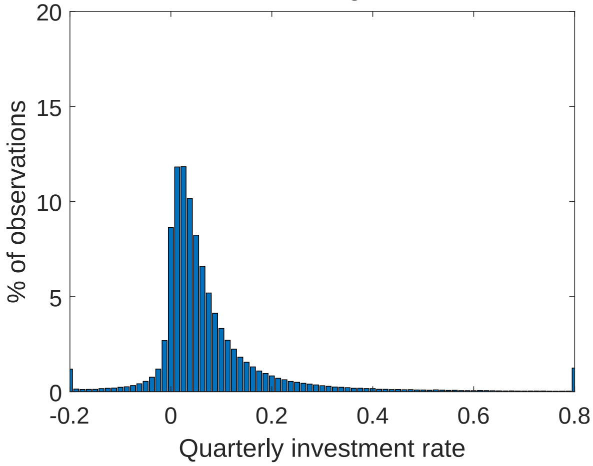
Authors	Year	Journal	Form
Berk, Green, and Naik	1999	JF	Irreversible investment
Hall	2001	AER	Asymmetric adjustment costs
Kogan	2001	JFE	Irreversible investment
Gomes, Kogan, and Zhang	2003	JPE	Irreversible investment
Kogan	2004	JFE	Irreversible investment
Carlson, Fisher, and Giammarino	2004	JF	Resale lower than purchase price
Zhang	2005	JF	Asymmetric adjustment costs
Cooper	2006	JF	Irreversible investment
Carlson, Fisher, and Giammarino	2006	JF	Irreversible investment
Kogan, Livdan, and Yaron	2009	JF	Irreversible investment
Livdan, Sapriza, and Zhang	2009	JF	Asymmetric adjustment costs
Carlson, Fisher, and Giammarino	2010	RFS	Irreversible investment
Gomes and Schmid	2010	JF	Irreversible investment
Tuzel	2010	RFS	Asymmetric adjustment costs
Belo and Lin	2012	RFS	Asymmetric adjustment costs
Ozdagli	2012	RFS	Resale lower than purchase price
Lin and Zhang	2013	JME	Asymmetric adjustment costs
Obreja	2013	RFS	Asymmetric adjustment costs
Belo,Lin, and Bazdresch	2014	JPE	Asymmetric adjustment costs
Kuehn and Schmid	2014	JF	Irreversible investment
Hackbarth and Johnson	2015	ReStud	Resale lower than purchase price
Belo, Li, Lin, and Zhao	2017	RFS	Asymmetric adjustment costs
Li	2017	MS	Asymmetric adjustment costs
Gu, Hackbarth, and Johnson	2018	RFS	Resale lower than purchase price
Bai, Hou, Kung, Li, and Zhang	2019	JFE	Asymmetric adjustment costs
Belo, Lin, and Yang	2019	RFS	Asymmetric adjustment costs
Gomes and Schmid	2021	JF	Partial irreversibility
Herskovic, Kind, and Kung	2023	JFE	Asymmetric adjustment costs

Figure S1 : Robustness, the Quarterly Gross Investment Rate Distribution in Compustat, Scaled by Net PPE, 1978Q1–2016Q4

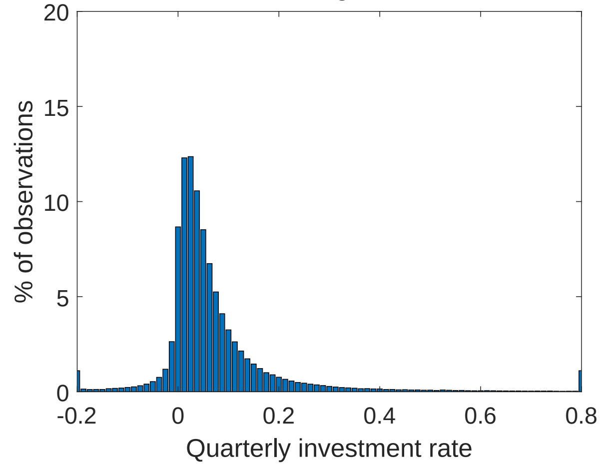
This figure shows ten robustness tests: (i) the first half sample, 1978Q1–1996Q2; (ii) the second half sample, 1996Q3–2016Q4; (iii) $(I - \text{CAPX})/K \leq 15\%$, which excludes firm-quarters with the difference between gross investment and capital expenditure (CAPX) higher than 15% of capital (CAPX is item CAPXQ minus item SPPEQ calculated from year-to-date figures [missing SPPEQ set to zero]); (iv) $(I - \text{CAPX})/K \leq 5\%$, which excludes firm-quarters with the difference between gross investment and CAPX higher than 5% of capital; (v) Age$>$ 3, which excludes the first 12 quarters for a given firm; (vi) Age$>$ 5, which excludes the first 60 quarters for a given firm; (vii) small ME, the small market equity sample; (viii) big ME, the big market equity sample; (ix) Small K, the small capital sample; and (x) Big K, the big capital sample. For last four tests, we use the NYSE breakpoints and split all fiscal quarters ending in a calendar quarter into two groups based on the begin-of-quarter market equity or capital. In all experiments, we continue to winsorize each quarter at the 1%–99% level in the full sample to ensure comparability across subsamples.



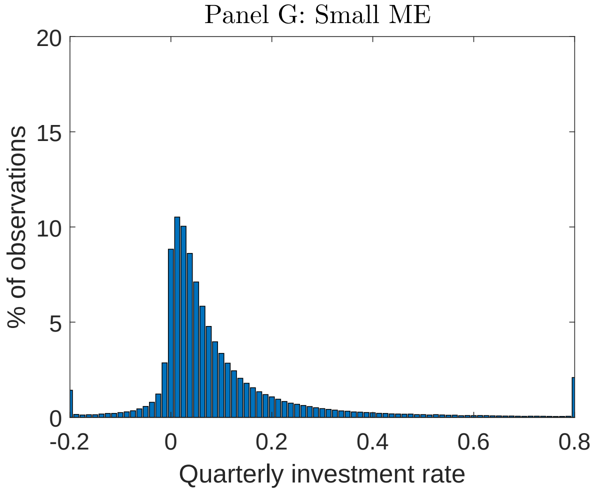
Panel E: Age>3



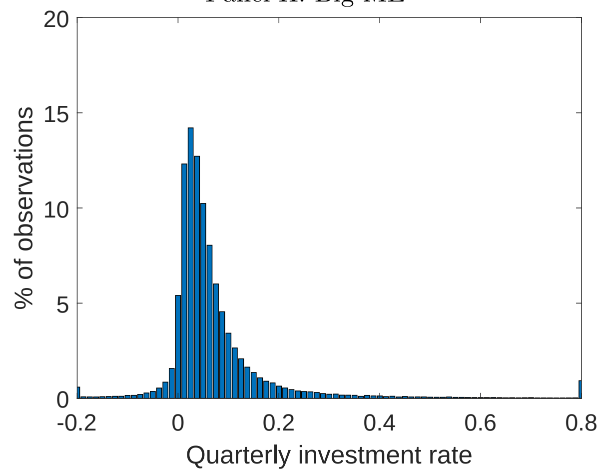
Panel F: Age>5



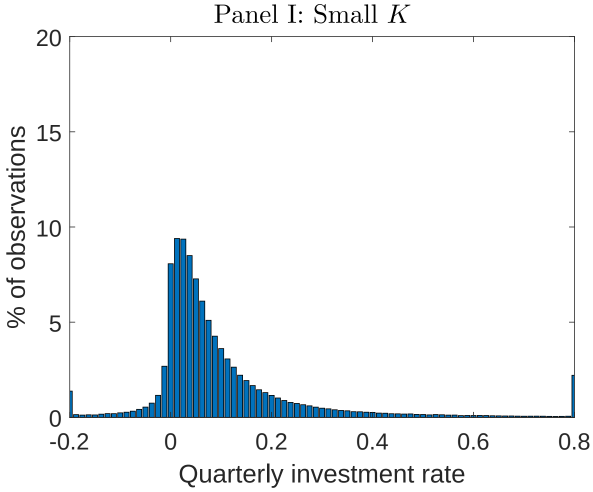
Panel G: Small ME



Panel H: Big ME



Panel I: Small K



Panel J: Big K

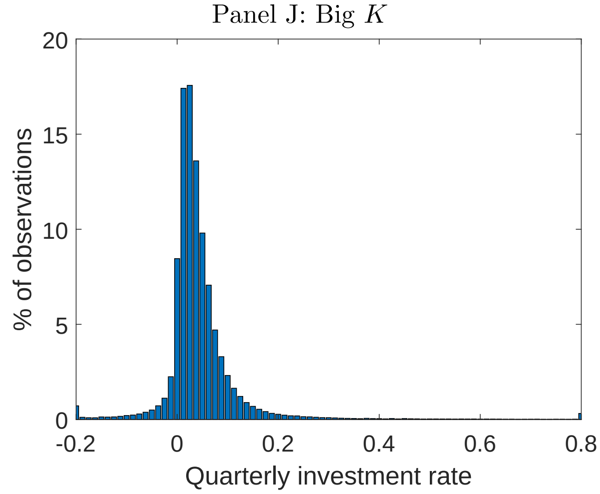
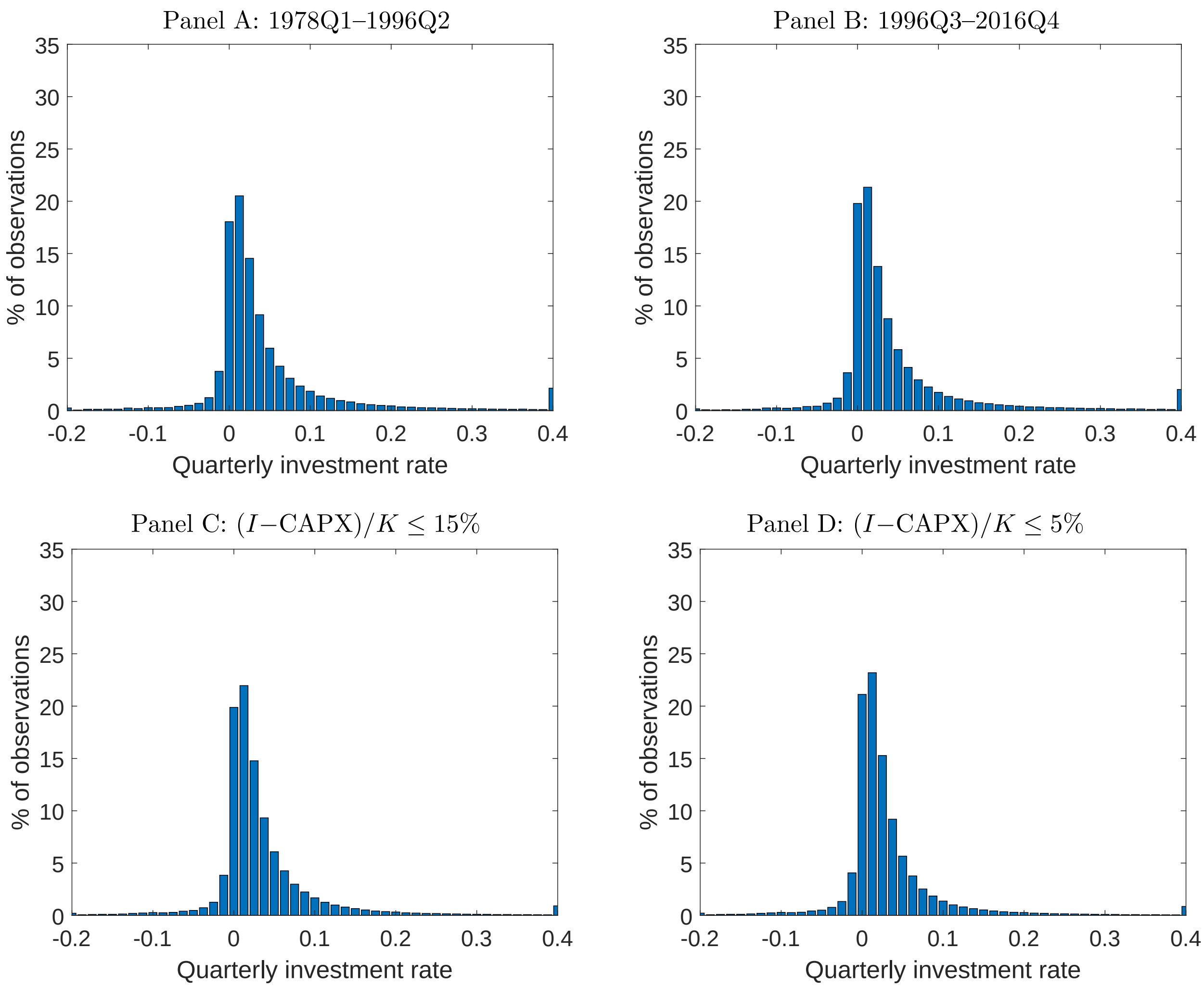


Figure S2 : Robustness, the Quarterly Gross Investment Rate Distribution in Compustat, Scaled by Gross PPE, 1978Q1–2016Q4

This figure shows ten robustness tests: (i) the first half sample, 1978Q1–1996Q2; (ii) the second half sample, 1996Q3–2016Q4; (iii) $(I-\text{CAPX})/K \leq 15\%$, which excludes firm-quarters with the difference between gross investment and capital expenditure (CAPX) higher than 15% of capital (CAPX is item CAPXQ minus item SPPEQ calculated from year-to-date figures [missing SPPEQ set to zero]); (iv) $(I-\text{CAPX})/K \leq 5\%$, which excludes firm-quarters with the difference between gross investment and CAPX higher than 5% of capital; (v) Age$>$ 3, which excludes the first 12 quarters for a given firm; (vi) Age$>$ 5, which excludes the first 60 quarters for a given firm; (vii) small ME, the small market equity sample; (viii) big ME, the big market equity sample; (ix) Small K, the small capital sample; and (x) Big K, the big capital sample. For last four tests, we use the NYSE breakpoints and split all fiscal quarters ending in a calendar quarter into two groups based on the begin-of-quarter market equity or capital. In all experiments, we continue to winsorize each quarter at the 1%–99% level in the full sample to ensure comparability across subsamples.



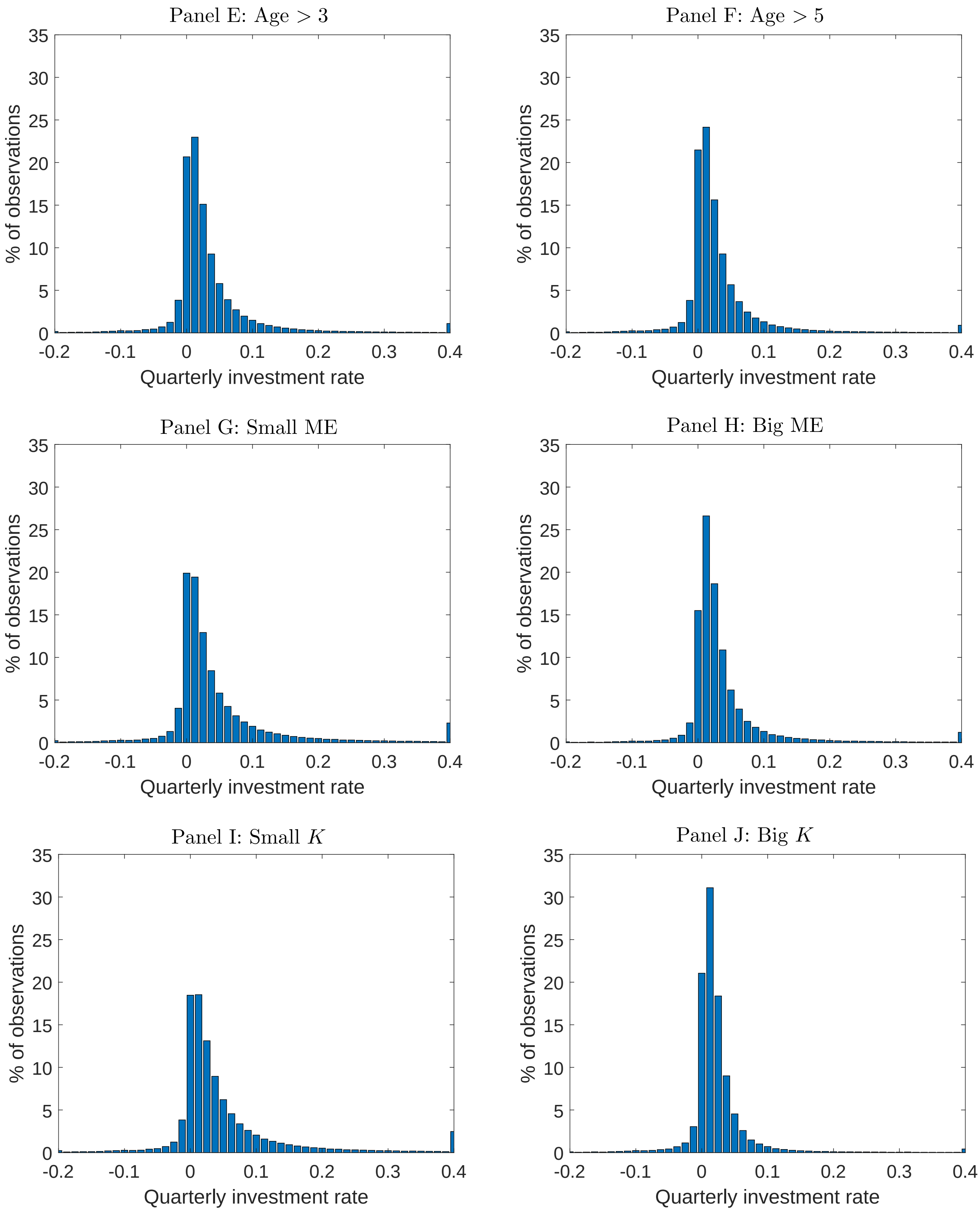
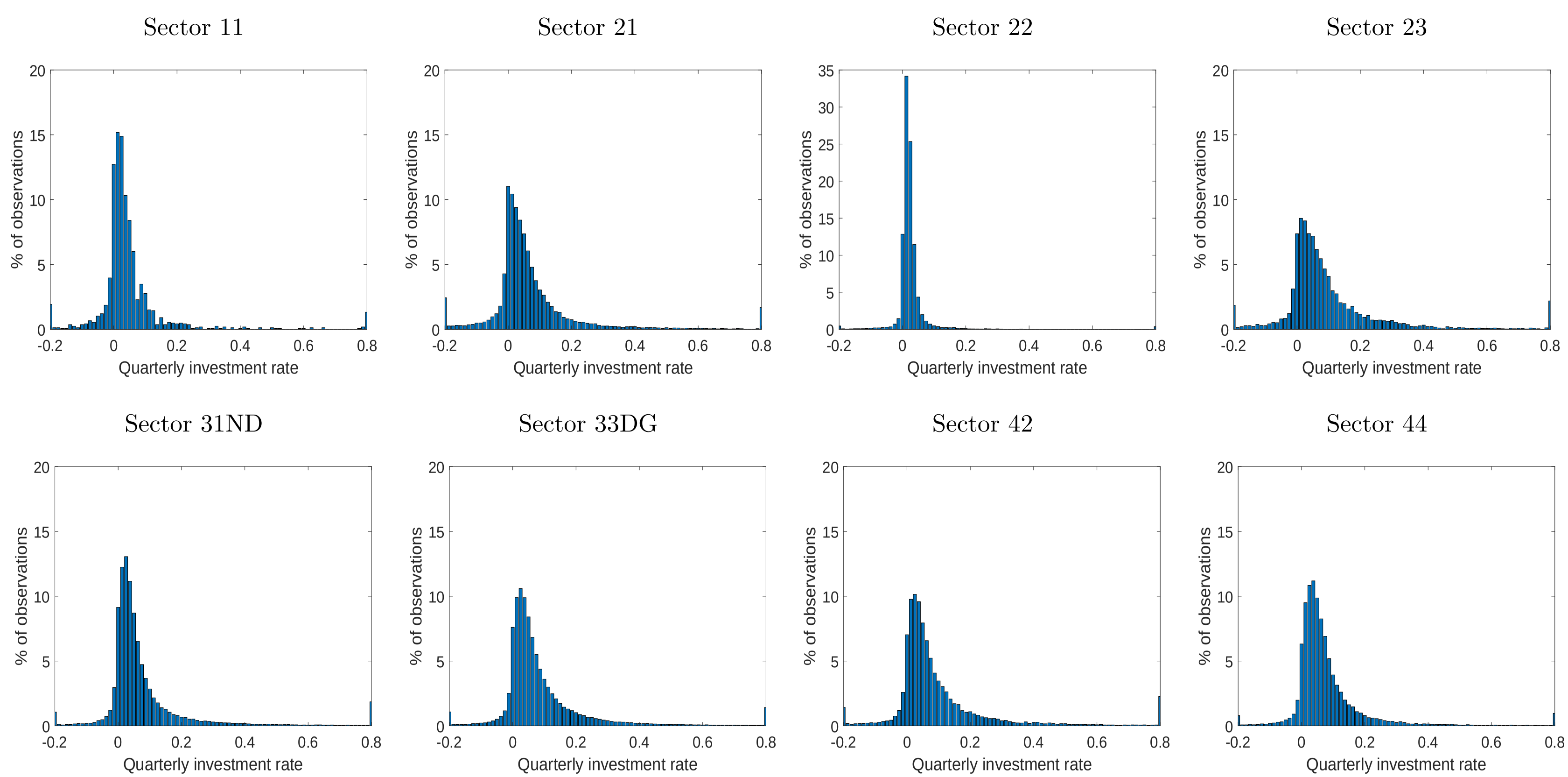
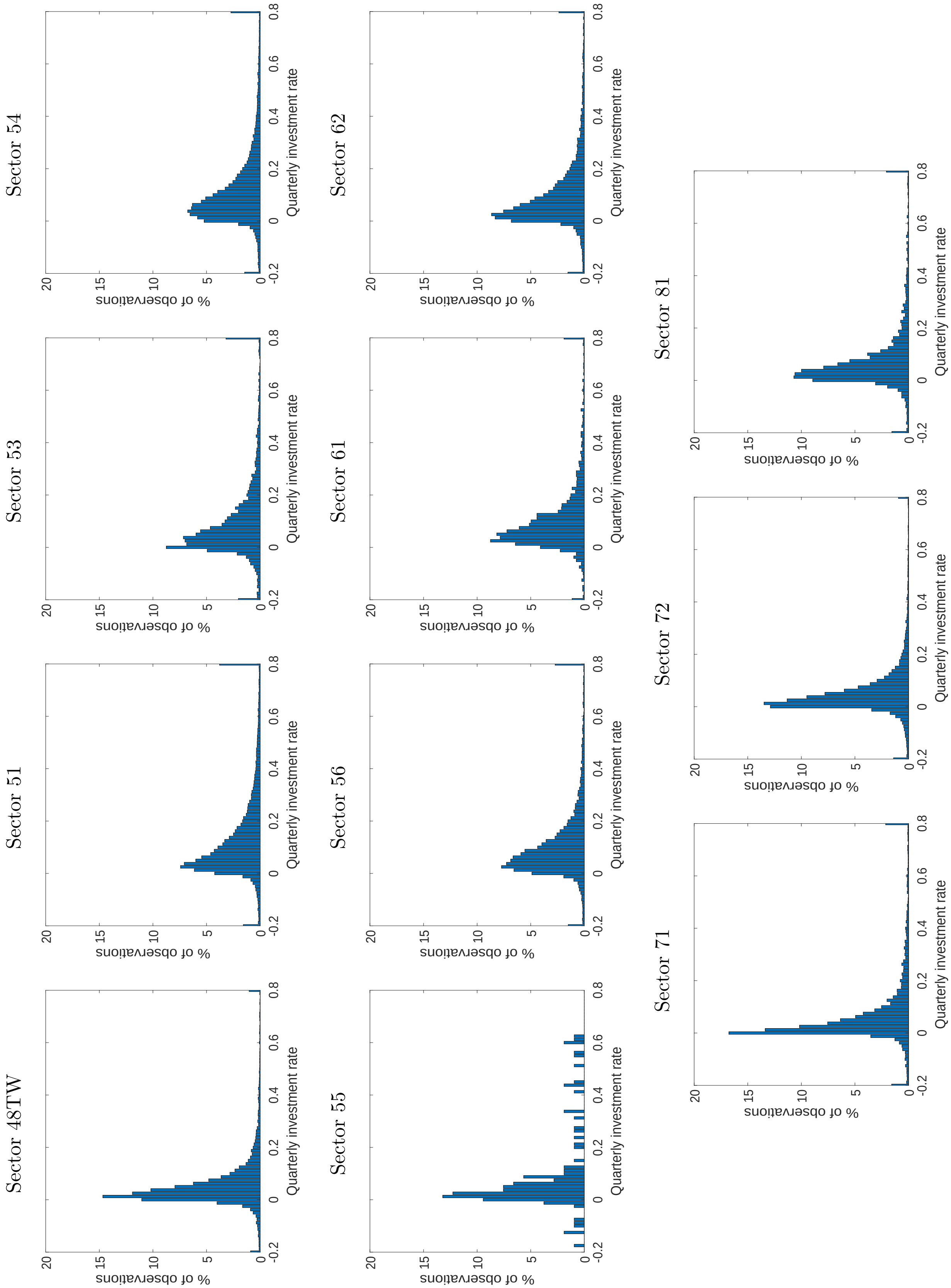

Panel E: Age > 3
Panel F: Age > 5
Panel G: Small ME
Panel H: Big ME
Panel I: Small K
Panel J: Big K
% of observations
Quarterly investment rate
0
5
10
15
20
25
30
35
-0.2
-0.1
0
0.1
0.2
0.3
0.4

Figure S3 : Robustness, the Quarterly Gross Investment Rate Distribution by 19 NAICS Nonfinancial Sectors, Scaled by Net PPE, 1978Q1–2016Q4

The 19 NAICS nonfinancial sectors are: Sector 11, Agriculture, forestry, fishing, and hunting; 21, Mining; 22, Utilities; 23, Construction; 31ND, Nondurable goods; 33DG, Durable goods; 42, Wholesale trade; 44, Retail trade; 48TW, Transportation and warehousing; 51, Information; 53, Real estate and rental and leasing; 54, Professional, scientific, and technical services; 55, Management of companies and enterprises; 56, Administrative and waste management services; 61, Educational services; 62, Health care and social assistance; 71, Arts, entertainment, and recreation; 72, Accommodation and food services; 81, Other services, except government.

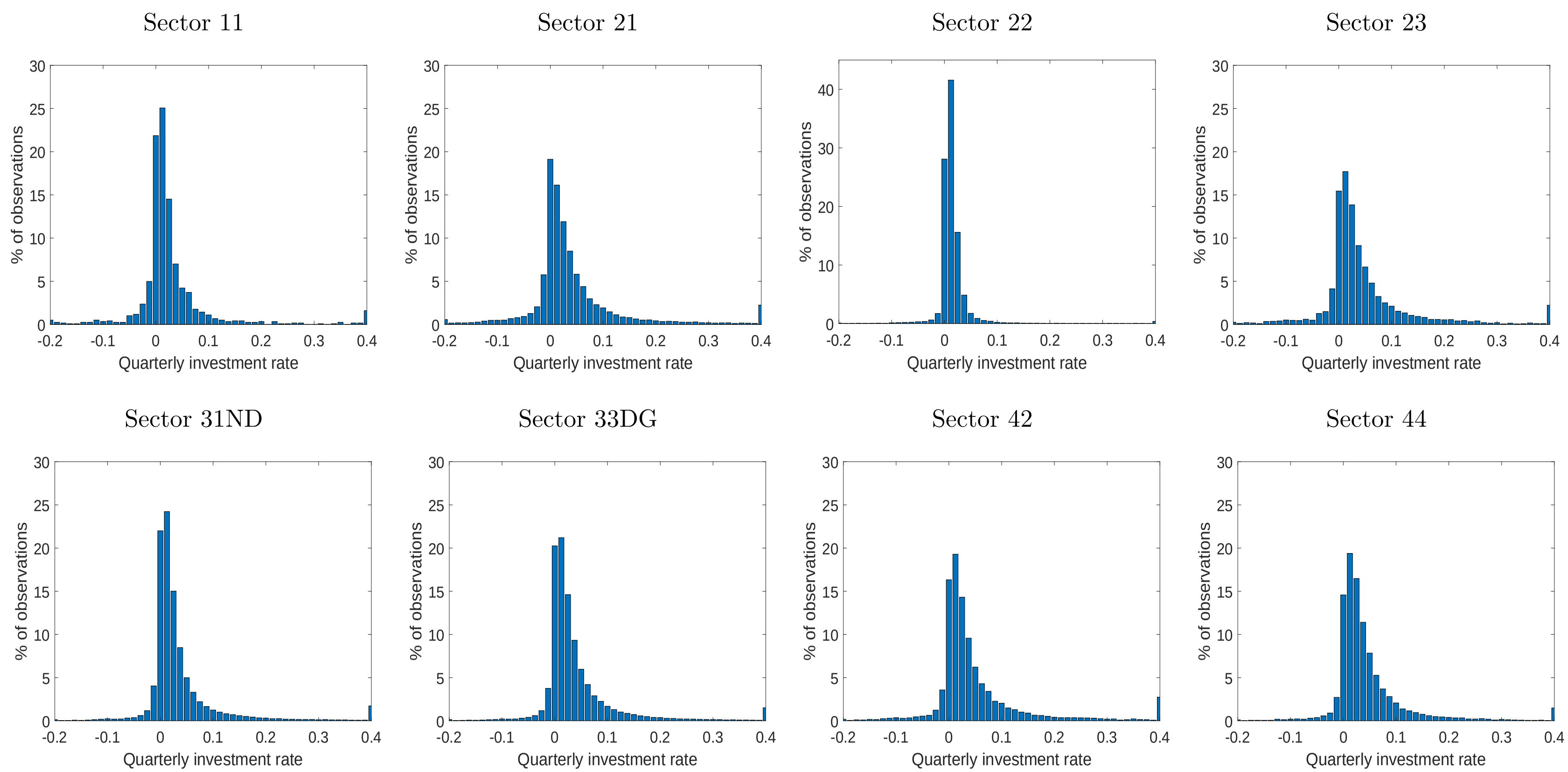


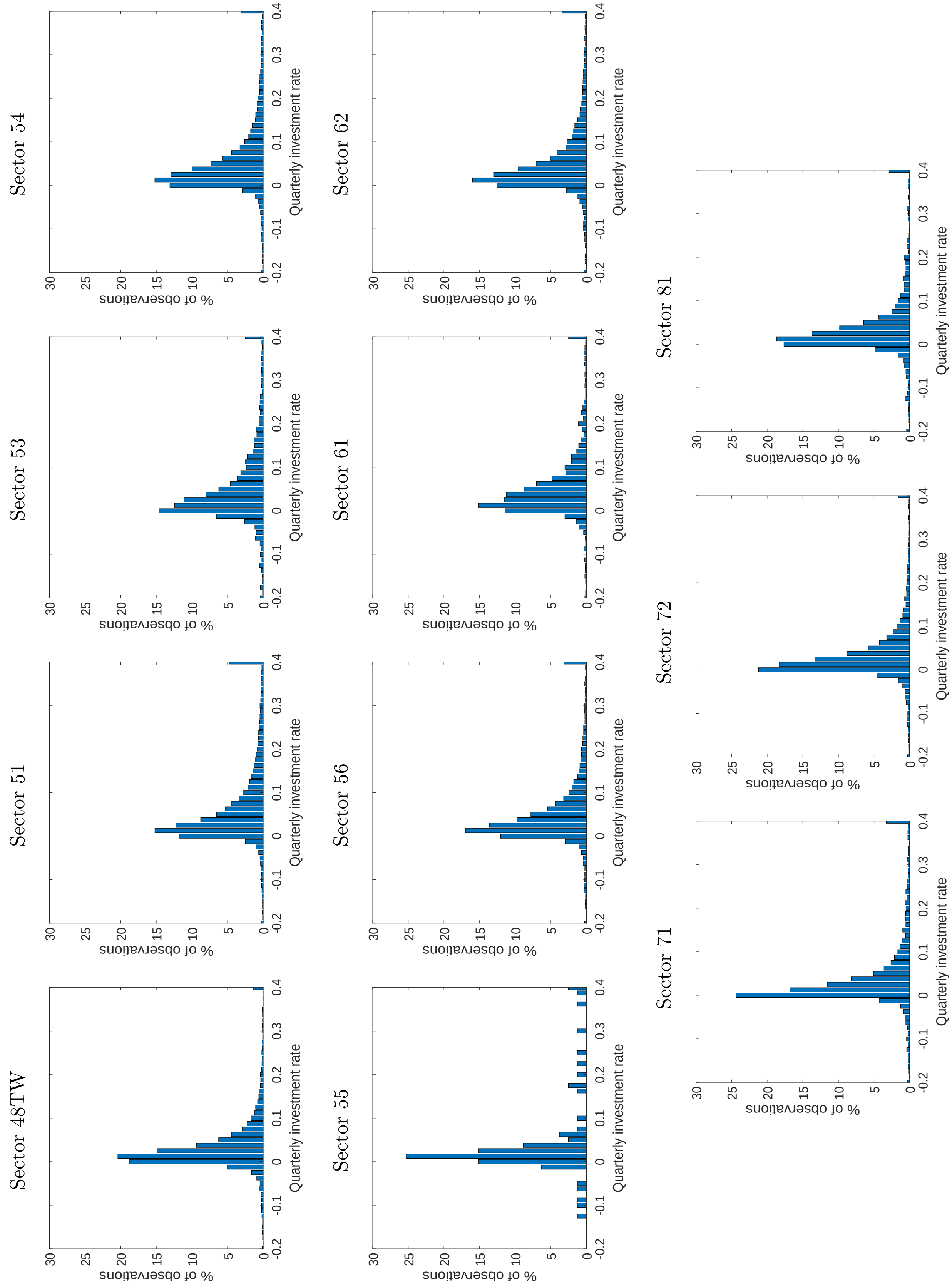


Sector 48TW
Sector 51
Sector 53
Sector 54
Sector 55
Sector 56
Sector 61
Sector 62
Sector 71
Sector 72
Sector 81
% of observations
Quarterly investment rate

Figure S4 : Robustness, the Quarterly Gross Investment Rate Distribution by 19 NAICS Nonfinancial Sectors, Scaled by Gross PPE, 1978Q1–2016Q4

The 19 NAICS nonfinancial sectors are: Sector 11, Agriculture, forestry, fishing, and hunting; 21, Mining; 22, Utilities; 23, Construction; 31ND, Nondurable goods; 33DG, Durable goods; 42, Wholesale trade; 44, Retail trade; 48TW, Transportation and warehousing; 51, Information; 53, Real estate and rental and leasing; 54, Professional, scientific, and technical services; 55, Management of companies and enterprises; 56, Administrative and waste management services; 61, Educational services; 62, Health care and social assistance; 71, Arts, entertainment, and recreation; 72, Accommodation and food services; 81, Other services, except government.




Sector 48TW
Sector 51
Sector 53
Sector 54
Sector 55
Sector 56
Sector 61
Sector 62
Sector 71
Sector 72
Sector 81
% of observations
Quarterly investment rate

D.1 "investment_rate_bea.do"

```
/*This program construct the investment rate in Clementi and Palazzo*/
clear
clear matrix

set more off
cd "/Users/administrator/Dropbox/Data_for_Lu"

/*Load Dataset: Quarterly CRSP-Compustat merged*/
use investment_data_may_2017_ccm2


/******************************/
/*Fiscal and Calendar Dates*****/
/******************************/
gen month=month(datadate)
gen year=year(datadate)

gen quarter=1 if month<4
replace quarter =2 if month>=4&month<7
replace quarter =3 if month>=7&month<10
replace quarter =4 if month>=10&month<=12

gen date = yq(year, quarter)
gen time = date
format time %tq


*ONLY USA INCORPORATED COMPANIES
*FIC: Incorporation Code - Foreign indicates the country in which a company is
incorporated.
drop if fic~="USA"


/*****************************************/
/*Drop if observations largerg than 2016q4*/
/*****************************************/
drop if date>=228

/*************************************/
/*Drop financials, utilities, and govt*/
/*************************************/
destring sic, replace force
drop if sic>=4900 & sic<=4999
drop if sic>=6000 & sic<=6999
drop if sic>=9000

/***************************/
/*Unique PERMCO within GVKEY*/
/***************************/
sort gvkey lpermco
```

```
by gvkey: gen aux=1 if lpermco~=lpermco[_n-1]&gvkey==gvkey[_n-1]
by gvkey: egen aux_max=max(aux)
drop if aux_max==1
drop aux*



/*********************************************************************/
/*Eliminate observations that have a double entry for the same GVKEY*/
/*********************************************************************/

*Each gvkey has multiple securities outstanding (i.e., multiple permno)
*permno is the issue identifier in CRSP, so we need to avoid double-counting
observations
*We count how many observations for each permno at the GVKEY level and then we keep
the
*accounting data relative to the permno with the largest numbers of available
observatons

*GVKEY is a unique six-digit number key assigned to each company

*PERMCO is a unique permanent company identification number assigned by CRSP to all
companies
*with issues on a CRSP File. This number is permanent for all securities issued by
this company regardless of name changes.

*PERMNO is a unique permanent security identification number assigned by CRSP to
each security.

*Count permno by gvkey
sort gvkey lpermno
by gvkey lpermno : gen number_obs=_N
sort gvkey datadate
*Drop all permno less than max permno
by gvkey : egen max_obs=max(number_obs )
drop if number_obs <max_obs

*What if we have a tie? Keep the permno with the lowest value (securiety issued
first)
sort gvkey  datadate lpermno
by gvkey  datadate: drop if lpermno~=lpermno[_n-1] &gvkey==gvkey[_n-1]

*If we still have multiple permno, we eliminate these obs
sort gvkey lpermno datadate
by gvkey: gen check=1 if lpermno ~=lpermno[_n-1]&gvkey ==gvkey[_n-1]
by gvkey: egen check_max=max(check)
drop if check_max==1
drop check*

*We can still have multiple identical entries that we eliminate
```

```
sort gvkey date
by gvkey: drop if date==date[_n-1]

*We end up with a dataset that has one security issue for each firm, we will
*match this security issue with CRSP data to perform the portfolio analysis


/*****************************************************/
/*Generate an indicator variable for fiscal year lags*/
/*****************************************************/
gen fdate = yq(fyearq, fqtr)
sort gvkey fdate
by gvkey: gen fdate_lag = fdate-fdate[_n-1]
/*Drop multiple observations within a fiscal quarter*/
by gvkey fdate, sort: drop if fdate==fdate[_n-1]


/*****************************************************************/
/*Check that we have one observation per gvkey per fiscal quarter*/
/*****************************************************************/
destring gvkey, replace
sort gvkey fdate
by gvkey fdate: egen check=count(gvkey)
tab check
drop check


/*****************************/
//Create book value of equity*/
/*****************************/

//Definition from French's website:
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/variable_defini
tions.html
//Book equity is constructed from Compustat data or collected from the Moody’s
Industrial, Financial, and Utilities manuals.
//BE is the book value of stockholders’ equity, plus balance sheet deferred taxes
and investment tax credit (if available), minus the book value of preferred stock.
//Depending on availability, we use the redemption, liquidation, or par value (in
that order) to estimate the book value of preferred stock.
//Stockholders’ equity is the value reported by Moody’s or Compustat, if it is
available. If not, we measure stockholders’ equity as the book value of common
equity
//plus the par value of preferred stock, or the book value of assets minus total
liabilities (in that order).
//See Davis, Fama, and French, 2000, “Characteristics, Covariances, and Average
Returns: 1929-1997,” Journal of Finance, for more details.

//Definition from Frazzini
//To obtain shareholders’ equity we use we use Stockholders’ Equity (SEQ) but if
```

```
not available,
//we use the sum of Common Equity (CEQ) and Preferred Stocks (PSTK). If both SEQ
and CEQ are unavailable,
//we proxy shareholders' equity by Total Assets (AT) minus the sum of Total
Liability (LT) and Minority Interest (MIB). To obtain book equity (BE),
//we subtract from shareholders' equity the preferred stock value (PSTKRV, PSTKL or
PSTK depending on availability).

//We follow Palazzo (2012) to construct BE at quarterly frequency

//BE =  Shareholders' equity + deferred taxes and investment tax credits - book
value of preferred stock
rename ceqq equity_common
rename seqq equity
rename txditcq tax_credit
rename pstkq preferred_par
rename pstkrq preferred_redemption
replace tax_credit=0 if  tax_credit==.


//First we build shareholders'equity
gen     shareholders_equity = equity
replace shareholders_equity= equity_common+preferred_par if
shareholders_equity==.&equity_common~=.&preferred_par~=.
replace shareholders_equity= atq-ltq if shareholders_equity==.&atq~=.&ltq~=.

//Then we add balance sheet deferred taxes and investment tax credit
replace shareholders_equity=shareholders_equity+ tax_credit


//In the last step we subtract the preferred stock value
gen     be= shareholders_equity -preferred_redemption     if
shareholders_equity~=.&preferred_redemption~=.
replace be =shareholders_equity -preferred_par            if
shareholders_equity~=.&preferred_redemption==.&preferred_par~=.

keep   datadate date fdate fdate_lag gvkey capxy sppey ppegtq ppentq prccq cshoq
aqcy atq   fqtr fyearq sic dp* be rdq lperm* /*
*/ cusip exchg tax_credit fyearq fqtr cheq preferred_par dlcq dlttq actq invtq


/*******************************************************************/
/**************CREATE THE NET INVESTMENT FLOW VARIABLE**************/
/*******************************************************************/


/*************************************************************/
/*Convert year-to-date (cumulative) data into quarterly flows*/
/*************************************************************/
```

```
*MERGE DEFLATOR DATA
sort date
merge date using deflator
drop if _merge~=3
drop _merge

//PPENT AND PPEGT IN REAL TERMS USING NON RESIDENTIAL FIXED ASSET DEFLATOR
replace ppent=100*ppent/deflator
replace ppegt=100*ppegt/deflator
replace capxy=100*capxy/deflator
replace sppey=100*sppey/deflator


sort gvkey fdate

/*Replace missing observations between ppen_{t-1} and ppen_{t+1}, where ppen_{t-1}
and ppen_{t+1} are non missing*/

*before replacement  585,871 non-missing values
gen dummy=0
by gvkey: replace dummy=1 if ppent==.&ppent[_n-1]~=.
by gvkey: replace dummy=0 if ppent==.&ppent[_n+1]>=.

replace ppent=(ppent[_n-1]+ppent[_n+1])/2 if dummy==1

*before replacement  587,248 non-missing values. We added 1,377 observations
(0.235%)

sort gvkey fdate
/*Gross investment: INV */
gen capx_gross=ppegt-ppegt[_n-1] if fdate_lag==1

/*NET investment:  INV - DEP */
gen capx_net=ppent-ppent[_n-1] if fdate_lag==1


/********************************************/
/* Generate the acquisition to assets ratio*/
/********************************************/
sort gvkey fyearq fqtr
gen aqc = aqcy
replace aqc = aqcy - aqcy[_n-1] if  fqtr~=1&fdate_lag==1
replace aqc = 0 if aqc==.
gen aqc_ratio=aqc/atq[_n-1] if fdate_lag==1
replace aqc_ratio=0 if aqc_ratio==.

replace aqcy = 0 if aqcy==.
sort gvkey fdate
gen aqc_ratio_annual=aqcy/atq[_n-4] if  fdate==fdate[_n-4]+4&fqtr==4
replace aqc_ratio_annual=0 if aqc_ratio_annual==.&fdate==fdate[_n-4]+4&fqtr==4
```

```
//Generate the investment flow using CAPEX and SPPE for comparison purposes
sort gvkey fyearq fqtr
gen capxq = capxy
by gvkey: replace capxq = capxy - capxy[_n-1] if  fqtr~=1&fdate_lag==1
gen sppeq = sppey
by gvkey: replace sppeq = sppey - sppey[_n-1] if  fqtr~=1&fdate_lag==1

keep   datadate date deflator fdate fdate_lag gvkey  sic ppentq ppegtq   capx_net
capx_gross/*
*/ capxq sppeq capxy sppey prccq cshoq  aqc_ratio*   dp* be rdq lperm* cusip exchg
fyearq fqtr tax_credit cheq preferred_par invtq dlcq dlttq actq
sort gvkey date

drop if capx_net==.
save temp, replace


/*********************************************************************/
/*************CREATE THE CAPITAL STOCK AT THE FIRM LEVEL*************/
/*********************************************************************/

gen k_stock=ppentq
drop if k_stock==.

sort gvkey date
replace k_stock=. if gvkey==gvkey[_n-1]

gen dummy=1 if k_stock~=.
drop if k_stock==.

keep gvkey date k_stock dummy
sort gvkey date

merge gvkey date using temp
sort gvkey date

drop _merge

drop fdate_lag
sort gvkey fdate
by gvkey: gen fdate_lag=fdate-fdate[_n-1]



//sort gvkey fdate
//Generate the capital stock recursively
by gvkey: replace k_stock=k_stock[_n-1] +capx_net if dummy~=1&fdate_lag==1
```

```
//Calculate number of observations
by gvkey : gen obs= _n

//Merge using BEA deflators
gen sic3=floor(sic/10)
sort sic3
merge sic3 using delta_bea

drop if _merge==2
drop _merge

//If delta at the 3-digit SIC code is missing we calculate the average at the
2-digit level
gen sic2=floor(sic/100)
bysort sic2: egen average_delta_sic2=mean(delta_bea)

replace delta_bea=average_delta_sic2 if delta_bea==.

//If delta at the 2-digit SIC code is missing we calculate the average at the
1-digit level
gen sic1=floor(sic/1000)
bysort sic1: egen average_delta_sic1=mean(delta_bea)

//If delta still missing, we use the unconditional mean
egen average_delta=mean(delta_bea)
replace delta_bea=average_delta if delta_bea==.


//Accounting depreciation for comparison purposes
sort gvkey fdate
by gvkey: gen delta_accounting=(capx_gross-capx_net)/k_stock[_n-1] if fdate_lag==1
egen   upperbound= pctile(delta_accounting) , p(99.5)
egen   lowerbound= pctile(delta_accounting) , p(0.5)
replace delta_accounting=. if delta_accounting<lowerbound
replace delta_accounting=. if delta_accounting>upperbound
drop upperbound lowerbound

drop average_delta* sic1 sic2 sic3

//Investment rate
drop capx_gross
sort gvkey fdate
by gvkey: gen capx_gross=capx_net+delta_bea*k_stock[_n-1] if fdate_lag==1
by gvkey: gen inv_rate_recursive   =  capx_gross/k_stock[_n-1]  if fdate_lag==1
by gvkey: gen inv_rate_annual   =  (capxy-sppey)/ppentq  if fdate==fdate[_n-4]+4
by gvkey: gen inv_rate_accounting   =  (capxq-sppeq)/ppentq[_n-1]  if fdate_lag==1
by gvkey: gen inv_rate_accounting_annual   =  (capxy-sppey)/ppentq[_n-4]  if
fdate==fdate[_n-4]+4
```

```
gen time = date
format time %tq

save investment_rate_data_bea, replace

/********************************************************************/
/********************************************************************/
/********************************************************************/
/*********************COMPUTE STATISTICS*************************/
/********************************************************************/
/********************************************************************/
/********************************************************************/


cd
"/Users/administrator/Desktop/work/Research/Investment_equity_returns/data/quarterl
y_analysis/May_2017/inv_rate"

use investment_rate_data_bea, clear

/*********************************************************************************/
/*Drop top and bottop 0.5% for investmet rates, now we work wiht truncated sample*/
/*********************************************************************************/

//Drop first 12 quarters
replace inv_rate_recursive=.  if obs <=12

//Set sample from 1978q1 to 2016q4 (168 quarters)
drop if date<72
drop if date>227

//Drop negative k_stock and investment rate less tha 100%
replace inv_rate_recursive=.  if k_stock<=0
replace inv_rate_recursive=.  if inv_rate_recursive<=-1

//Drop if acquisition larger than 5% of assets in a given quarter
drop if aqc_ratio  >0.05
drop if aqc_ratio  <-0.05

//Eliminate top and bottom 0.5% to dampen impact of outliers
egen   upperbound= pctile(inv_rate_recursive) , p(99.5)
egen   lowerbound= pctile(inv_rate_recursive) , p(0.5)
replace inv_rate_recursive=. if inv_rate_recursive>upperbound
replace inv_rate_recursive=. if inv_rate_recursive<lowerbound
drop upperbound lowerbound
gen inv_rate= inv_rate_recursive
drop if inv_rate==.


egen   upperbound= pctile(inv_rate_accounting) , p(99.5)
```

```
egen   lowerbound= pctile(inv_rate_accounting) , p(0.5)
replace inv_rate_accounting=. if inv_rate_accounting>upperbound
replace inv_rate_accounting=. if inv_rate_accounting<lowerbound
drop upperbound lowerbound


//Here we calculate the book-to-market ratio using an alternative approach used in
empirical asset pricing
//The market value of equity is common shares outstanding times the fiscal end
stock price
gen market_value=100*(prccq*cshoq+dlttq+preferred_par-invtq)/deflator
gen bm = (k_stock)/market_value


replace bm=. if bm<=0
egen   upperbound= pctile(bm) , p(99.5)
egen   lowerbound= pctile(bm) , p(0.5)
replace bm=. if bm>upperbound
replace bm=. if bm<lowerbound
drop upperbound lowerbound

sort gvkey date
by gvkey : drop if date==date[_n-1]

/********************************************/
/*Create an histogram for the investment rate*/
/********************************************/
gen hist = inv_rate
replace hist = . if hist >0.2
replace hist = . if hist <-0.2

histogram hist, bin(51) percent xtitle("Investment Rate", size(medlarge))
ytitle("Percentage", size(medlarge)) plotregion(style(none)) ylabel(, angle(0))
graph export ir_distribution_bea.eps, replace

histogram hist, bin(5) percent xtitle("Investment Rate", size(medlarge))
ytitle("Percentage", size(medlarge)) plotregion(style(none)) ylabel(, angle(0))
graph export ir_distribution_bea_large_bins.eps, replace


histogram delta_bea, bin(5) percent xtitle("Depreciation Rate", size(medlarge))
ytitle("Percentage", size(medlarge)) plotregion(style(none)) ylabel(, angle(0))

graph export depreciation_bea.eps, replace


//Clean annual investment rate
//Only consider end of the fiscal year values
replace inv_rate_annual =. if fqtr ~=4
replace inv_rate_accounting_annual =. if fqtr ~=4
```

```
replace inv_rate_annual =. if aqc_ratio_annual>0.05&aqc_ratio_annual<.
replace inv_rate_annual =. if aqc_ratio_annual<-0.05

replace inv_rate_accounting_annual =. if aqc_ratio_annual>0.05&aqc_ratio_annual<.
replace inv_rate_accounting_annual =. if aqc_ratio_annual<-0.05

egen   upperbound= pctile(inv_rate_annual) , p(99.5)
egen   lowerbound= pctile(inv_rate_annual) , p(0.5)
replace inv_rate_annual=. if inv_rate_annual>upperbound
replace inv_rate_annual=. if inv_rate_annual<lowerbound
drop upperbound lowerbound

egen   upperbound= pctile(inv_rate_accounting_annual) , p(99.5)
egen   lowerbound= pctile(inv_rate_accounting_annual) , p(0.5)
replace inv_rate_accounting_annual=. if inv_rate_accounting_annual>upperbound
replace inv_rate_accounting_annual=. if inv_rate_accounting_annual<lowerbound
drop upperbound lowerbound

sort gvkey date


save data_firm_level_bea, replace

/*******************************/
/*cross--sectional observations*/
/*******************************/
egen obs_inv_rate = count(inv_rate) , by(date)

/**********************/
/*cross--sectional mean*/
/**********************/
by date, sort: egen mean_bm = mean(bm)
by date, sort: egen mean_inv_rate_aux = mean(inv_rate)
by date, sort: egen mean_inv_rate   = max(mean_inv_rate_aux)
by date, sort: egen mean_inv_rate_pos_aux = mean(inv_rate) if inv_rate>0.01
by date, sort: egen mean_inv_rate_pos  = max(mean_inv_rate_pos_aux)
by date, sort: egen mean_inv_rate_neg_aux = mean(inv_rate) if inv_rate<-0.01
by date, sort: egen mean_inv_rate_neg  = max(mean_inv_rate_neg_aux)

/***********************************/
/*cross--sectional standard deviation*/
/***********************************/

by date, sort: egen sd_inv_rate_aux = sd(inv_rate)
by date, sort: egen sd_inv_rate   = max(sd_inv_rate_aux)

/************************/
/*cross--sectional median*/
/************************/
by date, sort: egen median_inv_rate = median(inv_rate)
```

```
by date, sort: egen median_inv_rate_pos_aux = median(inv_rate) if inv_rate>0.01
by date, sort: egen median_inv_rate_pos  = max(median_inv_rate_pos_aux)
by date, sort: egen median_inv_rate_neg_aux = median(inv_rate) if inv_rate<-0.01
by date, sort: egen median_inv_rate_neg  = max(median_inv_rate_neg_aux)

/*******************************/
/*cross--sectional inaction rate*/
/*******************************/
egen inaction_inv_rate = count(inv_rate) if abs(inv_rate)<0.01, by(date)
bysort date: gen inaction_rate_inv_rate_aux = inaction_inv_rate/obs_inv_rate
bysort date: egen inaction_rate_inv_rate = mean(inaction_rate_inv_rate_aux)

/*************************************/
/*cross--sectional negative investment*/
/*************************************/
egen neg_inv_rate = count(inv_rate) if inv_rate<-0.01 , by(date)
bysort date: gen neg_rate_inv_rate_aux = neg_inv_rate/obs_inv_rate
bysort date: egen neg_rate_inv_rate = mean(neg_rate_inv_rate_aux)

/*********************************/
/*cross--sectional positive spikes*/
/*********************************/
egen posspike_inv_rate = count(inv_rate) if inv_rate>=0.2 , by(date)
bysort date: gen posspike_rate_inv_rate_aux = posspike_inv_rate/obs_inv_rate
bysort date: egen posspike_rate_inv_rate = mean(posspike_rate_inv_rate_aux)

/*********************************/
/*cross--sectional negative spikes*/
/*********************************/
egen negspike_inv_rate = count(inv_rate) if inv_rate<=-0.2 , by(date)
bysort date: gen negspike_rate_inv_rate_aux = negspike_inv_rate/obs_inv_rate
bysort date: egen negspike_rate_inv_rate = mean(negspike_rate_inv_rate_aux)

/**********************************************/
/*Evaluate the investment rate serial correlation*/
/**********************************************/
drop fdate_lag
sort gvkey fdate
by gvkey: gen fdate_lag = fdate-fdate[_n-1]
by gvkey: gen inv_rate_lag = inv_rate[_n-1] if fdate_lag==1

/*************/
/*Fama-MacBeth*/
/*************/
gen serialcorr_coeff=0
gen serialcorr_const=0
sort date
set more off
   forval t=73 (1) 227 {
```

```
        reg inv_rate inv_rate_lag if date==`t'
        replace serialcorr_coeff = _b[inv_rate_lag] if date==`t'
        replace serialcorr_const = _b[_cons]  if date==`t'
   }

/************/
/*Pooled OLS*/
/************/
reg inv_rate inv_rate_lag
gen serialcorr_coeff_OLS=_b[inv_rate_lag]
gen serialcorr_const_OLS = _b[_cons]

/*****************************/
/*Create quarterly time-series*/
/*****************************/

sort  date
drop if date==date[_n-1]

sort date

replace negspike_rate_inv_rate=0 if negspike_rate_inv_rate==.
replace posspike_rate_inv_rate=0 if posspike_rate_inv_rate==.

keep date serialcorr* obs_inv_rate mean_inv_rate median_inv_rate_pos
median_inv_rate_neg sd_inv_rate median_inv_rate/*
*/   inaction_rate_inv_rate neg_rate_inv_rate  posspike_rate_inv_rate /*
*/   negspike_rate_inv_rate time mean_inv_rate_pos mean_inv_rate_neg mean_bm exchg
sic


/***************/
/*Export output*/
/***************/

format %6.0g mean_inv_rate sd_inv_rate median_inv_rate obs_inv_rate/*
*/     inaction_rate_inv_rate neg_rate_inv_rate posspike_rate_inv_rate
mean_inv_rate_pos mean_inv_rate_neg/*
*/     negspike_rate_inv_rate serialcorr_coeff_OLS serialcorr_const_OLS
median_inv_rate_pos median_inv_rate_neg mean_bm

sum

save investment_rate_stats_bea, replace
```

```
/*This program prepares the accounting data for portfolio formation*/

clear
clear matrix

cd "/Users/administrator/Dropbox/Data_for_Lu"

use data_firm_level_bea, clear



//Calendar dates
gen calendar_year=year(datadate)
gen calendar_quarter=quarter(datadate)

//Calendar date at quarterly frequency
gen dateq =yq(calendar_year,calendar_quarter)

//We move the quarterly date 2 quarters ahead in time to match with prices in
quarters at t+2
//This allows a six-month lag between CRSP data and accounting data
replace dateq=dateq+2

keep be inv_rate*   cusip lpermno lpermco dateq sic bm dlttq dlcq preferred_par
invtq k_stock  delta_bea datadate

replace preferred_par=0 if preferred_par==.
replace invtq=0 if invtq==.
replace dlcq =0 if dlcq ==.
replace dlttq=0 if dlttq==.

rename lpermno permno

//We drop missing and negative be observations
drop if be<0
drop if be==.

sort permno dateq

save accounting_data_cleaned.dta, replace
```

```
/*This program cleans the raw dataset*/
clear
clear matrix

set mem 2g
set matsize 800
set more off
cd "/Users/administrator/Dropbox/Data_for_Lu"
use  CRSP_may_2017

/*************************/
/*Generate Time variables*/
/*************************/
gen calendar_year=year(date)
gen calendar_month=month(date)
gen calendar_quarter=quarter(date)


/******************************/
/*Generate Financial variables*/
/******************************/
destring ret, replace force
gen size=abs(prc)*shrout/1000
gen price=abs(prc)
gen shares=shrout/1000
destring ret, replace force



/*****************************/
/*Drop duplicate rows in CRSP*/
/*****************************/
sort  permno calendar_year calendar_month
drop if date==date[_n-1]



/*****************************************/
/*We only consider ordinary common shares*/
/*****************************************/
drop if  shrcd~=10& shrcd~=11

/**************************************************/
/*We exclude observations relative to suspended ,
  halted or non listed shares in NYSE AMEX NASDAQ */
/**************************************************/
sort permno
by permno: egen max_exchcd=max(exchcd)
by permno: egen min_exchcd=min(exchcd)
```

```
drop if  min_exchcd<=0
drop if  max_exchcd>=4

destring dlret, replace force

/****************************************************/
/*Replace delisting return if missing following Shumway*/
/*If poor performance delisting return is missing I use*/
/*an average value of -30\%                          */
/****************************************************/
replace dlret=-0.30 if dlstcd==500&dlret==.
replace dlret=-0.30 if dlstcd==520&dlret==.
replace dlret=-0.30 if dlstcd==550&dlret==.
replace dlret=-0.30 if dlstcd==551&dlret==.
replace dlret=-0.30 if dlstcd==552&dlret==.
replace dlret=-0.30 if dlstcd==560&dlret==.
replace dlret=-0.30 if dlstcd==561&dlret==.
replace dlret=-0.30 if dlstcd==570&dlret==.
replace dlret=-0.30 if dlstcd==572&dlret==.
replace dlret=-0.30 if dlstcd==574&dlret==.
replace dlret=-0.30 if dlstcd==575&dlret==.
replace dlret=-0.30 if dlstcd==580&dlret==.
replace dlret=-0.30 if dlstcd==581&dlret==.
replace dlret=-0.30 if dlstcd==582&dlret==.
replace dlret=-0.30 if dlstcd==583&dlret==.
replace dlret=-0.30 if dlstcd==584&dlret==.
replace dlret=-0.30 if dlstcd==585&dlret==.
replace dlret=-0.30 if dlstcd==587&dlret==.
replace dlret=-0.30 if dlstcd==589&dlret==.
replace dlret=-0.30 if dlstcd==591&dlret==.

/****************************************************/
/*Replace delisting return if missing following Shumway*/
/*If NOT poor performance delisting return is missing  */
/*I use an average value of -100\%                     */
/****************************************************/
replace dlret=-1.00 if dlstcd>100&dlret==.&dlstcd~=.
/*******************************************************/
/*Assign the delisting return as the last available return*/
/*******************************************************/
replace ret= dlret if dlret~=.

//Create date at monthly frequency
drop date
gen date=ym(calendar_year, calendar_month)

//Create market cap lagged seven months, needed to calculate boof-to-market
sort permno calendar_year calendar_month
by permno: gen lag_size7=size[_n-7] if date==date[_n-7]+7
```

```
//Generate market cap in real values, deflator is CPI
sort date
merge date using CPI

drop if _merge~=3
drop _merge

replace size=100*size/cpi

//Generate one-month lagged market cap, needed for value weighting returns
sort permno date
by permno: gen lag_size=size[_n-1] if date==date[_n-1]+1

replace ret=100*ret

//Create date at quarterly frequency, needed to match with accounting data
gen dateq=yq(calendar_year, calendar_quarter)

sort permno dateq

save  CRSP_cleaned, replace
```

D.4 "merge_crsp_compustat.do"

```
/*This program cleans the raw dataset*/
clear
clear matrix

set mem 2g
set matsize 800
set more off

cd "/Users/administrator/Dropbox/Data_for_Lu"
use accounting_data_cleaned.dta

rename cusip cusip_cs


sort permno dateq

merge permno dateq  using CRSP_cleaned

tab _merge
drop if _merge~=3
drop _merge

//Calculate excess returns
sort date
merge date using FF_data
drop if _merge~=3
drop _merge

replace ret=ret-100*rf

//Generate market cap in real values, deflator is CPI
sort date
merge date using CPI

drop if _merge~=3
drop _merge



//Calculate book-to-market
sort permno date
by permno: gen be_me=be/lag_size7

drop bm
sort permno date
by permno: gen bm=k_stock/(100*(lag_size7+dlttq+dlcq+preferred_par-invtq)/cpi)

//Set size and inv_rate to missing if there is no available book-to-market
replace lag_size=. if be_me==.
replace inv_rate=. if be_me==.
```

```
replace bm=. if be_me==.
replace delta_bea=. if be_me==.

//This is log of real size
gen log_size=log(lag_size)

//NYSE dummy
gen     NYSE=0
replace NYSE=1 if hexcd==1


drop   hexcd   price

/*******************************************************************************/
/*Drop observations that do not match cusip from CS to cusip or ncusip from CRSP*/
/*******************************************************************************/

gen cusip6_cs=substr(cusip_cs,1,6)
gen ncusip6_crsp=substr(ncusip,1,6)
gen cusip6_crsp=substr(cusip,1,6)

gen dummy=1 if  ncusip6_crsp~=cusip6_cs
gen dummy1=1 if cusip6_crsp~=cusip6_cs
gen dummy2=dummy+dummy1
drop if dummy2~=.
drop dummy*

//Drop negative or zero values for book-market
drop if be_me<=0
winsor be_me, gen(aux) p(0.005)
replace be_me=aux
drop aux

replace bm=. if bm<=0
winsor bm, gen(aux) p(0.005)
replace bm=aux
drop aux


replace be_me=. if
calendar_month~=1&calendar_month~=4&calendar_month~=7&calendar_month~=10
replace inv_rate=. if
calendar_month~=1&calendar_month~=4&calendar_month~=7&calendar_month~=10
replace inv_rate_accounting=. if
calendar_month~=1&calendar_month~=4&calendar_month~=7&calendar_month~=10
replace inv_rate_annual=. if
calendar_month~=1&calendar_month~=4&calendar_month~=7&calendar_month~=10
replace delta_bea=. if
calendar_month~=1&calendar_month~=4&calendar_month~=7&calendar_month~=10
replace bm=. if
```

```
calendar_month~=1&calendar_month~=4&calendar_month~=7&calendar_month~=10
gen size_at_portfolio=lag_size if be_me~=.


keep permno lpermco  datadate date* calendar_month calendar_year log_size  ret
lag_size be_me  NYSE inv_rate inv_rate_a* lag_size sic bm size_at_portfolio
delta_bea k_stock

sort permno calendar_year calendar_month
save merged_crsp_compustat_data, replace
```

```
/*This program cleans the raw dataset*/
clear
clear matrix

set mem 2g
set matsize 800
set more off
cd "/Users/administrator/Dropbox/Data_for_Lu"
use merged_crsp_compustat_data, clear

drop if inv_rate==.
drop if be_me==.

drop date
rename dateq date

/********************************/
/*cross--sectional observations*/
/********************************/
egen obs_inv_rate = count(inv_rate) , by(date)

/**********************/
/*cross--sectional mean*/
/**********************/
by date, sort: egen mean_bm = mean(be_me)
by date, sort: egen mean_inv_rate_aux = mean(inv_rate)
by date, sort: egen mean_inv_rate   = max(mean_inv_rate_aux)
by date, sort: egen mean_inv_rate_pos_aux = mean(inv_rate) if inv_rate>0.01
by date, sort: egen mean_inv_rate_pos  = max(mean_inv_rate_pos_aux)
by date, sort: egen mean_inv_rate_neg_aux = mean(inv_rate) if inv_rate<-0.01
by date, sort: egen mean_inv_rate_neg  = max(mean_inv_rate_neg_aux)

/************************************/
/*cross--sectional standard deviation*/
/************************************/

by date, sort: egen sd_inv_rate_aux = sd(inv_rate)
by date, sort: egen sd_inv_rate   = max(sd_inv_rate_aux)

/***********************/
/*cross--sectional median*/
/***********************/
by date, sort: egen median_inv_rate = median(inv_rate)
by date, sort: egen median_inv_rate_pos_aux = median(inv_rate) if inv_rate>0.01
by date, sort: egen median_inv_rate_pos  = max(median_inv_rate_pos_aux)
by date, sort: egen median_inv_rate_neg_aux = median(inv_rate) if inv_rate<-0.01
by date, sort: egen median_inv_rate_neg  = max(median_inv_rate_neg_aux)

/********************************/
/*cross--sectional inaction rate*/
```

```
/*********************************/
egen inaction_inv_rate = count(inv_rate) if abs(inv_rate)<0.01, by(date)
bysort date: gen inaction_rate_inv_rate_aux = inaction_inv_rate/obs_inv_rate
bysort date: egen inaction_rate_inv_rate = mean(inaction_rate_inv_rate_aux)

/**************************************/
/*cross--sectional negative investment*/
/**************************************/
egen neg_inv_rate = count(inv_rate) if inv_rate<-0.01 , by(date)
bysort date: gen neg_rate_inv_rate_aux = neg_inv_rate/obs_inv_rate
bysort date: egen neg_rate_inv_rate = mean(neg_rate_inv_rate_aux)

/**********************************/
/*cross--sectional positive spikes*/
/**********************************/
egen posspike_inv_rate = count(inv_rate) if inv_rate>=0.2 , by(date)
bysort date: gen posspike_rate_inv_rate_aux = posspike_inv_rate/obs_inv_rate
bysort date: egen posspike_rate_inv_rate = mean(posspike_rate_inv_rate_aux)

/**********************************/
/*cross--sectional negative spikes*/
/**********************************/
egen negspike_inv_rate = count(inv_rate) if inv_rate<=-0.2 , by(date)
bysort date: gen negspike_rate_inv_rate_aux = negspike_inv_rate/obs_inv_rate
bysort date: egen negspike_rate_inv_rate = mean(negspike_rate_inv_rate_aux)

/***********************************************/
/*Evaluate the investment rate serial correlation*/
/***********************************************/
sort permno date
by permno: gen date_lag =date-date[_n-1]
by permno: gen inv_rate_lag = inv_rate[_n-1] if date_lag==1

/**************/
/*Fama-MacBeth*/
/**************/
gen serialcorr_coeff=0
gen serialcorr_const=0
sort date
set more off
   forval t=77 (1) 227 {

        reg inv_rate inv_rate_lag if date==`t'
        replace serialcorr_coeff = _b[inv_rate_lag] if date==`t'
        replace serialcorr_const = _b[_cons]  if date==`t'
   }

/************/
/*Pooled OLS*/
/************/
```

```
reg inv_rate inv_rate_lag
gen serialcorr_coeff_OLS=_b[inv_rate_lag]
gen serialcorr_const_OLS = _b[_cons]

/******************************/
/*Create quarterly time-series*/
/******************************/

save temp, replace

sort  date
drop if date==date[_n-1]

sort date

drop if calendar_year<=1978


replace negspike_rate_inv_rate=0 if negspike_rate_inv_rate==.
replace posspike_rate_inv_rate=0 if posspike_rate_inv_rate==.

keep date serialcorr* obs_inv_rate mean_inv_rate median_inv_rate_pos
median_inv_rate_neg sd_inv_rate median_inv_rate/*
*/   inaction_rate_inv_rate neg_rate_inv_rate  posspike_rate_inv_rate /*
*/   negspike_rate_inv_rate  mean_inv_rate_pos mean_inv_rate_neg mean_bm


/***************/
/*Export output*/
/***************/

format %6.0g mean_inv_rate sd_inv_rate median_inv_rate obs_inv_rate/*
*/     inaction_rate_inv_rate neg_rate_inv_rate posspike_rate_inv_rate
mean_inv_rate_pos mean_inv_rate_neg/*
*/     negspike_rate_inv_rate serialcorr_coeff_OLS serialcorr_const_OLS
median_inv_rate_pos median_inv_rate_neg mean_bm

sum

save investment_rate_stats_bea, replace
```