A Internet appendix: The Utilization Premium

OA.1 Variable description and construction

Asset growth. Asset growth is calculated as the year-on-year annual growth rate of total assets (Compustat Annual item AT) between years $t-1$ and t. This definition of asset growth is drawn from Cooper, Gulen, and Schill (2008).

Book-to-market (BE/ME). A firm's book-to-market ratio is constructed by following Daniel and Titman (2006). Book equity is defined as shareholders' equity minus the value of preferred stock. If available, shareholders' equity is set equal to stockholders' equity (Compustat Annual item SEQ). If stockholders' equity is missing, then common equity (Compustat Annual item CEQ) plus the par value of preferred stock (Compustat Annual item PSTK) is used instead. If neither of the two previous definitions of stockholders' equity can be constructed, then shareholders' equity is the difference between total assets (Compustat Annual item AT) and total liabilities (Compustat Annual item LT). For the value of preferred stock we use the redemption value (Compustat Annual item PSTKRV), the liquidating value (Compustat Annual item PSTKL), or the carrying value (Compustat Annual item PSTK), in that order of preference. We also add the value of deferred taxes and investment tax credits (Compustat Annual item TXDITC) to, and subtract the value of post retirement benefits (Compustat Annual item PRBA) from, the value of book equity if either variable is available. Finally, the book value of equity in the fiscal year ending in calendar year $t-1$ is divided by the market value of common equity from December of year $t-1$.

Capacity. The capacity estimate measures the maximum amount of output that an industry can produce, assuming the sufficient availability of inputs to production and a realistic work schedule. The FRB relies on a variety of sources to determine the capacity of each industry. The primary source of capacity data for manufacturing industries, which make up the bulk of our sample, is currently the Quarterly Survey of Plant Capacity Utilization (QPC). For approximately 20% industries, including a subset of manufacturers, capacity is reported in physical units obtained from government or trade sources, such as the United States Geological Survey. Finally, for a small proportion of industries for which neither of the aforementioned data sources are available, the FRB estimates capacity based on trends through peaks in production. Gilbert, Morin, and Raddock (2000) and Board of Govenors of the Federal Reserve System (2017) provide overviews of how the FRB measures capacity.

Capacity overhang (OVER). We construct a monthly measure of capacity overhang by following the procedure described by Aretz and Pope (2018). In particular we recursively estimate equation (1) of Aretz and Pope (2018) using total assets (Compustat Annual item AT) as our measure of installed capacity.

Debt growth. We measure the growth rate of a firm's debt by calculating the annual percentage change in outstanding total debt, expressed in real terms. We define total real debt as the sum of long-term debt (Compustat Annual item DLTT) and debt in current liabilities (Compustat Annual item DLC), scaled by the value of the consumer price index. When computing this quantity we require firms to have at least $10m of debt outstanding in year $t-1$.

Depreciation rate (BEA implied). To compute the BEA-implied depreciation rate we take the GDP-weighted average of the industry-level depreciation rates associated with equipment and structures.

Depreciation rate (Compustat implied). We compute Compustat-implied industry-level depreciation rates by first constructing firm-level depreciation rates. A firm's depreciation rate is defined as the firm's depreciation expense (Compustat annual item DP) minus the firm's amortization (Compustat annual item AM) scaled by net property, plant, and equipment (Compustat annual item PPENT). We then aggregate these firm-level depreciation rates to the industry-level by computing the value-weighted depreciation rate across all firms assigned to a particular industry.

Equity issuance. Gross equity issuance is defined as the sale of common and preferred stock (Compustat Annual item SSTK) divided by the lagged value of book equity, as per Belo et al. (2018).

Gross profitability (GP). Consistent with Novy Marx (2013), gross profitability is calculated as total revenue (Compustat Annual item REVT) minus the cost of goods sold (Compustat Annual item COGS), divided by total assets (Compustat Annual item AT).

Hiring rate. The hiring rate is computed following Belo et al. (2014). Specifically, the hiring rate in year t is the change in the number of employees (Compustat Annual item EMP) from year $t-1$ to year t, divided by the mean employees over years $t-1$ and t.

Idiosyncratic volatility (IVOL). Idiosyncratic volatility is computed in accordance with Ang et al. (2006). At the end of month t, a firm's idiosyncratic volatility over the past month is obtained by regressing its daily excess returns on the Fama and French (1993) factors, provided there are at least 15 valid daily returns in the month. Idiosyncratic volatility is then defined as the standard deviation of the residuals from the aforementioned regression.

Intangible investment rate (R&D / ME). We follow Lin (2012) and define a firm's intangible investment rate as the firm's R&D expense (Compustat Annual item XRD) divided by the firm's market capitalization.

Inventory growth. The inventory growth rate is defined following Belo and Lin (2012). That is, we compute the annual percentage change in each firm's inventory holdings (Compustat Annual item INVT) after converting the value of inventories to real terms.

Investment rate. We follow Stambaugh and Yuan (2017) and compute the investment rate as the change in gross property, plant, and equipment (Compustat Annual item PPEGT) plus the change in inventory (Compustat Annual item INVT) between years $t-1$ and t, divided by the value of total assets (Compustat Annual item AT) in year $t-1$.

Leverage. We define a firm's leverage ratio as long-term debt (Compustat Annual item DLTT) plus debt in current liabilities (Compustat Annual item DLC) divided by total assets (Compustat Annual item AT).

Market capitalization. A firm's end of month t market capitalization is computed as the firm's end of month t stock price (CRSP Monthly item PRC) multiplied by the firm's number of shares outstanding (CRSP Monthly item SHROUT).

Natural investment rate. Following Belo et al. (2014) the natural rate of investment is computed as capital expenditure (Compustat Annual item CAPX) minus the sales of property, planet, and equipment (Compustat Annual item SPPE) scaled by the average net property, planet, and equipment in years t and $t-1$ (Compustat Annual item PPENT). Missing values of SPPE are set to zero.

Organizational capital (OC). We construct the stock of a firm's organizational capital by following the perpetual inventory inventory method described by Eisfeldt and Papanikolaou (2013). That is, we recursively accumulate a firm's real selling, general and administrative expenses (Compustat Annual item XSGA) over time, and scale the stock of organizational capital by the firm's total assets (Compustat Annual item AT).

Return-on-assets (ROA). Following Imrohoroglu and Tuzel (2014) return on assets (ROA) is computed as net income before extraordinary items (Computat Annual item IB) minus preferred dividends (Compustat Annual item DVP), if available, plus deferred income taxes (Compustat Annual item TXDI), if available, scaled by total assets (Compustat Annual item AT).

TechMark. Recalling equation (OA.3.6), total factor productivity (TFP) is comprised of three distinct components: technology, time-varying markups, and time-varying capacity utilization rates. We isolate the components of TFP related to technology and markups, referred to as "TechMark," as follows. First, we obtain firm-level estimates of the natural logarithm of TFP from Imrohoroglu and Tuzel (2014). We refer to this variable as $\ln(\text{TFP}_{i,t})$. Next, we assign industry-level capacity utilization rates to individual firms by following the matching algorithm described in Section OA.3.3. We take the natural logarithm of these firm firm-level capacity utilization rates, and denote this quantity $\ln(\text{CU}_{i,t})$. Finally, we define the TechMark variable for firm i at time t as $\text{TechMark}_{i,t} = \ln(\text{TFP}_{i,t}) - \ln(\text{CU}_{i,t})$.

Total factor productivity (TFP). The firm-level estimates of TFP are drawn from Imrohoroglu and Tuzel (2014).

OA.2 Capacity utilization data and summary statistics

The public report on industrial capacity utilization covers 57 industries. These industries are defined at different levels of aggregation ranging from two- to six-digit North American Industry Classification System (NAICS) codes. Specifically, 12 of the industries are crude aggregates that span *multiple* two-digit NAICS codes. For example, one of these 12 aggregates includes the average capacity utilization rate of all manufacturers in the United States. We remove these 12 crude aggregates from our benchmark sample for two reasons. First, these aggregates do not provide new information as they are spanned by more granularly defined sub-industries that are also included in the sample. Second, these aggregates represents a considerable proportion of total market value and would consequently dominate the returns of the value-weighted portfolios we form in Section 1. Removing these 12 crude aggregates leaves us with a benchmark cross-section of 45 industries that features a mix of durable manufacturers, nondurable manufacturers, and miners and utilities.[25]

As the 45 industries included in the benchmark sample are defined from the relatively coarse two-digit NAICS code level to the most granular six-digit NAICS code level, the benchmark cross-section includes a number of overlapping industries.[26] For instance, the capacity utilization rate of food manufacturers is included in the utilization rate of two industries reported by the FRB: "Food," as well as "Food, beverage, and tobacco." Since removing overlapping industries from our benchmark sample would significantly reduce the number of cross-sectional assets, and thus, make certain asset pricing tests, such as portfolio double sorts, infeasible, we deal with this overlap in two ways. First, in robustness section OA.3.1 we remove the industries that overlap with others and conduct our baseline empirical tests in a sub-sample of 24 distinct non-overlapping industries, each of which corresponds to a unique three-digit NAICS code. Following the example above, this set of distinct industries includes both "Food"' and "Beverage and Tobacco" manufacturers, but excludes the composite index that covers both groups of manufacturers. Table OA.3.3 shows that the utilization premium is significant even within this narrower sample. Second, we verify that our results are not driven by any particular industry that dominates the sample (see Tables OA.3.13).

Monthly utilization data for 32 industries are available in January 1967, and data for an ad-

[25] A list of these 45 industries, along with each industry's sector, is provided in Table OA.2.1 of the Online Appendix.

[26] In particular, our final sample consists of one sector defined at the two-digit NAICS level, 27 subsectors defined at the three-digit NAICS level, 13 industry groups defined at the four-digit NAICS level, two industries defined at the five-digit NAICS level, and two U.S. industries defined at the six-digit NAICS level. In Section OA.3 we ensure that our results are robust to this heterogeneity in classification levels.

ditional 25 industries becomes available in January 1972.[27] The utilization data we collect ends in December 2015, when we commenced the empirical analysis of the paper. We verify in Section OA.3.1 that our results hold when we only consider the most recent half of the sample period.

OA.2.1 Summary statistics

Below, we describe the properties of the aggregate capacity utilization rate and report summary statistics related to the cross-section of industry-level utilization rates.

Figure OA.2.1 shows the annual growth rate of aggregate capacity utilization over the sample period. The figure shows that utilization fluctuates significantly over time and that the growth rate of aggregate utilization is procylical. The aggregate utilization rate drops during recessions, particularly during the Great Recession. The growth rate of aggregate utilization tends to slightly lead the business cycle, and has often served as an early warning for recessions. In five out of the seven recessions during our sample period the growth rate of utilization begins to drop prior to the start of the recession. The growth rate of capacity utilization increases during the technological revolution of the late 1990's, the housing bubble, and the recovery from the Great Recession.

As illustrated by equation (1), the capacity utilization rate is a combination of both industrial production and capacity. The former variable is studied extensively in the macroeconomic and finance literature, and features prominently in the context of asset pricing. For instance, Cooper et al. (2008) document a premium for firms with lower total asset growth. The growth rate of assets is directly linked to firms' output, and is consequently captured by the FRB's measure of industrial production. To establish the empirical novelty in examining capacity utilization, we examine the extent to which utilization fluctuates independently of industrial production using:

$$\Delta CU_t = \beta_0 + \beta_1 \Delta IP_t + \varepsilon_t. \tag{OA.2.1}$$

Here, ΔCU_t (ΔIP_t) is annual growth rate of aggregate capacity utilization (industrial production), and the residual ε_t captures the component of capacity utilization that is orthogonal to industrial production. Figure OA.2.1 also displays this orthogonal component over the sample period. The dynamics of this orthogonal component do not appear to reflect the dynamics of a white noise process. ε_t is smoother than utilization growth, and changes in ε_t are largely procylical. Similar to utilization growth, ε_t tends to drop during NBER recessions. In some instances the orthogonal component also deviates significantly from capacity utilization growth. For example, during the technology boom of mid-1990's, the orthogonal component declines whereas capacity utilization increases. The orthogonal component drops due to an acceleration in the growth of capacity that was likely facilitated by the technological advancements of the era (Bansak, Morin and Starr 2007).

Table OA.2.2 reports summary statistics for the capacity utilization rates of each sector in our benchmark sample. The average rate of aggregate capacity utilization rate is 79.91%. This figure for the U.S. is close to the average capacity utilization rates of 81.17%, 84.69%, and 82.49% for the Euro Zone, China, and Israel, respectively.[28]

The majority of the utilization data in the sample pertains to the manufacturing sector, with an almost even split between durable and nondurable manufacturing industries. The mean annual utilization rate is 77.39% (80.09%) for durable (nondurable) manufacturers. Each of these rates is statistically indistinguishable from the average rate of capacity utilization across all industries in the sample. The fact that the average utilization rate of the manufacturing sector, and of the durable and nondurable manufacturing subsectors, is not statistically different from the U.S. average alleviates the concern that our results are driven by ex-ante heterogeneity between sectors.

[27] There are only 11 monthly time-series reported between January 1948 to December 1966. As eleven industries is a very small cross-section, we do not consider the pre-1967 period in our benchmark sample.

[28] See `https://www.dallasfed.org/institute/oecd` for data on these utilization rates recorded by the Organization for Economic Cooperation and Development (OECD) and reported by the Federal Reserve Bank of Dallas.

Among mining industries and utilities the average utilization rate is 84.13%. This average rate is slightly higher than, and statistically different from, the average rate across all industries. Due to this difference in average capacity utilization rates we verify that our empirical results are robust to excluding mining industries and utilities from our sample. We also verify that our results still hold when we conduct tests using the growth rate of utilization that eliminates differences in levels by construction. The results of both of these tests are reported in Section OA.3.

Table OA.2.2 also reports the volatility and autocorrelation of utilization for the different sectors in our sample. The volatility of the capacity utilization rate is comparable across sectors and ranges from 6.67% per annum for mining to 8.29% per annum for durables. The autocorrelation ranges from 0.52 to 0.61, with an all-industry mean of 0.58. These statistics affirm the notion that the level of utilization follows similar dynamics regardless of sector.

Overall, Table OA.2.2 shows that the *unconditional* average rate of utilization is only slightly different between sectors. In particular, most differences between the average utilization rate of a sector and the average aggregate utilization rate are statistically indistinguishable from zero. In contrast to these unconditional differences, the asset pricing tests we conduct rely on *conditional* variation in utilization rates. Thus, our tests exploit the fact that the relative ranking of industries in terms of utilization changes over time. Untabulated results show that if we assume that utilization rates are constant over time, and try to utilize the small unconditional differences in the average rate of utilization between industries to perform the empirical tests, our results cease to hold.

Finally, Table OA.2.3 shows the correlation between capacity utilization and other industry-level production-based characteristics for the average industry in our sample. The characteristics considered include book-to-market, TFP, the hiring rate, sales-to-assets, and the investment rate. While utilization has a positive correlation with productivity, hiring, sales, and investment, these average correlation are fairly low. For example, the average correlation between the investment rate and utilization is only 0.16. This suggests that varying utilization constitutes a separate degree of freedom for managers to smooth dividends, and that any interaction between utilization and expected returns is likely to be independent of the well-established book-to-market and investment rate effects on risk premia.

Table OA.2.1: **Sample composition and industry specification**

Industry name	Sector	Overlap	Start year
Nonmetallic mineral product	D	No	1967
Primary metal	D	No	1967
Fabricated metal product	D	No	1967
Machinery	D	No	1967
Transportation equipment	D	No	1967
Motor vehicles and parts	D	Yes	1967
Aerospace and miscellaneous transportation eq.	D	Yes	1967
Furniture and related product	D	No	1967
Computers, communications eq., and semiconductors	D	Yes	1967
Wood product	D	No	1972
Iron and steel products	D	Yes	1972
Computer and electronic product	D	No	1972
Computer and peripheral equipment	D	Yes	1972
Communications equipment	D	Yes	1972
Semiconductor and other electronic component	D	Yes	1972

Continued on the next page...

Table OA.2.1 – Continued from the previous page

Industry name	Sector	Overlap	Start year
Electrical equipment, appliance, and component	D	No	1972
Automobile and light duty motor vehicle	D	Yes	1972
Miscellaneous	D	No	1972
Food, beverage, and tobacco	ND	Yes	1967
Leather and allied product	ND	No	1967
Paper	ND	No	1967
Petroleum and coal products	ND	No	1967
Chemical	ND	No	1967
Plastics and rubber products	ND	No	1967
Food	ND	No	1972
Beverage and tobacco product	ND	No	1972
Textile mills	ND	No	1972
Textiles and products	ND	Yes	1972
Textile product mills	ND	No	1972
Apparel	ND	No	1972
Apparel and leather goods	ND	Yes	1972
Printing and related support activities	ND	No	1972
Synthetic rubber	ND	Yes	1972
Plastics material and resin	ND	Yes	1972
Artificial and synthetic fibers and filaments	ND	Yes	1972
Mining	MU	No	1967
Metal ore mining	MU	Yes	1967
Nonmetallic mineral mining and quarrying	MU	Yes	1967
Electric power generation, transmission, and distribution	MU	Yes	1967
Electric and gas utilities	MU	Yes	1967
Natural gas distribution	MU	Yes	1967
Coal mining	MU	Yes	1967
Oil and gas extraction	MU	No	1972
Mining (except oil and gas)	MU	No	1972
Support activities for mining	MU	No	1972

The table lists the industries for which capacity utilization data is available at FRED. This set of industries comprises our benchmark sample. For each industry, the table specifies its name, its sector (D denotes the durable sector, ND denotes the nondurable sector, and MU refers to the mining and utilities sector), whether certain industry constituents overlap with other industries in the sample, and the first year in which the industry appears in the sample. All data ends at December 2015.

Table OA.2.2: **Summary statistics of the capacity utilization rate by sector**

Sector	N	Mean	t(Sector-All)	SD	AC(1)
All industries	45	79.91	–	7.38	0.58
Manufacturing	35	78.70	(-0.75)	7.58	0.57
Durable	18	77.39	(-1.58)	8.29	0.52
Nondurable	17	80.09	(0.12)	6.83	0.61
Mining and utilities	10	84.13	(2.35)	6.67	0.60

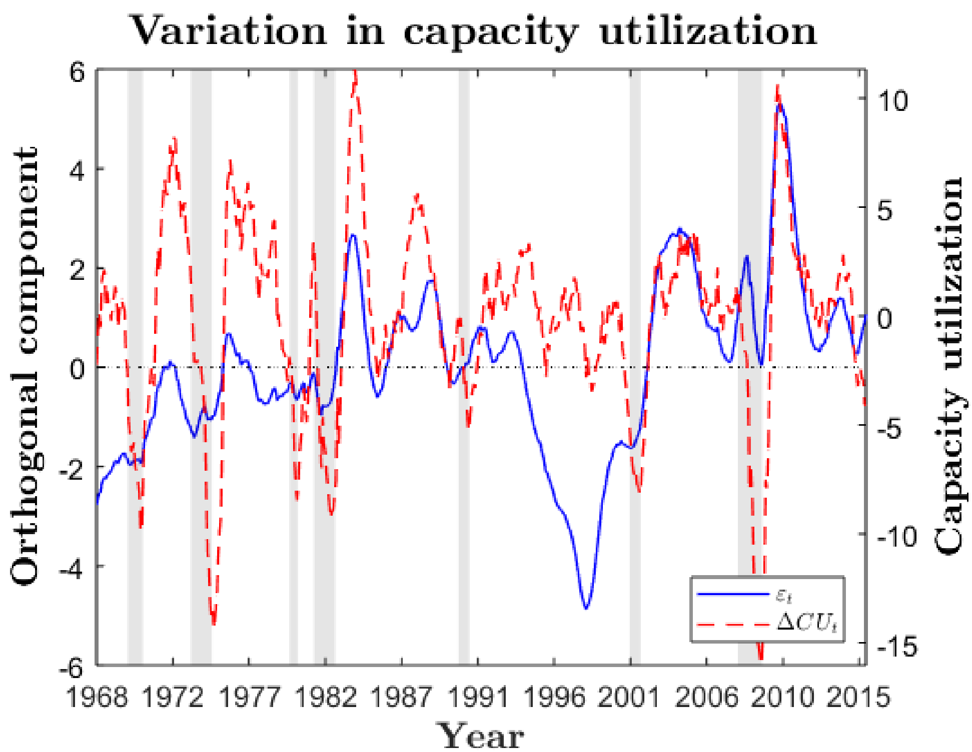
The table reports the mean, standard deviation (SD), and autocorrelation (AC(1)) of the annual capacity utilization rates for each sector in the sample. N represents the number of industries within each sector for which a time-series of capacity utilization data available on FRED. The column t(Sector-All) shows the Newey and West (1987) t-statistic, in parentheses, for the difference in the average capacity utilization rate between the sector denoted in the leftmost column and the average capacity utilization rate across all industries (the top row). The data spans the period 1967 to 2015.

Table OA.2.3: **Average industry-level correlation between production-based characteristics**

	CU	BE / ME	TFP	Hire Rate	Sales / Assets	I / K
CU	1.00	-0.15	0.11	0.13	0.15	0.16
BE / ME		1.00	0.04	-0.26	0.26	-0.00
TFP			1.00	0.23	0.29	0.48
Hire Rate				1.00	0.05	0.53
Sales / Assets					1.00	0.27
I / K						1.00

The table shows the correlation between pairs of industry-level characteristics, averaged over all industries in the sample. The characteristics are the capacity utilization rate (CU), the book-to-market ratio (BE/ME), total factor productivity (TFP), the hiring rate (Hire Rate), the ratio of sales-to assets (Sales / Assets), and the investment rate (I/K). At the end of each June from 1967 to 2015, each industry-level characteristic is constructed as the simple average of the characteristic of interest over all firms that belong to the industry at the point in time. For each industry, we then compute the correlation between industry-level characteristic X and industry-level characteristic Y over the sample period, and report the average value of this correlation across all industries in the sample. The data is annual and runs from 1967 to 2015. Additional details on the construction of each variable are provided in Section OA.1 of the Online Appendix.

Figure OA.2.1: **Capacity utilization: orthogonality from industrial production**



The figure shows the time-series of aggregate capacity utilization growth (dashed red line), as well as the component of capacity utilization growth that is orthogonal to industrial production growth (solid blue line). The orthogonal component of capacity utilization component is obtained from the residuals of the following projection: $\Delta CU_t = \beta_0 + \beta_1 \Delta IP_t + \varepsilon_t$, where CU is aggregate capacity utilization, IP is industrial production, and ε_t is the component of capacity utilization that is orthogonal to industrial production. The horizontal axis shows years and grey shaded regions denote NBER recessions. The right (left) vertical axis represents changes in (the orthogonal component of) capacity utilization. All growth rates are annual, and the sample period ranges from July 1967 to July 2015.

OA.3 Additional empirical results

OA.3.1 Methodological variations in portfolio formation

In this section we show that the capacity utilization spread is also robust to several implementation choices related to the portfolio formation procedure described in Section 1.2.

Variation in breakpoints. In the benchmark analysis we use the 10^{th} and 90^{th} percentiles of the cross-sectional distribution of capacity utilization rates as breakpoints for the low and high utilization portfolios. Here, we modify these breakpoints and sort industries into portfolios using either (i) quintiles or (ii) the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of capacity utilization rates. These choices of breakpoints either double or triple the number of industries in each of the low and high capacity utilization portfolios. The value- and equal-weighted returns of the quintile portfolios are reported in Panel A of Table OA.3.1, while the returns obtained by using the 30^{th} and 70^{th} percentiles of the distribution of utilization rates are reported in Panel B. Despite using coarser breakpoints, the value-weighted utilization spread in each panel is close to 5% per annum and statistically significant. This spread is less than 1% smaller in magnitude than the benchmark. Portfolio returns also tend to decrease as the average utilization rate of each portfolio increases, with the equal-weighted returns showing strict monotonicity in utilization.

Variation in the sample period. While our sample period spans July 1967 to December 2015, we consider the impact of breaking the sample in half and examining the utilization spread in the most recent subsample that starts in July 1991. This subsample analysis is reported in Table OA.3.2 and shows that the magnitude of the spread is larger in the recent subsample than it is over the entire sample period. While the value-weighted spread has a mean return of 5.67% per annum between July 1967 and December 2015, its mean return between July 1991 and December 2015 is 9.09% per annum. As the second half of the sample is populated by two major recessions, the recession of the early 2000s and the Great Recession, this result also shows that the utilization spread is largely countercyclical.

Variation in the sample of industries. Our benchmark results are based on a cross-section of 45 industries. However, as explained in Section OA.2, some of these industries are comprised of firms that belong to multiple industries in the sample. To ensure that our results are not driven by this feature of the data, we repeat our baseline analysis using a subsample of industries whose constituent firms are distinct from one another. These 24 no-overlap industries are listed in Table OA.2.1 and the results of repeating our benchmark portfolio sorts in this subsample of industries are shown in Table OA.3.3. The value-weighted utilization spread is close to 6.65% per annum in this subsample, and even larger than our benchmark spread of 5.7% per annum. The value-weighted (equal-weighted) utilization spread is still statistically significant at the 5% (10%) level.

Relatedly, Table OA.3.13 exclude various sectors of the economy (e.g., durables or mining and utilities) and shows that the utilization premium remains economically and statistically significant. Likewise, Table OA.3.14 shows the premium is robust to sorting on the growth rate of utilization thereby eliminating industry fixed effect in utilization's level.

Collectively, the evidence clearly shows, to the best of our ability and available data, that the utilization premium is not sensitive to the specific choice of industries comprising the sample.

Quarterly portfolio sorting. Table OA.3.4 shows that forming portfolios on a quarterly, rather than an annual basis also leads to an economically large and statistically significant capacity utilization spread. Here, we form portfolios at the end of each March, June, September, and December using the lagged capacity utilization rate, and hold each portfolio for three months before rebalancing. The value-weighted (equal-weighted) capacity utilization spread is 6.45% (5.23%) per annum. Both are statistically significant at the 5% level.

Importance of conditional sorting. In untabulated results we demonstrate the importance of the *conditional* portfolio sorting procedure described in Section 1.2. Specifically, we consider an alternative procedure in which each industry is permanently assigned to the first portfolio it is sorted into. This *unconditional* portfolio sort leads to a capacity utilization spread that is both economically and statistically insignificant. This result highlights that there is a significant degree of conditional variation in industry-level capacity utilization rates, and that this variation is important for generating the capacity utilization spread.

Table OA.3.1: **Capacity utilization spread: alternative portfolio breakpoints**

Panel A: Quintile portfolios				
	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	13.31	18.97	10.23	20.41
2	11.88	17.85	9.74	18.92
Medium	8.78	18.60	7.68	18.57
4	9.25	18.09	7.51	17.52
High (H)	8.44	17.79	5.87	18.64
Spread	4.87	14.07	4.35	11.80
(L-H)	(2.35)		(2.57)	

Panel B: 30^{th} and 70^{th} percentile breakpoints				
	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	12.62	17.86	9.66	19.72
Medium	10.09	18.43	8.51	17.85
High (H)	7.75	17.18	6.22	17.86
Spread	4.86	13.44	3.44	10.43
(L-H)	(2.47)		(2.26)	

The table reports annual returns of portfolios sorted on the basis of capacity utilization, as well as the spread between the low (L) and high (H) capacity utilization portfolios. The construction of the portfolios is identical to the benchmark analysis, except for the portfolio breakpoints used. In Panel A, we sort the cross-section of industries into five portfolios based on quintile breakpoints. In Panel B, we sort the cross-section of industries into three portfolios based on the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of capacity utilization rates. Mean refers to the average annual return and SD denotes the standard deviation of annual returns. Parentheses report t-statistics computed using Newey and West (1987) standard errors. The portfolios are formed at the end of each June from 1991 to 2015 and are rebalanced annually.

Table OA.3.2: **Capacity utilization spread: results based on the recent subsample**

	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	15.29	22.74	11.72	21.81
Medium	10.04	16.95	8.24	16.86
High (H)	6.20	20.49	4.39	20.40
Spread	9.09	20.66	7.34	17.53
(L-H)	(2.39)		(2.22)	

The table reports annual returns of portfolios sorted on the basis of capacity utilization, as well as the spread between the low (L) and high (H) capacity utilization portfolios. The construction of the portfolios is identical to the benchmark analysis, except that the sample period only includes the recent period from July 1991 to December 2015. Mean refers to the average annual return and SD denotes the standard deviation of annual returns. Parentheses report t-statistics computed using Newey and West (1987) standard errors. The portfolios are formed at the end of each June from 1991 to 2015 and are rebalanced annually.

Table OA.3.3: **Capacity utilization spread: results based on non-overlapping industries**

	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	15.43	20.56	11.52	20.25
Medium	10.36	16.70	7.94	18.17
High (H)	8.77	20.24	7.08	20.30
Spread	6.65	18.32	4.45	17.40
(L-H)	(2.51)		(1.77)	

The table reports annual returns of five portfolios sorted on the basis of capacity utilization, as well as the spread between the low (L) and high (H) capacity utilization portfolios. The construction of the portfolios is identical to the benchmark analysis, except that the sample of industries is restricted to those industries whose constituents do not belong to multiple industries in the sample (see Table OA.2.1, Column 3, for the list of these non-overlapping industries). Mean refers to the average annual return, SD denotes the standard deviation of annual returns. Parentheses report t-statistics computed using Newey and West (1987) standard errors. The portfolios are formed at the end of each June from 1967 to 2015 and are rebalanced annually, with portfolio returns spanning July 1967 to December 2015.

Table OA.3.4: **Capacity utilization spread: quarterly portfolio formation**

Portfolio	Value-weighted		Equal-weighted	
	Mean	SD	Mean	SD
Low (L)	13.66	21.91	10.33	21.52
Medium	10.69	16.65	8.23	17.55
High (H)	7.20	20.48	5.10	21.00
Spread	6.45	18.57	5.23	16.55
(L-H)	(2.41)		(2.21)	

The table reports annual returns of three portfolios sorted on the basis of capacity utilization, as well as the spread between the low (L) and high (H) capacity utilization portfolios. Here, portfolios are formed at the end of each March, June, September, and December on the basis of each industry's lagged capacity utilization rate. Each of these portfolios is then held for three months before rebalancing. Mean refers to the average annual return, SD denotes the standard deviation of annual returns. Parentheses report t-statistics computed using Newey and West (1987) standard errors. The portfolios are formed at the end of each quarter from 1967 to 2015, with portfolio returns spanning April 1967 to December 2015.

OA.3.2 Conditional CAPM and time-varying betas to aggregate productivity

In this section we show that time-varying exposures to the market (aggregate productivity) fully absorb the utilization premium.

Conditional CAPM. We estimate the conditional CAPM to show that time variation in risk exposures are important to reconcile the utilization premium. Here, we estimate the conditional CAPM following the procedure outlined by Lewellen and Nagel (2006). That is, *within* each year of the sample period, we estimate an unconditional CAPM regression that projects each portfolio's excess returns on excess market returns. We then compute the time-series average of the corresponding CAPM alphas and betas. The results, reported in Table OA.3.5, confirm that a single-factor model is able to explain the utilization premium. Accounting for time-variation in CAPM betas *reduces* the economic magnitude of the average CAPM alphas compared to the unconditional specification and renders the CAPM alpha statistically insignificant.

Table OA.3.5: **Capacity utilization spread: conditional CAPM alphas and betas**

Portfolio	CAPM alpha α	$t(\alpha)$	CAPM beta β	$t(\beta)$
Low (L)	0.81	(0.33)	1.15	(9.66)
Medium	-1.05	(-0.48)	1.06	(11.70)
High (H)	-3.99	(-0.57)	0.92	(10.85)
Spread (L-H)	3.43	(1.34)	0.18	(2.35)

The reports the alphas and betas obtained from estimating the conditional CAPM following the methodology proposed by Lewellen and Nagel (2006). We implement this analysis to obtain the conditional CAPM alphas and betas of each value-weighted utilization-sorted portfolio in two steps as follows. First, within each year of the sample period, we estimate an unconditional CAPM regression and record the alpha and beta associated with each portfolio. Second, we compute the time-series average of the estimated alphas and betas, and report these time-series averages in the table. We annualize each alpha by multiplying the time-series average of the monthly alphas obtained in the second step of the procedure by 12. Parentheses report Newey and West (1987) t-statistics, and the sample is from July 1967 to December 2015.

Non-linear model exposures. We complement the former evidence using other proxies of aggregate productivity and a non-linear model specification. By construction, the exposure of each portfolio to aggregate productivity is varies over time and with the business cycle (if the loading on the non-linear term is non-zero). We augment equation (2) with a quadratic aggregate productivity term and estimate the following regression:

$$Ret^e_{i,t} = \beta_{0,i} + \beta_{1,i}\text{Agg-Prod}_t + \beta_{2,i}\text{Agg-Prod}^2_t + \varepsilon_{i,t}. \tag{OA.3.1}$$

Here, $Ret^e_{i,t}$ is the value-weighted excess return of the portfolio of interest, Agg-Prod_t is a proxy for aggregate productivity, and $\beta_{1,i}$ ($\beta_{2,i}$) captures the exposure of portfolio i to the linear (quadratic) effect of aggregate productivity.

We implement this analysis by considering three proxies for aggregate productivity: (i) the market return, (ii) utilization-adjusted TFP growth from Fernald (2012), and (iii) labor productivity from the BLS. We also combine the slope coefficients on the linear and quadratic terms to form the (total) productivity beta of portfolio i (denoted by $\beta_{i,prod}$) as:

$$\beta_{i,prod} = \mathrm{E}\left[\frac{\partial Ret^e_i}{\partial \text{Agg-Prod}}\right] = \beta_{1,i} + 2\beta_{2,i}\mathrm{E}\left[\text{Agg-Prod}\right], \tag{OA.3.2}$$

and compute the standard errors associated with $\beta_{i,prod}$ using the Delta method. We report the results of this analysis in Table OA.3.6.

Table OA.3.6: **Exposure of CU-sorted portfolios to aggregate productivity proxies**

Portfolio	Market returns		Util.-adjusted TFP		Labor productivity	
	β_{prod}	$t(\beta_{prod})$	β_{prod}	$t(\beta_{prod})$	β_{prod}	$t(\beta_{prod})$
Low (L)	1.37	(9.12)	1.18	(2.78)	0.88	(1.81)
Medium	1.25	(11.12)	0.76	(2.27)	0.54	(1.33)
High (H)	1.07	(6.53)	0.78	(2.02)	0.54	(1.18)
Spread (L-H)	0.30	(3.53)	0.40	(1.91)	0.34	(1.94)
Intercept	1.27	(0.41)	4.46	(1.33)	3.08	(0.99)

The table reports the exposures of portfolios sorted on capacity utilization to three different aggregate productivity proxies. The regression we estimate is: $Ret^e_{i,t} = \beta_{0,i} + \beta_{1,i}\text{Agg-Prod}_t + \beta_{2,i}\text{Agg-Prod}^2_t + \varepsilon_{i,t}$, where $Ret^e_{i,t}$ is the value-weighted excess return of portfolio i, Agg-Proxy is a proxy of aggregate productivity, and β_1 (β_2) is the sensitivity of portfolio i's excess return to the linear (quadratic) aggregate productivity term. We combine the linear and quadratic sensitivities to form the productivity beta (β_{prod}) following equation (OA.3.2), and report β_{prod} in the table. Here, Agg-Proxy is either (i) excess market returns, (ii) utilization-adjusted TFP growth from Fernald (2012), or (iii) labor productivity growth from the BLS. Monthly returns are aggregated to the quarterly frequency so that each regression is estimated using quarterly data. Newey and West (1987) t-statistics associated with each exposure are reported in parentheses, with the standard errors associated with β_{prod} computed using the delta method. "Intercept" refers to the annualized value of β_0 (obtained by multiplying β_0 by four) from projecting the utilization spread on each productivity proxy. Finally, the sample spans July 1967 to December 2015.

Comparing Table OA.3.6, which features a non-linear relation to aggregate productivity, to Table 2, which features only a linear relation, delivers three key takeaways. First, accounting for the non-linear term, the low utilization portfolio's exposure to aggregate productivity remains significantly larger than the high utilization portfolio's exposure to productivity. Second, with non-linear exposures to aggregate productivity, the differences in productivity betas (L-H) in Table OA.3.6 are at least as large, if not larger, than those reported in Table 2. Third, when we measure aggregate productivity using excess market returns, the economic magnitude of the intercept falls to 1.27% per annum and becomes statistically insignificant with a t-statistic of 0.41. Likewise, the intercept from projecting the utilization premium on utilization-adjusted TFP growth from Fernald (2012) becomes statistically insignificant.

Relation between conditional exposure and utilization. We show that capacity utilization correlates with industries' *conditional* exposure to aggregate productivity (CAPM betas). We establish this link in two steps. First, for each industry j and each month t, we estimate the following time-series regression using the past 60 months of stock return data

$$R^e_{j,t-60\to t} = \alpha_{j,t} + \beta_{j,t} MKTRF_{t-60\to t} + \varepsilon_{j,t-60\to t}, \qquad \text{(OA.3.3)}$$

where, $R^e_{j,t-60\to t}$ denotes the excess stock return of industry j, $MKTRF$ represents the market excess return, and $\beta_{j,t}$ is the conditional exposure of industry j to aggregate productivity at time t. Estimating this time-series regression for each industry and each month of our sample delivers a panel containing each industry's conditional risk exposure between July 1972 and December 2015.

Next, we estimate the following panel regression that projects the risk exposures onto capacity utilization, and other monthly industry-level variables, observable at time t

$$\beta_{j,t} = \alpha_j + \delta_t + \beta_{CU} CU_{j,t} + \boldsymbol{\beta X'_{j,t}} + \varepsilon_{j,t}. \qquad \text{(OA.3.4)}$$

Here, $CU_{j,t}$ denotes the utilization rate of industry j in month t. β_{CU} captures the degree to which capacity utilization is, on average, informative about conditional risk exposures. $\boldsymbol{X}_{j,t}$ is a vector of other industry-level controls that vary at the monthly frequency, including the size, idiosyncratic volatility, and momentum of industry j. α_j is an industry fixed effect, and δ_t is a time fixed effect that absorbs changes in conditional risk exposures common to all industries. We standardize each independent variable (to aid the interpretation of the marginal effects across specifications) and compute t-statistics using standard errors clustered by both industry and time.

Panel A of Table OA.3.7 shows that without controlling for other industry-level characteristics, the capacity utilization rate is highly informative about an industry's conditional CAPM beta. For instance, Column 1 shows that a one standard deviation increase in capacity utilization is associated with a -0.09 *decrease* in an industry's conditional market beta. Controlling for both time and industry fixed effects in Column 3 yields similar results. Moreover, Panel B indicates that the negative association between capacity utilization rates and conditional market betas remains significant when additional control variables are added to each specification.

Table OA.3.7: **Sensitivity of conditional CAPM betas to capacity utilization**

	Panel A: No controls			Panel B: Controls		
	(1)	(2)	(3)	(4)	(5)	(6)
CU	-0.09	-0.09	-0.05	-0.09	-0.07	-0.05
	(-3.87)	(-3.09)	(-2.33)	(-3.86)	(-2.76)	(-2.38)
Size				-0.35	-1.00	0.20
				(-1.22)	(-1.70)	(1.66)
IVOL				0.49	0.99	0.02
				(1.93)	(1.62)	(0.22)
Momentum				-0.07	0.07	-0.15
				(-0.65)	(0.29)	(-4.81)
Time FE	No	Yes	Yes	No	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
R2	0.05	0.12	0.55	0.07	0.17	0.55
Obs.	23363	23363	23363	23363	23363	23363

The reports the relation between conditional CAPM betas, capacity utilization rates, and other industry-level characteristics. We implement this analysis in two steps. In the first step, we obtain the conditional CAPM beta of each industry j in each month t by estimating the time-series regression represented by equation (OA.3.3) using the past 60 months of stock return data. In the second step, we project the conditional CAPM beta of each industry j at each time t on a set of industry-level characteristics including the industry's (i) capacity utilization rate, (ii) size, (iii) idiosyncratic return volatility, and (iv) return momentum. The specification of the panel regression we use in this second step is represented by equation (OA.3.4). The regressions in Panel A feature no additional control variables beyond the capacity utilization rate, whereas the regressions in Panel B include all control variables. We scale each independent variable by its unconditional standard deviation prior to estimating and reporting each slope coefficient. Parentheses represent t-statistics computed using standard errors clustered by industry and month. Finally, the sample period is from July 1967 to December 2015.

Non-linear model: robustness using quintiles. Table OA.3.8 repeats the non-linear model analysis using five, rather than three, portfolios sorted on utilization rates. We use the market portfolio as a proxy for aggregate productivity. Even with quintile portfolios: (i) productivity exposures tend to decline with utilization, and (ii) the linear and non-linear productivity exposures of the low utilization portfolio are significantly higher than those of the high utilization portfolio.

Panel B also shows that the non-linear alpha is only 0.67% per annum and statistically insignificant.

Table OA.3.8: **Exposures of quintile portfolios to aggregate productivity**

Portfolio	Panel A: Linear β_{prod}	Panel A: Linear $t(\beta_{prod})$	Panel B: Non-linear β_{prod}	Panel B: Non-linear $t(\beta_{prod})$
Low (L)	1.29	(11.76)	1.33	(10.70)
2	1.27	(11.68)	1.30	(10.68)
Medium	1.28	(10.47)	1.28	(9.53)
4	1.16	(10.96)	1.19	(10.38)
High (H)	1.08	(8.72)	1.08	(7.38)
Spread (L-H)	0.21	(2.74)	0.25	(3.00)
Intercept (L-H)	3.82	(1.79)	0.67	(0.26)

The table reports the exposures of portfolios sorted on capacity utilization to aggregate productivity. The regression we estimate is: $Ret^e_{i,t} = \beta_{0,i} + \beta_{1,i}\text{Agg-Prod}_t + \beta_{2,i}\text{Agg-Prod}^2_t + \varepsilon_{i,t}$, where $Ret^e_{i,t}$ is the value-weighted excess return of portfolio i, Agg-Proxy is the aggregate productivity proxy, and β_1 (β_2) is the sensitivity of portfolio i's excess return to the linear (quadratic) aggregate productivity term. We combine the linear and quadratic sensitivities to form the productivity beta (β_{prod}) following equation (OA.3.2), and report β_{prod} in the table. Here, Agg-Proxy is measured using excess market returns. Monthly returns are aggregated to the quarterly frequency so that each regression is estimated using quarterly data. Newey and West (1987) t-statistics associated with each exposure are reported in parentheses, with the standard errors associated with β_{prod} computed using the delta method. "Intercept" refers to the annualized value of β_0 (obtained by multiplying β_0 by four) from projecting the utilization spread on the productivity proxy. Finally, the sample spans July 1967 to December 2015.

OA.3.3 Assigning industry-level utilization rates to CRSP/Compustat firms

When applicable, we assign each firm a capacity utilization rate that corresponds to the utilization rate of the industry to which the firm belongs. However, recall from Section OA.2 that our sample is comprised of industries that are defined with different degrees of granularity. This means that some firms may be matched to more than one industry. We execute the following matching algorithm to ensure that each firm is matched to the most granularity defined industry to which it belongs. We start by assigning utilization rates to all firms that belong to a six-digit NAICS code industry for which utilization data is available. We then consider the five-digit NAICS code industries and identify the constituents of these industries that were not previously assigned a utilization rate. These firms are then assigned a utilization rate corresponding to a five-digit NAICS code industry. This procedure then continues to the four-, three-, and two-digit NAICS code industries, in that order. If a previously unmatched firm belongs to two or more N-digit NAICS code industries, then we assign the firm the utilization rate of its "parent" (N-1)-digit NAICS code industry. Any firms unmatched at the end of this procedure are removed from the sample.

OA.3.4 Independence from value, investment, and organizational capital

Independence using Fama-Macbeth regressions. Table 5 reports the results of the Fama and MacBeth (1973) regression analysis outlined by equation (3). Columns (1) to (6) show that the coefficient on utilization is negative and significant at the 5% level when we include another investment-related characteristic in the regressions, including TFP, hiring rate, investment rate, capital overhang, and book-to-market ratios. The signs of the average slope coefficients of the other characteristics are consistent with the spreads associated with each characteristic of interest.[29]

Columns (7) to (9) control for several investment-related characteristics jointly. In all cases, utilization's predictive power for returns remains significant at the 5% level. Lastly, Column (10) considers all of the aforementioned characteristics in addition to size, organizational capital, and past returns.[30] Compared to Columns (1) to (9), the slope coefficient on utilization in Column (10) is largely similar. The slope coefficients on the other characteristics also remain similar, with the exception of the loading on TFP that flips sign from negative to positive. However, this change in sign does not compromise the validity of the TFP spread as a number of the investment-related characteristics included in this specification are relatively highly correlated. In all, the regression evidence shows that the relation between utilization and stock returns is materially distinct from to the known relations between returns and each of TFP, hiring, investment, and B/M.

Independence using portfolio double sorts. We corroborate the results of the Fama and MacBeth (1973) analysis by conducting conditional portfolio double sorts. That is, we conditionally sort the sample of industries into portfolios along two dimensions. The first dimension corresponds to either (i) book-to-market ratios, (ii) physical investment rates, or (iii) ratios of organizational capital to assets, while the second dimension reflects the capacity utilization rate.

The motivation behind this exercise is to control for one variable of interest (e.g., B/M) and show that controlling for the first dimension, the utilization premium remains significant. We focus on book-to-market ratios and investment rates as these are the only two characteristics in Panel B of Table 3 that are significantly different between the extreme utilization portfolios and also command a risk premium that is aligned with the utilization spread. While the difference in organizational capital is insignificant across the utilization portfolios, we conduct this double sort to rule out the possibility that low utilization is mechanically driven by a higher reliance on intangible capital.

The double sorts are implemented as follows. At the end of each June from 1967 to 2015 we first sort the cross-section of firms into three portfolios based on either their (i) book-to-market ratios, (ii) investment rates, or (iii) organizational capital-to-asset ratios. We use the 30^{th} and 70^{th} percentiles of the firm-level cross-sectional distribution of each characteristic to assign each firm to one of three portfolios. Next, within each of these three characteristic-sorted portfolios, we further sort firms into three additional portfolios on the basis of capacity utilization. We also use the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of capacity utilization rates in March of the same year to determine portfolio membership in this second step. This process produces nine portfolios that are each held from the beginning of July in year t to the end of June in year $t+1$, at which point in time all portfolios are rebalanced.[31]

Table OA.3.9 reports the results of the bivariate portfolio sorts on the basis of both value- and

[29] Imrohoroglu and Tuzel (2014) show that low TFP predicts high returns, Belo et al. (2014) find low hiring is associated with high returns, Titman, Wei, and Xie (2004) documents the relation between low investment rates and high returns, Aretz and Pope (2018) find higher capacity overhang predicts lower returns, and Fama and French (1993) discuss how both low market capitalization and high market-to-book ratios predict high returns.

[30] Untabulated results show that adding sector fixed effects to column (10) of Table 5 produces quantitatively similar results.

[31] The portfolio breakpoints used in the bivariate sorting procedure described above (the 30^{th} and 70^{th} percentiles) are necessary to ensure that there is a sufficient number of firms in each of the *nine* doubles-sorted portfolios (note that Table OA.3.1 shows that the utilization premium is sizable and significant using a univariate sort that is based on these breakpoints).

equal-weighted portfolio returns. The rightmost column of each panel shows the capacity utilization spread, along with its associated p-value, within portfolios that control for a characteristic of interest. Panels A and B report the results obtained by first controlling for book-to-market ratios, Panels C and D report the results obtained by first controlling for investment rates, while Panels E and F report the results obtained by first controlling for organizational capital. Finally, each panel of the table also reports the p-value from a joint test on the null hypothesis that the capacity utilization spread across all three characteristic-sorted portfolios is zero.

The results show that after controlling for either book-to-market ratios, investment rate, organizational capital, the capacity utilization spread remains positive in 15 out of 18 cases. The utilization spread is also quantitatively large and statistically significant in most cases.

Panel A shows that, keeping book-to-market ratios relatively constant, the equal-weighted capacity utilization spread is significantly different from zero at the 10% level within the low book-to-market portfolio and at the 5% level for both the medium and the high book-to-market portfolios. The joint p-value across the three spread portfolios is under 8%. The value-weighted returns reported in Panel B show that the capacity utilization spread is most pronounced among growth firms. Within this low book-to-market portfolio, the capacity utilization spread exceeds 6% per annum. The p-value of 0.016 associated with the joint test in Panel B shows that the three value-weighted utilization spreads are statistically significant after conditioning on book-to-market.

Panels C and D show that, regardless of whether portfolio returns are value-weighted or equal-weighted, the capacity utilization spread typically exceeds 4% per annum within the portfolios of low and medium investment rate firms. In each of these cases the utilization spread is significantly different from zero at better than the 1% level. While the capacity utilization spread is not significant within the high investment rate portfolios, the joint test reported in each Panel is still rejected at the 5% level. Panel C (Panel D) shows that, conditioning on investment rates, the three equal-weighted (value-weighted) capacity utilization spreads are jointly and significantly different from zero at the 1% (5%) level.

Lastly, Panels E and F indicate that the utilization spread exists *within* organizational capital portfolios. In particular, the value-weighted utilization spread among firms with medium amounts of organization capital exceeds 4% per annum, and is statistically significant at better than the 1% level. Moreover, the joint test reported in each panel is rejected at the 5% level or better.

Table OA.3.9: **Controlling for book-to-market ratios, investment rates, and organizational capital: double-sort analysis**

		Panel A: Capacity Utilization (EW)					Panel B: Capacity Utilization (VW)				
		Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)	Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)
Low (L)	BE/ME	9.40	8.10	6.26	3.14	(p=0.079)	11.63	10.53	5.55	6.08	(p=0.001)
Medium		16.92	13.36	12.82	4.10	(p=0.012)	13.29	10.45	11.01	2.27	(p=0.084)
High (H)		19.89	17.49	16.66	3.23	(p=0.039)	14.47	12.70	14.49	-0.02	(p=0.503)
					Joint test	(p=0.076)				Joint test	(p=0.016)

		Panel C: Capacity Utilization (EW)					Panel D: Capacity Utilization (VW)				
		Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)	Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)
Low (L)	I/K	20.06	17.74	12.62	7.44	($p < 0.001$)	15.52	11.30	10.55	4.97	(p=0.007)
Medium		17.07	13.65	13.08	3.99	(p=0.005)	13.35	10.64	9.07	4.28	(p=0.008)
High (H)		10.60	7.73	9.02	1.59	(p=0.249)	9.96	9.26	7.74	2.23	(p=0.180)
					Joint test	($p < 0.001$)				Joint test	(p=0.045)

		Panel E: Capacity Utilization (EW)					Panel F: Capacity Utilization (VW)				
		Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)	Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)
Low (L)	OC / AT	10.39	7.70	13.27	-2.87	(p=0.861)	10.18	10.11	11.16	-0.98	(p=0.661)
Medium		17.30	13.55	14.52	2.78	(p=0.021)	15.57	12.46	11.31	4.26	(p=0.007)
High (H)		20.19	17.50	16.60	3.59	(p=0.027)	16.57	12.20	14.63	1.94	(p=0.240)
					Joint test	(p=0.018)				Joint test	(p=0.038)

The table reports portfolio returns obtained from conditional double-sort procedures, where the controlling variable (i.e., the first dimension sorting variable) is either a firm's book-to-market ratio, investment rate, or organization capital-to-assets ratio, and the second sorting variable is a firm's rate of capacity utilization. The sorting algorithm is as follows: First, at the end of each June, we sort the cross-section of firms into three portfolios on the basis of either the book-to-market ratio, the investment rate, or organizational capital using the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of the characteristic of interest. Second, within each portfolio formed on the basis of the first sorting variable, we further sort firms into three additional portfolios on the basis of capacity utilization, using the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of capacity utilization rates in March of the same year. This process produces nine portfolios that are each held from the beginning of July in year t to the end of June in year $t+1$, at which point in time all portfolios are rebalanced. Portfolio returns are reported for both equal-weighted ("EW", Panels A, C, and E) and value-weighted ("VW", Panels B, D, and F) schemes. The rightmost column of each Panel shows the capacity utilization spread, along with its associated p-value, within portfolios that are first sorted on the controlling variable. These p-values are constructed using Newey and West (1987) standard errors. Each Panel also reports the p-value from a joint test on the null hypothesis that the capacity utilization spread across all three characteristic-sorted portfolios is zero. Panels A and B report the results obtained by first controlling for book-to-market ratios, Panels C and D report the results obtained by first controlling for investment rates, while Panels E and F report the results obtained by first controlling for organizational capital. The sample period is from July 1967 to December 2015.

Independence using projections. As an additional method of showing the independence between the utilization premium and other spread, we consider the following set of projections that regress the utilization premium at time t on another spread, including a constant

$$Utilizationpremium_t = \alpha + \beta X_t^{Other} + \varepsilon_t. \tag{OA.3.5}$$

Here, X_t^{Other} represents the returns of the book-to-market, TFP, investment, size, profitability, momentum, or idiosyncratic volatility spread at the t. If the utilization premium is simply a linear transformation of the value premium (or any one of the other spreads represented by X_t^{Other}), then the constant in this regression (α) will be zero. In contrast, the results in Panel A of Table OA.3.20 indicate that the utilization premium is economically and statistically distinct from these spreads. For instance, projecting the utilization premium on the value premium results in an alpha of 5.11%

per annum (t-statistic of 2.04). Thus, the utilization premium is materially distinct from a host of (potentially related) risk premia.

OA.3.5 Independence from capital overhang

Aretz and Pope (2018) document that firms with higher capital overhang, or firms' whose installed productive capacities exceed their optimal amounts of capacity, have lower expected returns. The authors refer to these firms as possessing "capacity overhang." While the Fama and MacBeth (1973) regressions in Section 1.6 show that the utilization premium and the overhang spread are empirically distinct, the conceptual similarity between these margins motivates us to discuss how the notion of capacity utilization materially differs from that of capacity overhang. We also complement the regression analysis by showing that utilization and overhang each have a distinct impact on stock returns using portfolio double sorts.

Recalling equation (1), capacity utilization is defined as the ratio of a firm's actual output to its maximum potential output (its capacity). On the other hand, capacity overhang is the difference between a firm's *installed* capital stock and its *optimal* (value maximizing) level of capital. Intuitively, capacity utilization and capacity overhang are negatively related since a firm that desires to downscale can reduce its output by lowering the utilization of its existing capital. At the same time, the level of the firm's optimal capital stock also drops. If capital adjustments are not frictionless, then these frictions create a wedge between installed and optimal capacity, resulting in capacity overhang. Consequently, capacity utilization tends to decrease at the same time that overhang tends to increase.

The negative correlation between utilization and overhang is neither theoretically perfect nor empirically large in magnitude. Theoretically, the reason for this less than perfect correlation is that low capacity utilization is a result of a *costless and optimal* policy to keep some machines idle.[32] This optimal decision to reduce the utilization of capital does not hinge on any installation frictions or adjustment costs. In contrast, capacity overhang depends crucially on the degree to which investment is irreversible, as influenced by frictions such as convex adjustment costs. While low capacity utilization is optimal in states of low productivity, a non-zero amount of capacity overhang can never represent the first-best outcome for a firm. Consequently, capacity overhang should always be zero in a frictionless economy, whereas capacity utilization may still fluctuate depending on a firm's productivity.

While capacity overhang and capacity utilization are conceptually distinct, the rest of this section examines whether the two effects are also empirically distinct. Since Aretz and Pope (2018) document that high capacity overhang is associated with low expected returns there is no ex-ante reason to believe that the overhang effect is driving the capacity utilization spread. This is because low capacity utilization firms tend to have both high returns and high amounts of capacity overhang. Nonetheless, we perform portfolio double sorts to ensure that the capacity utilization spread is empirically separate from the overhang effect. We show that, controlling for capital adjustment frictions and the degree of irreversibility via the overhang measure of Aretz and Pope (2018), the capacity utilization spread survives. Moreover, controlling for the frictionless production decisions represented by capacity utilization, the overhang effect also survives.

[32] Keeping machines idle in bad states is not only costless, but may also benefit the firm by preserving capital for future use in more productive states.

To implement this analysis we construct a measure of firm-level capital overhang based on the statistical procedure described by Aretz and Pope (2018), summarized in Section OA.1 of the Online Appendix. Following the discussion on the conceptual relation between capacity utilization and capacity overhang, Table OA.3.10 shows the correlation between overhang and utilization for each industry in our sample.[33] The magnitude of the correlation between the two variables decreases with the degree of aggregation. When we aggregate all firms in our sample, the correlation between capacity utilization and capacity overhang is negative, as expected, and amounts to -0.52. When we compute the correlation between these two variables on an industry-by-industry basis and average these pairwise correlations, the result is a modest average correlation of -0.32. The 95% confidence interval for this cross-sectional correlation shows a high degree of dispersion and ranges from -0.71 to 0.11. Panel C of this table reports that the average firm-level correlation drops to -0.11, and shows that this correlation becomes even more dispersed in the cross-section of firms. These results collectively highlight the fact that while capacity utilization and overhang are conceptually negatively related, the empirical correlation between these two variables is low.

Table OA.3.11 reports the results of performing portfolio double sorts along the dimensions of capacity utilization and capacity overhang using the conditional double-sort analysis as described in Section OA.3.4. Panel A shows the average annual capacity utilization spread within three capacity overhang sorted portfolios when all returns are equal-weighted. The capacity utilization spread is positive and statistically significant within each overhang portfolio. The utilization spread is also jointly significant across all three overhang portfolios. Panel B shows that the results are similar when returns are value-weighted. Panels C and D report that the results are largely similar after changing the order of the sorts. Controlling for capacity utilization, the joint tests in Panels C and D show that the capacity overhang spread is positive and statistically significant on an equal-weighted basis, but is insignificant on a value-weighted basis.

[33]The industry-level capacity overhang measure is obtained by computing the average overhang for all firms that belong to each industry at each point in time. We also note that our sample is only comprised of manufacturing, mining, and utilities firms, whereas the sample of Aretz and Pope (2018) includes the entire Compustat universe, excluding financial firms and utilities.

Table OA.3.10: **Correlation between capacity utilization and capacity overhang**

Panel A: Correlation by Industry		
Industry name	Sector	$\rho_{CU,OVER}$
Food, beverage, and tobacco	ND	-0.741
Printing and related support activities	ND	-0.728
Textile mills	ND	-0.690
Wood product	D	-0.687
Textiles and products	ND	-0.683
Beverage and tobacco product	ND	-0.642
Textile product mills	ND	-0.543
Computer and electronic product	D	-0.537
Food	ND	-0.534
Machinery	D	-0.528
Nonmetallic mineral product	D	-0.502
Support activities for mining	MU	-0.500
Coal mining	MU	-0.472
Computers, communications eq., and semiconductors	D	-0.469
Metal ore mining	MU	-0.450
Communications equipment	D	-0.392
Paper	ND	-0.366
Mining	MU	-0.348
Leather and allied product	ND	-0.345
Transportation equipment	D	-0.342
Mining (except oil and gas)	MU	-0.342
Semiconductor and other electronic component	D	-0.329
Automobile and light duty motor vehicle	D	-0.310
Motor vehicles and parts	D	-0.300
Primary metal	D	-0.254
Artificial and synthetic fibers and filaments	ND	-0.250
Chemical	ND	-0.248
Fabricated metal product	D	-0.214
Electrical equipment, appliance, and component	D	-0.195
Aerospace and miscellaneous transportation eq.	D	-0.183
Computer and peripheral equipment	D	-0.182
Apparel	ND	-0.170

Continued on the next page...

Table OA.3.10 – Continued from the previous page

Panel A: Correlation by Industry

Industry name	Sector	$\rho_{CU,OVER}$
Nonmetallic mineral mining and quarrying	MU	-0.144
Apparel and leather goods	ND	-0.140
Furniture and related product	D	-0.134
Plastics and rubber products	ND	-0.047
Plastics material and resin	ND	0.002
Iron and steel products	D	0.004
Petroleum and coal products	ND	0.071
Miscellaneous	D	0.083
Oil and gas extraction	MU	0.157
Synthetic rubber	ND	0.242

Panel B: Industry-level Summary Statistics

Statistic	Mean	Median	p5	p95
$\rho_{CU,OVER}$	-0.32	-0.34	-0.71	0.11

Panel C: Firm-level Summary Statistics

$\rho_{CU,OVER}$	-0.11	-0.13	-0.66	0.51

Panel A shows the correlation between industry-level capacity utilization and industry-level capital overhang for each industry in the sample. Overhang at the industry level is computed as the simple average of firm-level overhang rates for all firms that belong to each industry. Panel B reports summary statistics for the industry-level correlations between capacity utilization and capacity overhang that are reported in Panel A. These summary statistics include the cross-sectional mean, median, 5th and 95th percentiles of the distribution of industry-level correlation coefficients. Panel C reports these same summary statistics for firm-level correlations between capacity utilization and capacity overhang.

Table OA.3.11: **Double-sorted portfolios: capacity utilization versus capacity overhang**

		Panel A: Capacity Utilization (EW)					Panel B: Capacity Utilization (VW)				
		Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)	Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)
Low (L)	Overhang	19.27	16.14	15.99	3.28	(p=0.026)	17.03	12.21	11.75	5.28	(p=0.025)
Medium		17.59	14.04	14.54	3.05	(p=0.023)	14.14	10.38	11.21	2.93	(p=0.059)
High (H)		14.10	10.29	8.78	5.32	(p=0.021)	12.84	10.83	7.90	4.94	(p=0.007)
					Joint test	(p=0.061)				Joint test	(p=0.079)

		Panel C: Overhang (EW)					Panel D: Overhang (VW)				
		Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)	Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)
Low (L)	CU	19.00	17.29	13.48	5.52	(p<0.001)	16.03	13.84	12.24	3.79	(p=0.050)
Medium		15.49	14.00	10.98	4.51	(p<0.001)	10.27	10.21	10.24	0.03	(p=0.494)
High (H)		16.31	13.55	9.07	7.23	(p<0.001)	12.29	10.35	8.87	3.42	(p=0.029)
					Joint test	(p<0.001)				Joint test	(p=0.153)

The table reports portfolio returns obtained from conditional double-sort procedures in which one sorting variable is capacity overhang and other sorting variable is capacity utilization. Two cases are considered: in Panels A and B the controlling variable (i.e., the first dimension sorting variable) is overhang, and the variable used in the second-stage sort is capacity utilization. In Panels C and D, the order is flipped: the first (second) stage sorting variable is capacity utilization (overhang). The sorting algorithm is as follows. First, at the end of each June, we sort the cross-section of firms into three portfolios on the basis of the first sorting variable, using the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of the variable of interest. Second, within each portfolio formed on the basis of the first sorting variable, we sort firms into three additional portfolios on the basis of the second sorting variable, using the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of the variable. This process produces nine portfolios that are each held from the beginning of July in year t to the end of June in year $t+1$, at which point in time all portfolios are rebalanced. Both equal-weighted ("EW", Panels A and C) and value-weighted ("VW", Panels B and D) portfolio returns are reported. The rightmost column of each Panel shows the spread on the basis of the second sorting variable, along with the p-value associated with null hypothesis this spread is zero. These p-values are constructed using Newey and West (1987) standard errors. Each Panel also reports the p-value from a joint test on the null hypothesis that the three spreads obtained by forming portfolios in the second stage are jointly equal to zero. The sample period is from July 1967 to December 2015.

OA.3.6 Dissecting the TFP spread

Imrohoroglu and Tuzel (2014) show that low productivity (TFP) firms earn a high risk premium. While the results of Fama and MacBeth (1973) regressions reported in Table 5 show that utilization predicts return beyond TFP, this section explores the relation and the distinction between the capacity utilization premium and the productivity premium. The relation between the two can be motivated by the general form of a firm's production function:

$$Y = \underbrace{\text{Technology} \times \text{Markups} \times \text{Utilization}}_{\text{Total factor productivity (TFP)}} \cdot F(K, L), \qquad \text{(OA.3.6)}$$

where $F(\cdot)$ is a production function over capital (K) and labor (L). The residual obtained by projecting output on factor-share weighted capital and labor provides the Solow residual, or an estimate for TFP. This TFP can then be decomposed into three elements: technology shocks, time-varying markups, and time-varying capacity utilization.

We begin by examining whether the TFP spread exists in our sample of manufacturing firms, mining firms, and utilities. This is necessary because our sample is more constrained compared to the sample in Imrohoroglu and Tuzel (2014). The results of replicating the TFP spread in our subsample of firms are reported in Panel A of Table OA.3.12. The equal-weighted TFP spread amount to 4.22% per annum and is statistically significant.[34]

[34] The value-weighted TFP spread is positive yet statistically insignificant using our subsample and time frame.

Table OA.3.12: **Dissecting the productivity spread**

Panel A: Univariate sorts on TFP				
	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	11.68	23.11	17.00	25.04
Medium	12.25	17.24	15.25	20.23
High (H)	11.16	16.18	12.78	20.73
Spread	0.51	13.79	4.22	12.23
(L-H)	(0.25)		(2.26)	

Panel B: TFP spread controlling for CU						
		TFP (EW)				
		Low (L)	Medium	High (H)	Spread(L-H)	p(Spread)
Low (L)	CU	19.24	17.09	14.68	4.56	(p=0.004)
Medium	CU	16.30	13.81	12.09	4.21	(p=0.007)
High (H)	CU	14.02	15.01	12.24	1.78	(p=0.143)
					Joint test	(p=0.031)

Panel C: Univariate sorts on TechMark				
	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	11.84	22.45	16.89	24.54
Medium	11.64	17.16	15.05	20.28
High (H)	11.66	16.17	13.60	20.91
Spread	0.18	12.90	3.29	11.52
(L-H)	(0.09)		(1.87)	

Panel D: Unconditional correlations		
ρ(CU,TFP)	ρ(CU,TechMark)	ρ(TFP,TechMark)
0.39	0.28	0.96

Panel A reports both the annual returns of value- and equal-weighted portfolios formed on total factor productivity (TFP), and the spread between low and high TFP (or productivity) portfolios. Mean (SD) refers to the average (standard deviation) of annual returns, and parentheses report Newey and West (1987) robust t-statistics. Panel B reports equal-weighted portfolio returns obtained from a double sort procedure in which firms are first sorted into three portfolios on the basis of capacity utilization (CU). Within each portfolio, firms are further sorted into three portfolios on the basis of TFP. The rightmost column of the panel show the p-value from a test on the null hypothesis that each TFP spread is zero, as well as a test on null hypothesis that the three spreads are jointly equal to zero. Panel C reports the annual returns of three portfolios sorted on the technology and markups (TechMark) component of TFP. In each of Panels A, B, and C, portfolio breakpoints are based on the 30^{th} and 70^{th} percentiles of the cross-sectional distribution of the characteristic of interest. Panel D shows the pairwise correlations between equal-weighted univariate spreads formed on CU, TechMark, and TFP. The sample period is from July 1967 to June 2015, when the TFP data becomes unavailable. Additional details on the construction of each variable are provided in Section OA.1 of the Online Appendix.

Since capacity utilization is an underlying fundamental component of TFP, we begin by examining whether the TFP spread remains positive when controlling for capacity utilization. We conduct this analysis using a firm-level dependent double sort as described in Section OA.3.4. In other words, we construct the productivity spread within capacity utilization sorted portfolios. The results are reported in Panel B of Table OA.3.12 and show that the TFP spread is 4.56%, 4.21%,

For this reason we only focus on equal-weighted returns in this subsection.

and 1.78% per annum within the portfolio of firms with low, medium, and high rates of capacity utilization, respectively. A joint test on the magnitude of the productivity premium across the three capacity utilization portfolios is statistically significant at the 5% level. This suggests that the productivity premium is distinct from the capacity utilization spread.

Next, we construct a measure for the technology and markup (TechMark) components of TFP by taking the difference between TFP and the capacity utilization rate, as motivated by equation (OA.3.6). This allows us to isolate the component of TFP that is separate from capacity utilization, and examine the relation between this orthogonal component of TFP and stock returns. We sort firms into portfolios based on the TechMark measure at the end of each June and report the results of these univariate sorts in Panel C of Table OA.3.12. The annualized spread between low and high TechMark firms is 3.29% and statistically significant.

Taken together, the results above indicate that the TFP premium is driven by two *distinct* underlying spreads: the TechMark and utilization spreads. Each of these spreads is statistically significant and economically large. We shed light on the contribution of each of these components to the overall productivity spread in Panel D of Table OA.3.12. This panel shows that the correlation between the TFP spread and the capacity utilization (TechMark) spread is 0.39 (0.96).

OA.3.7 Controlling for sectoral effects: Within-sector spread

Table 4 shows that some durable industries are often sorted into the low utilization portfolio, whereas mining industries and utilities often exhibit high capacity utilization rates. The former fact raises the concern that the utilization premium may be a manifestation of the durability spread of Gomes, Kogan, and Yogo (2009). That is, the utilization spread may reflect the know fact that durable manufacturers are riskier than nondurable manufacturers. The latter fact raises the concern that the utilization spread is dominated by one particular sector and may reflect ex-ante heterogeneity between different sectors, as opposed to reflecting a risk premium that exists within sectors. We alleviate both concerns below.

First, only three (two) of the five industries that are most commonly sorted into the low (high) capacity utilization portfolio are durable (nondurable) manufacturers (recall Table 4). Furthermore, the most common industry constituents of the high capacity utilization portfolio are not nondurable manufacturers, as may be expected if the utilization spread were strongly associated with the durability premium.

Second, in the left panel of Table OA.3.13 we examine the utilization premium within a subsample of industries that only includes durable manufacturers. Specifically, we sort the cross-section of 18 durable manufacturers into three portfolios based on the level of capacity utilization following our benchmark sorting procedure. The capacity utilization spread *within* this subsample of durable manufacturers amounts to 5.85% per annum, and is statistically significant. This demonstrates that the utilization spread is also a within-sector phenomenon that is materially unrelated to the ex-ante heterogeneous exposures of durable and nondurable manufacturers to aggregate risk.

Table OA.3.13: **Capacity utilization spread: inclusion and exclusion of major sectors**

	Only Durable Sector		Excluding Mining & Utilities Sector	
Portfolio	Mean	SD	Mean	SD
Low (L)	15.08	24.23	14.39	21.63
Medium	10.39	22.11	10.73	17.88
High (H)	9.23	23.77	9.12	20.21
Spread	5.85	19.08	5.27	17.31
(L-H)	(2.13)		(2.12)	

The table reports the annual returns of portfolios sorted on the basis of capacity utilization, as well as the spread between the low (L) and high (H) capacity utilization portfolios when specific sectors are included or excluded from the sample. The left panel shows the results when the sample includes only industries that are classified as durable goods manufacturers. The right panel shows the results when the sample excludes all mining industries and utilities. The table reports the average value-weighted return (Mean) and standard deviation (SD) of each portfolio's returns. t-statistics, reported in parentheses, are computed using Newey and West (1987) standard errors. The sample period is between July 1967 to December 2015.

Third, we examine the magnitude of the capacity utilization spread when we exclude the only sector that heavily populates the high utilization portfolio: mining and utilities. The mining and utilities sector is also unique in that its average level of capacity utilization over the sample period is statistically different from that of all other industries (see Table OA.2.2). The right Panel of Table OA.3.13 shows the results of sorting all non-mining industries into three portfolios on the basis of capacity utilization. Excluding mining industries and utilities from the sample does not change our baseline results. The utilization spread remains positive, yielding an average return of about 5.3% annually, and statistically significant at the 5% level.

Fourth, in our benchmark analysis we sort industries into portfolios based on the level of each industry's utilization rate. Here we modify this approach by sorting industries into portfolios based on the year-on-year *growth* rate, instead of the *level*, of utilization. Using the growth rate removes any (potential) differences in the average level of utilization across industries. The portfolio formation procedure follows that in Section 1.2, apart from the use of growth rates. The results are reported in Table OA.3.14 and show that the value-weighted (equal-weighted) utilization spread is 4.80% (5.74%) per annum and is significant at the 5% level. Portfolio returns are also monotonically decreasing in the utilization growth rate.

Table OA.3.14: **Capacity utilization spread: sorting on growth rates**

	Value-weighted		Equal-weighted	
Portfolio	Mean	SD	Mean	SD
Low (L)	14.49	21.41	11.53	21.92
Medium	10.05	16.63	7.78	17.59
High (H)	9.69	20.59	5.79	20.93
Spread	4.80	16.93	5.74	16.41
(L-H)	(2.00)		(2.45)	

The table reports the annual returns of three portfolios sorted on the basis of capacity utilization growth, as well as the spread between the low (L) and high (H) utilization growth portfolios. The construction of the portfolios is identical to the benchmark analysis, except that portfolios are sorted on the basis of the growth rate of utilization rather than the level of utilization. The growth rate of utilization is measured between March of years t and $t-1$. Mean refers to the average annual return and SD denotes the standard deviation of annual raw returns, and the parentheses report t-statistics computed using Newey and West (1987) standard errors. The portfolios are formed at the end of each June from 1968 to 2015 and are rebalanced annually, with portfolio returns spanning July 1968 to December 2015

Lastly, we complement the empirical evidence above with a theoretical exercise in Section OA.4.5. We consider the implications of ex-ante parameter heterogeneity on the model-implied utilization premium. Parameter heterogeneity captures any cross-sectoral differences in depreciation, adjustment costs, or elasticity of depreciation to utilization. We show that such heterogeneities contribute only marginally to the utilization premium.

OA.3.8 Robustness: firm-level capacity utilization

For robustness, we construct a proxy for the unobservable capacity utilization rate at the *firm*-level. First, for each industry, we project the utilization rate of industry j at time t ($CU_{j,t}$) on salient industry-level production-related characteristics, contained in the vector $\boldsymbol{X}_{j,t}$. For the dependent variable, $CU_{j,t}$, we use either the raw industry utilization or industry-demeaned utilization rate. The latter approach ensures that the fitted value of this projection is not affected by fixed differences in average utilization across industries. The choice of $\boldsymbol{X}_{j,t}$ is motivated by the model in Section 2. We use the logarithms of size and book-to-market, the investment-to-capital ratio, IVOL, and TFP.[35] The regression is

$$CU_{j,t} = \beta_{j,0} + \boldsymbol{\beta}_j \boldsymbol{X}_{j,t} + \varepsilon_{j,t} \tag{OA.3.7}$$

By estimating this projection separately for each industry, the relation between utilization and the characteristics, as measured by $\hat{\boldsymbol{\beta}}_j$, is specific to industry j. The average R^2 of this projection across industries is sizable at 33%, suggesting that the regressors well-span utilization at the industry level.

[35] Our empirical analysis shows that utilization affects the exposure of firms to aggregate productivity. Consequently, for $\boldsymbol{X}_{j,t}$ we choose firm-level variables that are known to also correlate with firms' exposure to this factor. In particular, Zhang (2005) shows that aggregate productivity exposure interacts with book-to-market and investment. Imrohoroglu and Tuzel (2014) establish the relation between firm-level TFP and aggregate productivty. Ai and Kiku (2016) document that idiosyncratic volatility serves as proxy for underlying growth options, whose riskiness depends on aggregate consumption. Importantly, in untabulated results we verify that our findings are robust to either removing particular characteristics (e.g., IVOL) or adding additional characteristics (e.g., hiring rates).

Second, the proxy for the utilization rate of a firm i that belongs to industry j at time t (denoted $\hat{CU}_{i,j,t}$) is obtained by combining the estimated slope coefficients for industry j, obtained via equation (OA.3.7), with the observable characteristics of firm i, denoted $\boldsymbol{X}_{i,j,t}$,

$$\hat{CU}_{i,j,t} = \hat{\beta}_{j,0} + \hat{\boldsymbol{\beta}}_j \boldsymbol{X}_{i,j,t}. \tag{OA.3.8}$$

This procedure allows the utilization proxy to vary between firms *within* the same industry. We use the firm-level utilization proxy to sort firms into portfolios as per Section 1.2, and report the results in Table OA.3.15.

Table OA.3.15: **Capacity utilization spread: proxy for firm-level utilization rates**

	Utilization		De-meaned utilization	
Portfolio	Mean	SD	Mean	SD
Low (L)	12.65	19.92	11.92	17.67
Medium	11.98	15.93	11.80	16.50
High (H)	7.66	21.58	6.77	22.08
Spread	4.98	15.39	5.14	15.48
(L-H)	(2.32)		(2.02)	

The table reports the annual value-weighted returns of portfolios sorted on the basis of estimated firm-level capacity utilization rates, as well as the spread between the low (L) and high (H) utilization portfolios. The table reports the average value-weighted return (Mean) and standard deviation (SD) of each portfolio's returns, and all portfolios are formed by following the procedure described in Section 1.2. t-statistics, reported in parentheses, are computed using Newey and West (1987) standard errors. The sample period is between July 1967 to December 2015.

The table shows that the firm-level utilization premium is about 5% per annum and statistically significant. Similar results are obtained using both raw or industry-demeaned utilization rates in projection (OA.3.7) (i.e., excluding industry fixed effects). The relation between firm-level utilization and average returns also remains monotonically decreasing in either case. The findings above help to further illustrate that the benchmark utilization premium is not driven by ex-ante heterogeneity across industries.

OA.3.9 Supplemental tables

Table OA.3.16: **Transition matrix of constituents between capacity utilization portfolios**

Portfolio in	Portfolio in year $t+1$		
year t	Low	Medium	High
Low	0.746	0.254	0.000
Medium	0.033	0.939	0.027
High	0.011	0.232	0.758

The table shows the probability of an industry sorted into portfolio $i \in$ {Low, Medium, High} in year t, where i is the row index, being sorted into portfolio $j \in$ {Low, Medium, High} in year $t+1$, where j is the column index. The transition probabilities are computed using annual capacity utilization data from June 1967 to December 2015. Industries are sorted into portfolios at the end of each June following the portfolio formation procedure described in Section 1.2.

Table OA.3.17: **Capacity utilization spread: decile portfolios**

	Panel A: Portfolio returns		Panel B: Market betas	
Portfolio	Mean	SD	β	$t(\beta)$
Low (L)	13.64	21.23	1.24	(11.74)
2	12.74	18.84	1.16	(14.51)
3	10.89	17.81	1.14	(14.00)
4	11.39	19.48	1.17	(17.34)
Medium	10.56	19.97	1.20	(11.78)
6	9.26	18.35	1.16	(14.51)
7	10.65	18.07	1.08	(12.71)
8	9.02	18.69	1.07	(11.39)
9	9.94	17.54	1.03	(13.57)
High (H)	7.96	20.22	1.05	(11.59)
Spread	5.67	17.71	0.19	(2.96)
(L-H)	(2.31)			

The table reports annual returns (Panel A) and CAPM betas (Panel B) of portfolios sorted on the basis of capacity utilization, as well as the spread between the returns and risk exposures of the low (L) and high (H) capacity utilization portfolios. The construction of the portfolios is identical to the benchmark analysis, except for the portfolio breakpoints used. Specifically, we sort the cross-section of industries into portfolios based on decile breakpoints. The mean in Panel A refers to the average annual return, obtained by multiplying the average monthly return by 12, and SD denotes the annualized standard deviation of monthly returns. Parentheses report Newey and West (1987) t-statistics. All portfolios are formed at the end of each June and are rebalanced annually. The sample is from July 1967 to December 2015. The risk exposures in Panel B correspond to β_1 from the regression $Ret^e_{i,t} = \beta_0 + \beta_1 \text{MKTRF}_t + \varepsilon_{i,t}$, where $Ret^e_{i,t}$ is the value-weighted excess return of a portfolio and MKTRF_t is the excess market return, a proxy of aggregate productivity. Parentheses report Newey and West (1987) t-statistics. All portfolios are formed at the end of each June and are rebalanced annually. The sample is from July 1967 to December 2015.

Table OA.3.18: **Fama-MacBeth regressions: industry-level**

	Panel A: Annual		Panel B: Monthly	
	(1)	(2)	(3)	(4)
β_{CU}	-1.77	-1.52	-0.15	-0.14
	(-2.16)	(-2.01)	(-2.04)	(-1.95)
Sector FE	-	Yes	-	Yes
Controls	Yes	Yes	Yes	Yes
R^2	0.475	0.540	0.459	0.533

The table reports the results of Fama-MacBeth regressions in which future industry-level excess returns are regressed on current utilization rates and industry-level controls. We run the following cross-sectional regression in which the dependent variable is either an industry's annual excess return from the start of July in year t to the end of June in year $t+1$ (Panel A), or an industry's monthly excess return from the end of month t to the end of month $t+1$ (Panel B). The independent variables are the utilization rate and a vector of the industry's characteristics (controls), $\boldsymbol{X}_t$ measured at the end of June in year t (Panel A) or every month if available (Panel B): $R_{i,t\to t+1} = \beta_0 + \beta_u u_{i,t} + \boldsymbol{\beta}_t' \boldsymbol{X}_{i,t} + \varepsilon_{i,t\to t+1}$. The controls $\boldsymbol{X}_t$ include total factor productivity (TFP), the hiring rate (HIRE), the natural investment rate (I/K), capacity overhang (OVER), the ratio of organization capital to assets (OC / AT), the natural logarithm of the market value of equity (ln(ME)), the natural logarithm of the book-to-market (ln(B/M)) ratio, and lagged annual return (RET_{t-1}). Here, all industry-level characteristics (other than CU) are calculated as the average characteristic across all firms assigned to a given industry at each point in time. After running these cross-sectional regressions we compute the time-series average of each element of the vectors the estimated slope coefficients. Parenthesis report Newey and West (1987) t-statistics. Columns 1 to 4 show the results when all characteristics are used in multivariate regressions, while Columns 2 and 4 also including sector fixed effects. Each control variable is standardized by dividing it by its unconditional standard deviation. The table also report the time-series average of the R^2 obtained from each set of cross-sectional regressions.

Table OA.3.19: **Fama-MacBeth regressions: excluding small and microcap firms**

	(1)	(2)	(3)
CU	-1.56	-1.52	-2.02
	(-1.92)	(-2.26)	(-3.81)
TFP		1.27	1.20
		(2.84)	(2.65)
HIRE		-4.20	-4.15
		(-2.72)	(-2.84)
I/K		-2.30	-2.22
		(-2.17)	(-2.36)
OVER		-3.62	-3.83
		(-2.98)	(-3.32)
OC / AT		0.80	0.62
		(0.65)	(0.65)
ln(ME)		-0.18	-0.32
		(-0.17)	(-0.29)
ln(B/M)		3.71	3.69
		(4.18)	(4.27)
RET_{t-1}		-0.34	-0.29
		(-0.22)	(-0.21)
Sector FE	-	-	Yes
R^2	0.012	0.119	0.147

The table reports the results of Fama-MacBeth regressions in which future excess returns are regressed on current characteristics. In each year t we first identify firms in our sample that have a market capitalization rate less than the median market capitalization at that point in time, and remove these firms from the cross-section of firms (i.e., we restrict our focus to firms with large market capitalizations). We then run the following cross-sectional regression in which the dependent variable is a firm's annual excess return from the start of July in year t to the end of June in year $t+1$, and the independent variables are a vector of the firm's characteristics, $\boldsymbol{X}_t$ measured at the end of June in year t: $R_{i,t\to t+1} = \beta_0 + \boldsymbol{\beta}_t' \boldsymbol{X}_{i,t} + \varepsilon_{i,t\to t+1} \quad \forall t \in \{1967, \ldots, 2014\}$. The characteristics considered are capacity utilization (CU), total factor productivity (TFP), the hiring rate (HIRE), the natural investment rate (I/K), capacity overhang (OVER), the ratio of organization capital to assets (OC / AT), the natural logarithm of the market value of equity (ln(ME)), the natural logarithm of the book-to-market (ln(B/M)) ratio, and lagged annual return (RET_{t-1}). After running these cross-sectional regressions we compute the time-series average of each element of the vectors the estimated slope coefficients, $\{\hat{\boldsymbol{\beta}}_t\}_{t=1967}^{2014}$. Each column reports the average slope coefficients for the characteristics of interest. Parenthesis report Newey and West (1987) t-statistics. Column 1 shows the results when the capacity utilization rate is the only predictor, while Columns 2 and 3 show the results when all characteristics are used in multivariate regressions. Column 3 also including sector fixed effects. Each control variable is standardized by dividing it by its unconditional standard deviation. The table also report the time-series average of the R^2 obtained from each set of cross-sectional regressions. The first regression is run in 1967 and the last regression is run in 2014, when the TFP data becomes unavailable.

Table OA.3.20: **Distinction between the utilization premium and other spreads: projection evidence**

Other spread	BE/ME	TFP	I/K	ME	Profit.	MOM	IVOL
			Panel A: Empirical results				
α^{Data}	5.11	5.40	4.74	5.39	6.33	7.06	6.76
	(2.04)	(2.20)	(1.99)	(2.21)	(2.63)	(2.84)	(2.72)
			Panel B: Model-implied results				
α^{Model}	4.90	6.64	3.90	6.58	-	-	-

The table considers the distinction between the utilization premium and other spreads in both the data (Panel A) and the model (Panel B) through the lens of time-series regressions. Specifically, each panel reports the results of a univariate time-series regression in which returns associated with the utilization premium are projected on returns associated with another spread. The regression we estimate is given by $R_t^{CU} = \alpha + \beta R_t^{Other} + \varepsilon_t$, where R_t^{CU} represents the utilization premium in month t and R_t^{Other} represents the return associated with another spread in the same month. The other spreads we consider are: (i) the value premium (BE/ME), (ii) the productivity premium (TFP), (iii) the investment premium (I/K), (iv) the size effect (ME), (v) the profitability premium, (vi) the momentum effect, and (vii) the idiosyncratic volatility effect. The time-series regressions in Panel A are estimated using monthly data between July 1967 and December 2015, and then annualized. Parentheses report Newey and West (1987) t-statistics. The time-series regressions in Panel B are estimated using model-implied data from population-sample simulations of our extended models in Section OA.5, that feature two aggregate shocks (aggregate productivity and investment efficiency). Specifically, BE/ME, I/K and ME spreads are implied by the model in Section OA.5.1, while the TFP premium is implied by the model in Section OA.5.2. Panel B then reports the values of the model-implied alphas obtained by estimating the aforementioned regressions.

Table OA.3.21: **Return predictability using utilization and productivity proxies-**

	Panel A: Industry-level			Panel B: Firm-level		
CU	-1.85	-2.28	-2.34	-1.69	-2.30	-2.55
	(-1.79)	(-2.06)	(-2.24)	(-2.09)	(-2.69)	(-3.10)
ln (Sales / Capital)	-0.14		-0.09	0.50		0.21
	(-0.43)		(-0.30)	(0.25)		(0.10)
ln (Sales / Employees)		0.02	0.02		2.82	3.19
		(1.34)	(1.31)		(2.90)	(3.19)
R^2	0.126	0.104	0.170	0.027	0.030	0.043

The table reports the results of Fama-MacBeth regressions in which future industry-level (Panel A) or firm-level (Panel B) excess returns are regressed on current utilization rates and productivity proxies. We run the following cross-sectional regression in which the dependent variable is either an industry's annual excess return from the start of July in year t to the end of June in year $t+1$ (Panel A), or a firm's annual excess return from the start of July in year t to the end of June in year $t+1$ (Panel B): $R_{i,t\to t+1} = \beta_0 + \boldsymbol{\beta}_t' \boldsymbol{X}_{i,t} + \varepsilon_{i,t\to t+1} \quad \forall t \in \{1967, \ldots, 2014\}$. Here, the independent variables X_t are the utilization rate, the logarithmic ratio of sales to physical capital, and the logarithmic ratio of sales-to-employees, proxies for productivity. The industry-level characteristics (other than CU) in Panel A are calculated as the average characteristic across all firms assigned to a given industry at each point in time. After running these cross-sectional regressions we compute the time-series average of each element of the vectors the estimated slope coefficients. Parenthesis report Newey and West (1987) t-statistics. The table also report the time-series average of the R^2 obtained from each set of cross-sectional regressions. Each control variable is standardized by dividing it by its unconditional standard deviation.

OA.4 Additional theoretical results: Asset pricing

OA.4.1 Discussion of the model's assumptions

Countercyclical price of risk. The model assumes a countercyclical market price of risk to break the symmetry between high and low utilization firms in the presence of symmetric convex adjustment costs. While we do not model a micro foundation for this cyclicality, this countercyclicality can arise in a general equilibrium setup by assuming habit preferences or time-varying volatility that is countercyclical (e.g., Bansal and Yaron (2004)).

Fixed cost. The model features a real option to disinvest. While we could, in principle, also introduce a fixed cost for expanding capacity to the model (and make positive investment a real option also), we refrain for doing so to keep the dimensionality of the model's parameters low. Since the prior literature emphasizes that the adjustment costs of disinvestment are larger than those of investment (e.g., Zhang (2005)), our model captures this notion in a parsimonious manner. However, we relax this assumption and use a more comprehensive adjustment cost function in Section OA.6.4. We also provide an extensive discussion on the role of the disinvestment option under Section OA.4.3. There, we show that the non-convex adjustment cost breaks the link between utilization and investment margins.

Sources of risk. Motivated by the empirical evidence in Sections 1.4, our framework relies on exposures to a single priced state variable: productivity. Despite having only a single aggregate shock, the model can generate a distinct value premium and utilization premium, as shown in Section OA.4.2. The distinction arises from the real option nature of investment in the model. This real option to disinvest creates a substitution between a firm's utilization and investment policies, and also introduces additional degrees of non-linearity to model. However, we relax the assumption of a single aggregate shock in Section OA.5.1. Introducing an additional priced source of risk enhances the quantitative distinction between the utilization and value premia.

Wages and labor. Lastly, we demonstrate the importance of wages and flexible labor in the model. Our wage specification is shared with Jones and Tuzel (2013), Belo et al. (2014), and Imrohoroglu and Tuzel (2014), among many other studies. As in the aforementioned studies, our model does not feature a fixed cost that is paid unconditionally every period as in Zhang (2005) (thereby creating operating leverage). In contrast, our model relies on the labor margin which leads to a similar amplification of risk. To see this, consider equation (19). Utilization depends on aggregate productivity with the elasticity A_x that rises when $\omega < 1$. That is, the optimal choice of utilization (of all firms) is sensitive to changes in aggregate productivity though the dispersion in firms' choices of labor. This influences the risk exposures of firms' to aggregate productivity without explicitly introducing operating leverage to the model.

Table OA.4.1 show the model's sensitivity to both flexible labor and ω. Lower (higher) ω has a negligible effect on the utilization premium, but increases (decreases) the equity premium. For example, the equity premium changes by 0.82% per annum when we switch between $\omega = 0$ and $\omega = 1$. In contrast, the utilization premium only changes by about 0.20% per annum. This is in line with equation (19): lower ω increases the risk for all firms. In line with Imrohoroglu and Tuzel (2014), we solve a model with fixed labor (that is, with $L_{i,t} = 1$ for all firms). The resulting utilization premium is considerably smaller (about 1.35% per annum). Thus, flexible labor (or fixed costs of production) are required to create sufficient dispersion in risk exposures across firms.

Table OA.4.1: **Model-implied moments across alternative calibrations for labor and wages**

Row	Model	Time-series		Cross-sectional		Risk premia	
		$S_{TS}(ik)$	$\rho_1(ik)$	$\sigma_{CS}(ik)$	$S_{CS}(ik)$	$E\left[R^M\right]$	$E\left[R^{CU}\right]$
	Data						
(1)		0.67	0.52	0.16	1.89	6.28	5.67
	Baseline						
(2)		0.65	0.58	0.11	1.19	5.13	5.12
	Different ω						
(3)	Low ($\omega = 0.00$)	0.67	0.59	0.11	1.17	5.55	5.08
(4)	High ($\omega = 1.00$)	0.55	0.56	0.10	1.26	4.73	5.28
	Fixed labor						
(5)		-1.87	0.75	0.01	0.23	7.90	1.35

The table reports model-implied population moments related to the time-series and cross-section of investment rates, as well as risk premia, under various calibrations that perturb wages and labor. The table reports the time-series skewness ($S_{TS}(ik)$) and the first-order autocorrelation ($\rho(ik)$) of firm-level investment rates and the cross-sectional dispersion ($\sigma_{CS}(ik)$) and skewness ($S_{CS}(ik)$) of investment rates. In addition, the table also reports the equity premium (($E\left[R^{CU}\right]$)) and the capacity utilization premium ($E\left[R^{CU}\right]$) obtained by sorting the cross-section of model-implied returns association with each calibration on capacity utilization rates. These risk premia are expressed as an annualized percentage. Each alternative calibration is identical to the benchmark calibration in all ways except for altering the sensitivity of wages to aggregate productivity (ω) or fixing the quantity of labor available to each firm. All moments are based on a simulations of 1,000 firms over 40,0000 periods (years). Finally, the top row of the table reports the empirical counterpart of each moment, while the second row of the table reports the value of each moment in our benchmark calibration and model.

OA.4.2 Model-implied distinction between utilization and book-to-market (or investment)

This section discusses the distinction between the utilization premium and the value premium (i.e., investment-related premia) within our benchmark model. To illustrate this point, we report the results of a conditional portfolio double sort of model-implied stock returns on book-to-market ratios and capacity utilization rates. The portfolio formation procedure follows the discussion in Section 3.1. The results of the analysis are reported in Table OA.4.2 and show that the utilization premium also exists *within* book-to-market portfolios.

Table OA.4.2 shows the model produces a sizable utilization premium *within* book-to-market portfolios. There are two reasons why our single-shock model is capable of simultaneously generating a spread along these two separate dimensions. First, despite the comovement between investment, utilization, and book-to-market in the model (all relate to Tobin's q), the correlation between the latter two margins is not perfect. Notably, our model features a real option (i.e., the fixed cost f) that induces "wait and see" periods of investment inaction. In these "wait and see" periods of investment inaction, utilization and investment do not comove as utilization *substitutes* the decision to exercise the costly real option to shed capital. Second, while both utilization and book-to-market are linked to the same aggregate shock, these relations to the aggregate shock are *non linear* because of time-varying betas. Both margins vary over time, and as they do, both

Table OA.4.2: **Conditional double sort in the model**

	Low CU	Medium CU	High CU	Spread (L-H)
Low B/M	6.42	4.96	3.48	2.95
Medium B/M	9.28	7.34	5.69	3.59
High B/M	10.02	8.99	7.60	2.42

The table shows the model-implied equal-weighted returns obtained from a conditional double-sort procedure in which the control variable (i.e., the first dimension sorting variable) is the book-to-market ratio and the second dimension sort variable is the capacity utilization rate. The portfolios are constructed as follows. First, in each period firms are sorted into three portfolios based on the cross-section of book-to-market ratios from period $t-1$ using the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of book-to-market ratios. Next, within each book-to-market portfolio, firms are further sorted into three additional portfolios on the basis of capacity utilization in period $t-1$ using the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of capacity utilization rates. This procedure produces nine portfolios that are held for one period, and are then rebalanced. The table also shows the capacity utilization spread associated with each book-to-market portfolio. Here, model implied moments are based on one simulation of the model that features 1,000 firms and 40,000 periods (years.)

margins change firms' conditional exposures to aggregate productivity shocks.

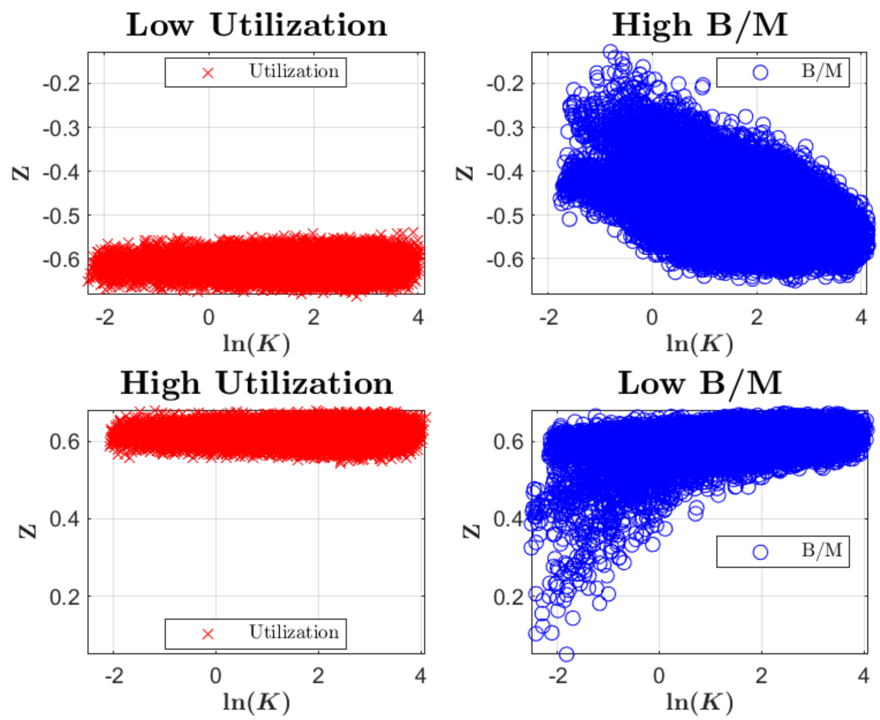
To further illustrate the distinction between low/high utilization firms and high/low B/M firms in our benchmark single-shock model, Figures OA.4.1 and OA.4.2 plot how firms sorted into portfolios according to their utilization rates and B/M ratios differ along two key dimensions: the natural logarithm of their capital stocks (represented by the x-axis and denoted by K) and the level of their idiosyncratic productivities (represented by the y-axis and denoted by Z). We focus on these two variables because they represent the two firm-level state variables in our model.

We produce these figures in three steps. First, we sort the cross-section of model-implied firms in each year t into portfolios based on either their utilization rates or their book-to-market ratios in year $t-1$. When forming these portfolios we use the 10^{th} and 90^{th} percentiles of the cross-sectional distribution of each firm-level characteristic as portfolio breakpoints. This mimics the portfolio sort procedure we implement empirically. Next, we compute the mean value of both the capital stock and the idiosyncratic productivity of all firms assigned to each portfolio. Finally, we produce scatterplots and heat maps that visually describes the level of capital and idiosyncratic productivity for each key portfolio on the portfolio formation dates. Each plot is obtained from a model-implied population simulation that spans 40,000 periods and 1,000 firms.

The scatterplots in the top panel of Figure OA.4.1 display the properties of the low utilization portfolio and the high B/M portfolio. The figure shows that while both low utilization and high B/M firms tend to have large capital stocks and low idiosyncratic productivities, there are differences between the two portfolios in the *magnitude* of these two margins. In particular, low utilization firms are associated with low idiosyncratic productivity (Z) regardless of their capital stocks, whereas there is a substitution between Z and K for high book-to-market (i.e., value) firms. The idiosyncratic productivity of a value firm can be higher than that of a low utilization firm provided the value firm's capital stock is sufficiently small (i.e., low K). Similarly, the bottom panel of the figure shows the types of firms assigned to the high utilization and low B/M portfolios. Once again, high utilization firms are often not synonymous with low B/M firms.

Notably, a firm's exposure to aggregate productivity shocks is ultimately a function of the firm's underlying *state variables* Z and K. By projecting the cross-section of firms onto the (Z, K) plane, these plots help to illustrate why our single-shock model is able to generate both a utilization

Figure OA.4.1: **Utilization and B/M sorted firms: Scatter plots of Z and $\ln(K)$**

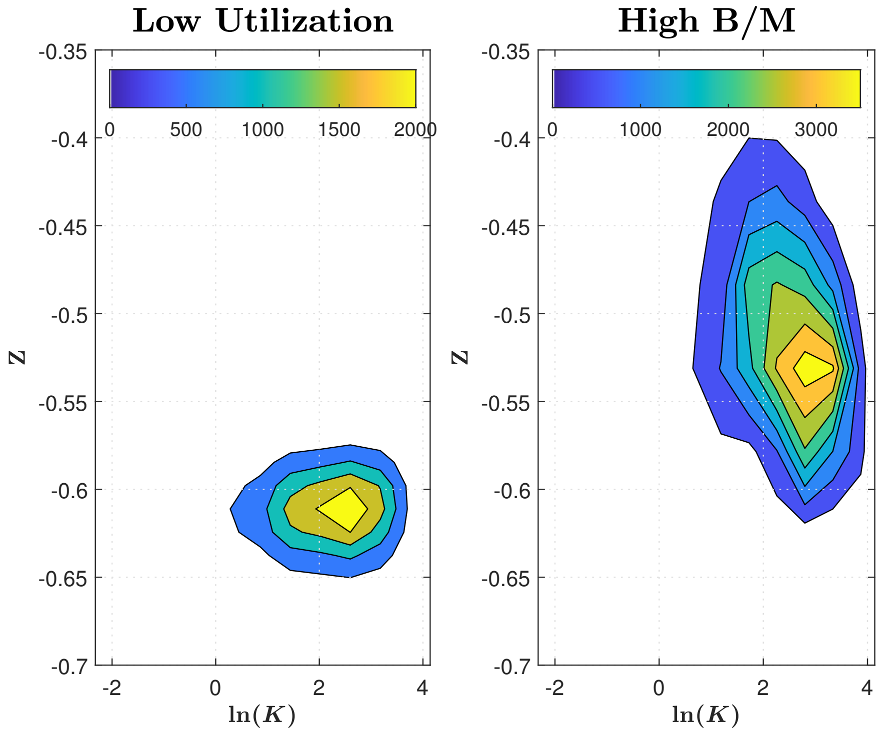


The figure reports scatterplots that display the average capital stock (represented by the x-axis) and the average level of idiosyncratic volatility (represented by the y-axis) of model-implied firms sorted into portfolios on the basis of their capacity utilization rates and book-to-market ratios. We produce these figures in two steps. First, we sort the cross-section of model-implied firms in each year t into portfolios based on either their utilization rates or their book-to-market ratios in year $t-1$. When forming these portfolios we use the 10^{th} and 90^{th} percentiles of the cross-sectional distribution of each firm-level characteristic as portfolio breakpoints. This mimics the portfolio sort procedure we implement empirically. Second, we compute the mean value of both the capital stock and the idiosyncratic productivity of all firms assigned to each portfolio. Each plot is obtained from a model-implied population simulation that spans 40,000 periods and 1,000 firms.

premium and value premium. This is because the types of firms that are sorted into the low utilization portfolio are often quite different from the types of firms that are sorted into the high B/M portfolio. As such, these portfolios have visually different exposures to aggregate productivity risk and consequently earn different expected returns. This ultimately gives rise to relatively distinct utilization and investment premia within our framework.

While the former scatterplots show a number of distinct differences between the firms assigned to the various utilization- and B/M-sorted portfolios, the large number of overlapping points on these diagrams can mask the underlying *distributions* of firms in the economy. To address this point, Figure OA.4.2 plots heat maps that display the types of firms assigned to the key portfolios, along with associated the frequencies of portfolio membership as a function of Z and K. The contours of the low (high) utilization (B/M) portfolios are almost non-overlapping when we account for the frequency of portfolio membership. This once again highlights how a value firm in our single-shock model is not necessarily synonymous with a low utilization firm.

Figure OA.4.2: **Low utilization versus value firms: heat maps of Z and $\ln(K)$**



The figure reports heat maps that display the average capital stock (represented by the x-axis) and the average level of idiosyncratic volatility (represented by the y-axis) of model-implied firms sorted into portfolios on the basis of their capacity utilization rates and book-to-market ratios. We produce these figures in two steps. First, we sort the cross-section of model-implied firms in each year t into portfolios based on either their utilization rates or their book-to-market ratios in year $t-1$. When forming these portfolios we use the 10^{th} and 90^{th} percentiles of the cross-sectional distribution of each firm-level characteristic as portfolio breakpoints. This mimics the portfolio sort procedure we implement empirically. Second, we compute the mean value of both the capital stock and the idiosyncratic productivity of all firms assigned to each portfolio. Each plot is obtained from a model-implied population simulation that spans 40,000 periods and 1,000 firms.

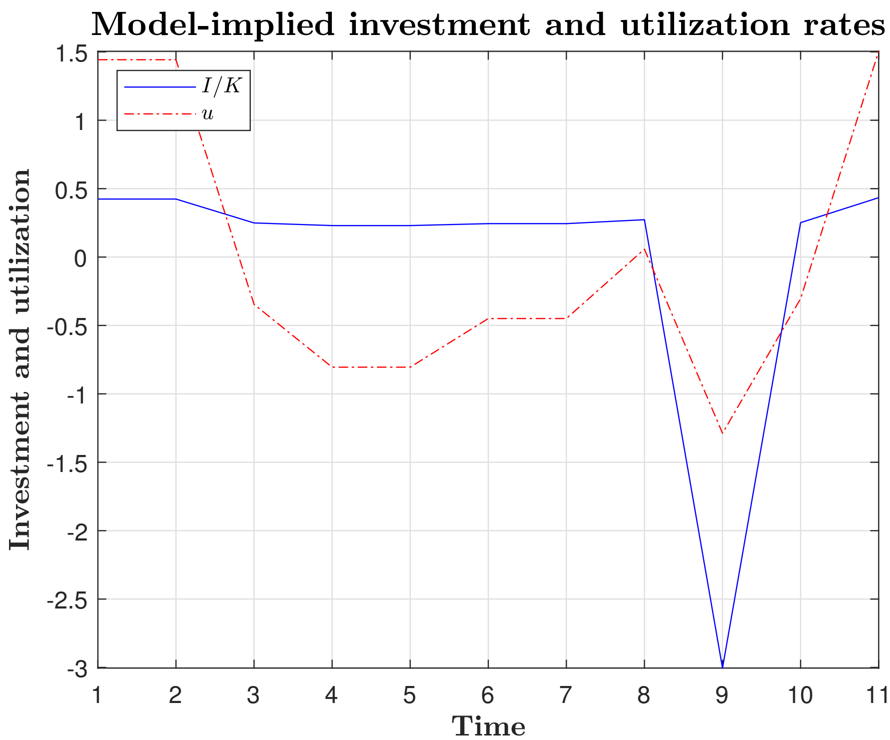
OA.4.3 The role of non-convex disinvestment costs

When the fixed cost of disinvestment is positive ($f > 0$), the decision to shed capital becomes a real option. This means that firms may not necessarily want to sell their capital immediately when a negative shock hits the economy. This is because there are multiple periods of investment inaction in which firms "want and see" if productivity will recover before committing to the partially irreversible decision to shed physical capital. It is these periods of investment inactions that show the role of variable capacity utilization as serving as a *substitute* for selling capital.

Consequently, the fixed cost of disinvestment creates economically important distinctions between capital *utilization* and capital *investment*. Namely, a positive disinvestment cost implies that (i) utilization leads investment during economic downturns, an empirical observation that is commonly used by forecasters, and (ii) the utilization premium features distinction from investment-based spreads, even in a single-shock model. We discuss each role of $f > 0$ in turn.

Utilization as a leading economic indicator. Consider Figure OA.4.3 that displays the (normalized) paths of physical investment and capacity utilization for a model-simulated firm

Figure OA.4.3: **Substitution between investment and utilization**



The figure displays the (normalized) paths of physical investment and capacity utilization for a model-simulated firm around an economic downturn.

around an economic downturn. While this representative firm *slightly* reduces its physical investment rate between periods two and three, its physical investment rate is unchanged between periods three and eight. This is because the firm is stuck in a "wait and see" period of investment inaction while the firm decides whether to exercise its costly real option to shed capital. When the firm ultimately commits to shed physical capital in period nine, its investment rate drops sharply before *slightly* increasing in period 11. Overall, the firm's investment policy is characterized by a lengthy period of investment inaction followed by a "lumpy" decision to disinvest capital.

In contrast, the firm's choice of utilization drops substantially from period two to period six and therefore *predicts* the sharp disinvestment of physical capital that is only observed in period nine. Moreover, the firm's utilization rate continues to vary over time even while its physical investment rate is fixed between periods three and eight. From periods two to eight, the lower utilization rate *substitutes* the firm's investment decision for the purpose of smoothing its dividends. That is, the lower utilization rate implies a lower depreciation that, all else equal, raises the firm's dividend. This lower depreciation rate has the same impact on the firm's dividends as selling machines.

The fact that capacity utilization *leads* investment is a prominent empirical regularity. For instance, the Federal Reserve Board often cites the capacity utilization rate as serving as a useful predictor of output (Koenig et al., 1996) and other business cycle fluctuations (Corrado and Mattey, 1997). This leading behavior would be difficult to observe in a model *without* non-convex adjustment costs and real options. Moreover, without the substitution between investment and utilization that is induced by the real option to disinvest, the firm's investment-related decisions would be too

Table OA.4.3: **Utilization premium within B/M-sorted portfolios: the role of f**

	Panel A: $f > 0$				
	Low CU	Medium CU	High CU	Spread (L-H)	
Low B/M	6.42	4.96	3.48	2.95	
Medium B/M	9.28	7.34	5.69	3.59	
High B/M	10.02	8.99	7.60	2.42	
	Panel B: $f = 0$				
	Low CU	Medium CU	High CU	Spread (L-H)	% change
Low B/M	6.97	6.35	5.69	1.28	56.61%
Medium B/M	8.17	7.30	6.83	1.33	62.95%
High B/M	8.95	8.24	7.89	1.06	56.20%

The table reports the results of conditional portfolio double sorts that construct the utilization premium *within* book-to-market sorted portfolios using our benchmark model from Section 2 with the fixed cost of disinvestment (i.e., $f > 0$) in Panel A and without the fixed cost of disinvestment (i.e., $f = 0$) in Panel B. The portfolio sorts in each panel are conducted as follows. First, we simulate a cross-section of 1,000 model-implied firms for 40,000 time periods. Then, we then conduct a portfolio double sort in which we first sort the cross-section of firms on the basis of book-to-market ratios. That is, on each date t we sort the cross-section of firms into one of three portfolios on the basis of their book-to-market ratios on date $t-1$. We assign firms to these portfolios using the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of book-to-market ratios. Next, *within* each of these three book-to-market-sorted portfolios, we further sort the cross-section of firms into one of three portfolios based on the cross-sectional distribution of capacity utilization rates on date $t-1$. Here, we use the same portfolio breakpoints as in the previous step. This produces a series of nine portfolios that we hold for one period, at which point in time all portfolios are rebalanced. The table then reports the average return of each portfolio, as well as the spread in the returns between the low and high utilization-sorted portfolios. Moreover, in Panel B, we report the percentage change between the conditional utilization premium obtained in Panel A, from the economy in which $f > 0$, to the conditional utilization premium obtained in Panel B, from the economy in which $f = 0$.

highly correlated with utilization rates (counterfactually so).

Distinction from investment-based spreads. The real option to disinvest also introduces a large degree of non-linearity to the production-based model in Section 2. In turn, this non-linear relation between the fundamental aggregate productivity shock and both investment and capacity utilization decisions allows the model to produce book-to-market and utilization spreads that are relatively distinct despite the model only featuring a single source of risk: aggregate productivity. Section OA.4.2 further elaborates on the economic mechanism that separates the utilization premium and the value premium within the baseline model.

To illustrate the importance of $f > 0$ for this between the utilization premium and investment premia, we start by simulating a cross-section of 1,000 model-implied firms for 40,000 time periods in an economy that features a positive fixed cost of disinvestment. We then conduct a portfolio double sort in which we first sort the cross-section of firms on the basis of book-to-market ratios. That is, on each date t we sort the cross-section of firms into one of three portfolios on the basis of their book-to-market ratios on date $t-1$. We assign firms to these portfolios using the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of book-to-market ratios. Next, *within* each of these three book-to-market-sorted portfolios, we further sort the cross-section of firms into one of three portfolios based on the cross-sectional distribution of capacity utilization rates on date $t-1$. We once again use the same portfolio breakpoints as the previous step. This produces a series of nine portfolios that we hold for one period, at which point in time all portfolios are rebalanced.

Panel A of Table OA.4.3 reports the average annual portfolio returns associated with this

Table OA.4.4: **The role of flexible utilization for key moments: the case of $f = 0$**

Row	Model	Time-series $S_{TS}(ik)$	Time-series $\rho_1(ik)$	Cross-sectional $\sigma_{CS}(ik)$	Cross-sectional $S_{CS}(ik)$	Risk premia $E\left[R^{CU}\right]$	Risk premia $E\left[R^{bm}\right]$
	Data						
(1)		0.67	0.52	0.16	1.89	5.67	3.71
	Baseline						
(2)		0.65	0.58	0.11	1.19	5.13	3.77
	Baseline without fixed cost						
(3)		1.04	0.63	0.11	1.32	4.51	3.27
	Baseline without utilization						
(4)		-0.27	0.63	0.07	0.07	-	3.00
	Baseline without utilization and fixed cost						
(5)		0.53	0.70	0.08	0.58	-	2.65

The table reports model-implied population moments related to the time-series and cross-section of investment rates, as well as risk premia, under various calibrations of the model featuring no fixed disinvestment cost ($f = 0$). The table reports the time-series skewness ($S_{TS}(ik)$), and the first-order autocorrelation ($\rho(ik)$) of firm-level investment rates, as well as the cross-sectional dispersion ($\sigma_{CS}(ik)$) and skewness ($S_{CS}(ik)$) of investment rates. In addition, the table also reports the value premium ($E\left[R^{bm}\right]$) and investment premium ($E\left[R^{ik}\right]$) obtained by sorting the cross-section of model-implied returns association with each calibration on book-to-market ratios and investment rates, respectively. These risk premia are expressed as an annualized percentage. Each alternative calibration is identical to the benchmark calibration in all ways except for $f = 0$ in rows (3) to (5), and $\lambda \rightarrow \infty$ in rows (4) and (5). All moments are based on a simulations of 1,000 firms over 40,0000 periods (years). Finally, the top row of the table also reports the empirical counterpart of each moment.

analysis. The panel indicates that even though our baseline model only features a *single* priced source of risk, the model is capable of generating a utilization premium that is largely distinct from the value premium. Notably, the utilization premium ranges from about 2.5% to 3.5% per annum even after we condition on book-to-market ratios. While smaller than the unconditional spread, this conditional spread falls within the empirical confidence interval. When we set $f = 0$ and repeat this analysis, Panel B shows that the relative magnitude of the utilization premium falls by between 55% and 62% compared to Panel A. Moreover, a utilization premium of about 1.2% is below the lower bound of the empirical confidence interval. The stark difference between Panels A and B highlights the key role of non-convex disinvestment costs for separating the utilization premium from the value premium in the context of our benchmark model.

Sensitivity to fixed disinvestment costs. Table OA.4.4 reports a sensitivity analysis to show how the key investment-related moments and risk premia change when we switch both the fixed cost of disinvestment and variable capacity utilization on and off. There are two key takeaways.

First, row (3) shows the unconditional level of the utilization premium remains sizable even when $f = 0$. That is, without the fixed cost of disinvestment, the utilization premium falls by 0.62% (from 5.13% in row (2) to 4.51% in row (3)). This represents as 12% decline in the relative magnitude of the utilization premium compared to the baseline model. While this decline in the magnitude of the spread is quite sizable, the model-implied utilization spread in row (3) still falls within the empirical confidence interval of the spread in the data.

Second, an economy with flexible utilization can improve the model's fit to the data compared to an economy with fixed utilization, even if there is no fixed cost of disinvestment and $f = 0$. Comparing row (3) to row (5) shows that setting $f = 0$ but letting utilization vary helps the model to generate a sizable value premium, holding all else constant. The value premium in row (5) is 0.62% smaller than the value premium in row (3). This indicates that fixing utilization induces the value premium to fall, in relative terms, by 19% compared to the economy underlying row (3). Row (5) also shows that the cross-sectional skewness of investment is more than halved when utilization is fixed, and investment becomes too autocorrelated compared to the data.

OA.4.4 Sensitivity analysis of the model for risk premia

Below, we numerically illustrate the intuition for the utilization spread discussed in Section 3.2. We show the sensitivity of the spread to ingredients (1)–(3) of our model (the quadratic capital adjustment cost, fixed cost of disinvestment, and countercyclical market price of risk, respectively). We also show that the utilization spread is largely unaffected by perturbing the parameters governing the evolution of aggregate or idiosyncratic productivity. Table OA.4.5 presents these results, and reports the mean value-weighted of the utilization spread. The table also reports the mean and volatility of the equity risk premium in the model under each alternative calibration.

The results in rows (2) and (3) show that when the extent of the first friction, the quadratic capital adjustment costs, is perturbed, the magnitudes of the utilization spread changes. As this friction is increased in row (3), the magnitude of the utilization spread increases. With higher adjustment costs, firms can less readily alter the level of their capital stocks, and low utilization implies more underlying capital risk.

Row (4) considers an economy is which the second ingredient, the fixed cost of capital disinvestment, is removed but the remaining two frictions are held constant. The utilization spread still exists, although its magnitude is decreased by almost 1% per annum. The decrease in the utilization spread reflects how removing the fixed cost of disinvestment better allows firms to shed their capital stock instead of substituting disinvestment with temporary declines in utilization. However, the fact that the utilization spread remains sizable indicates that firms still cannot fully absorb productivity shocks into their capital stock. Next, rows (5), (6), and (7) consider the role of the third ingredient, the countercyclical market price of risk. In particular, row (6) illustrates how a more countercyclical market price of risk translates into a higher equity risk premium and volatility of aggregate market returns, as well as an increased utilization spread. This occurs because the asymmetry between good and bad aggregate productivity is widened. Row (7) indicates that both the equity risk premium and capacity utilization spread are severely diminished with an acyclical market price of risk.

Rows (8) and (9) show how the utilization spread changes as the persistence of aggregate productivity changes. The results in row (8) show that when aggregate productivity is less persistent, the magnitude and the volatility of the equity risk premium decrease. Similarly, the mean utilization premium falls slightly. The opposite patterns emerge in row (9), when the persistence of aggregate productivity increases.

Rows (10) and (11) of the table display how the utilization premium and equity risk premium both fall (rise) when aggregate productivity becomes less (more) volatile. The same patterns hold true for the volatility of the equity risk premium. Importantly, in rows (8)–(11), the model implied utilization premium changes by at most 0.45% in absolute value compared to the benchmark case, and falls within the empirical 95% confidence interval.

Finally, rows (12) to (15) display the sensitivity of key asset-pricing moments to perturbations in the parameters governing the dynamics of idiosyncratic productivity. The results indicate that when the persistence (ρ_z) or the volatility (σ_z) of idiosyncratic productivity increases, the capacity

Table OA.4.5: **Model-implied capacity utilization spread across alternative calibrations of the model**

Row	Model	$E\left[R^M\right]$	$\sigma\left(R^M\right)$	$E\left[R^{CU}\right]$
	Baseline			
(1)		5.39	20.88	5.13
	Different ϕ			
(2)	Low ($\phi = 1.40$)	5.39	20.85	5.06
(3)	High ($\phi = 1.60$)	5.39	20.92	5.19
	No fixed cost			
(4)		5.72	20.43	4.51
	Different γ_1			
(5)	Low ($\gamma_1 = -8.60$)	5.26	20.47	5.07
(6)	High ($\gamma_1 = -9.00$)	5.52	21.30	5.19
(7)	Acyclical ($\gamma_1 = 0$)	1.88	9.86	2.68
	Different ρ_x			
(8)	Low ($\rho_x = 0.899$)	4.57	17.56	4.71
(9)	High ($\rho_x = 0.945$)	6.48	24.99	5.50
	Different σ_x			
(10)	Low ($\sigma_x = 0.0137$)	4.95	19.75	4.91
(11)	High ($\sigma_x = 0.0143$)	5.84	22.02	5.34
	Different ρ_z			
(12)	Low ($\rho_z = 0.585$)	5.58	20.89	4.87
(13)	High ($\rho_z = 0.615$)	5.18	20.88	5.37
	Different σ_z			
(14)	Low ($\sigma_z = 0.2925$)	5.54	20.87	4.88
(15)	High ($\sigma_z = 0.3075$)	5.24	20.89	5.38

The table reports model-implied population moments under various calibrations. The table reports the equity premium ($E[R^M]$), the volatility of the market return ($\sigma\left(R^M\right)$), and the level of the capacity utilization spread ($E[R^{CU}]$). Each moment is reported as an annual percentage. and each alternative calibration is identical to the benchmark calibration in all ways except for altering the specified parameter of interest. The parameters altered are the fixed cost of disinvestment (f), the quadratic capital adjustment cost (ϕ), the cyclicality of the market price of risk (γ_1), the persistence of the aggregate productivity process (ρ_x), the volatility of the aggregate productivity process (σ_x), the persistence of the idiosyncratic productivity process (ρ_z), and the volatility of idiosyncratic productivity (σ_z). All moments are based on a simulations of 1,000 firms over 40,0000 periods (years).

utilization premium rises.

OA.4.5 Sectoral heterogeneity and the utilization premium

We design a simulation-based experiment to put an upper bound on how much ex-ante sectoral-level differences (e.g., difference in adjustment costs and the flexibility of utilization across different industries) may affect the utilization spread. Let x denote a parameter of interest, let x_0 denote the value of x in our benchmark calibration, and let N represent the number of firms in the economy. First, we solve the model twice: once when x is doubled to $x_U = 2x_0$, and once when x is halved to $x_D = 0.5x_0$. Next, we simulate $N/2$ firms implied by the model solved for each x_U and x_D, and combine these two simulations into one economy of N firms. These two simulations aim to capture *extreme* differences between industries in terms of parameter x. Finally, we sort the N firms on the basis of utilization in an identical fashion to our baseline model results. The difference between the model-implied utilization premium here versus our benchmark quantifies an upper bound on the effect that ex-ante heterogeneity in parameter x has on the premium.

Table OA.4.6: **Capacity utilization spread: sensitivity to ex-ante heterogeneity**

	Hetero. in ϕ		Hetero. in λ		Hetero. in δ_k		Hetero. in f	
Portfolio	$E\left[R^{CU}\right]$	β	$E\left[R^{CU}\right]$	β	$E\left[R^{CU}\right]$	β	$E\left[R^{CU}\right]$	β
Low (L)	9.82	1.21	9.60	1.19	9.71	1.26	9.81	1.21
Medium	6.99	1.10	6.99	1.09	6.15	1.10	7.01	1.12
High (H)	4.52	0.97	4.23	1.01	3.60	0.97	4.50	1.01
Spread (L-H)	5.30	0.23	5.37	0.18	6.11	0.29	5.31	0.21

The Table show the model-implied capacity utilization spread when firms in the economy show ex-ante heterogeneity in some parameter of interest x. The parameter of interest x is either ϕ, the quadratic capital adjustment costs, λ, the elasticity of depreciation to utilization, δ_k, the average depreciation rate, or f, the fixed cost of disinvestment. For each parameter, we follow the simulation procedure described in Section OA.4.5 and set N to 1,000. Each simulation encompasses 40,000 periods, and we sort all firms on the basis of utilization in an identical fashion to Table 8. A firm is sorted into the high (low) utilization portfolio if its level of capacity utilization is above (below) the 90^{th} (10^{th}) percentile of the cross-sectional distribution of capacity utilization rates in the previous period. To compute β in the model, the volatility of market returns in the model is scaled to match the volatility of market returns in the data.

We consider heterogeneity in four parameters: the depreciation rate, δ_k, elasticity of depreciation to utilization, λ, convex adjustment cost, ϕ, and fixed disinvestment cost, f. The results are reported in Table OA.4.6. In comparison to Panel B of Table 8, heterogeneity in ϕ, λ, δ_k and f add up to 0.06%, 0.13%, 0.87% and 0.20% per annum, respectively, to the premium. Thus, ex-ante sectoral heterogeneity only implies a marginal amplification of the model-implied utilization spread.

OA.4.6 Numerical model solution

We solve the model numerically using value function iteration. The value function and the optimal policies implied by the firm's maximization problem in equation (12) are solved on a grid in a discrete state space. The grid for capital stock, K, features 501 grid points, with the endpoints of the grid chosen to be nonbinding. The aggregate productivity process, x, and the idiosyncratic productivity process, z, are each driven by an independent and identically distributed normal distribution. While each of these state variables has continuous support in the model, each variable needs to be transformed into a finite number of states to implement the numerical solution algorithm. We use the method of Tauchen and Hussey (1991) to discretized the z process into 11 states. Because the method of Tauchen and Hussey (1991) does not work well for persistent processes, namely those with a persistence parameter greater than 0.90, we use the method of Rouwenhorst (1995) to discretize x into 5 states. Once the discrete state space has been constructed, conditional expectations are computed using matrix multiplication and the firm's maximization problem is solved using a global search routine. All results are robust to choosing finer grids.

OA.5 Model extensions

OA.5.1 A model with investment efficiency shocks

We extend our baseline single aggregate shock model from Section 2 to a model with two aggregate shocks in order to generate distinct economic mechanisms that drive the utilization premium and other investment-based spreads. Specifically, we introduce an aggregate "investment efficiency" shock to our model that captures the spirit of the capital-embodied shocks in Papanikolaou (2011), Garlappi and Song (2017), and many other studies.

Introducing these investment efficiency shocks to the model leads to periods in which the marginal efficiency of investment is higher, meaning that investment can be transformed into capital in a more efficient manner as in Papanikolaou (2011) and Justiniano et al. (2011). This increased investment efficiency can arise as a result of an improvement in the quality of investment goods or through lower frictions to the process through which investment goods are transformed into productive capital. Justiniano et al. (2011) argue that these investment frictions can also be related to the efficiency of the financial sector (i.e., a reduced-form financial accelerator as in Carlstrom and Fuerst (1997)).

Consequently, investment efficiency shocks reduce the installation cost of new vintages of capital (i.e., lower adjustment costs). This means that investment efficiency shocks will affect firms' investment policies but will not affect their choices of utilization rates. Thus, introducing these investment efficiency shocks creates a separation between the investment-related spreads (e.g., the value premium), which are driven by these investment efficiency shocks, and the utilization premium, which remains primarily driven by aggregate productivity shocks.

Model setup. We make two main changes to the setup of our baseline model described in Section 2.1. First, we alter the convex capital adjustment cost function to

$$\frac{\phi - s_t}{2}\left(\frac{I_{i,t}}{K_{i,t}} - \delta\left(u_{i,t}\right)\right)^2 K_{i,t}, \tag{OA.5.1}$$

where s_t is the degree of aggregate investment efficiency. Note that a positive s_t shock decreases the amount of adjustment costs. We assume that $\phi >> \sigma_s$, such that $\phi - s_t$ is always positive. Alternatively, we could posit that adjustment costs are proportional to $\phi/\exp(s_t)$ (see, e.g., Belo et al. (2014)). However, we opt for the former linear specification to avoid any asymmetric effects between positive versus negative s_t shocks due to convexity. The process for s_t follows

$$s_{t+1} = \rho_s s_t + \varepsilon^s_{t+1}, \text{ where } \varepsilon^s_{t+1} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_s^2\right). \tag{OA.5.2}$$

Second, we also alter the SDF to feature this priced "investment efficiency" shock

$$\ln\left(M_{t+1}\right) = \ln\left(\beta\right) - \gamma_t \varepsilon^x_{t+1} - SIGN_S \cdot \gamma_s \varepsilon^s_{t+1} - \frac{1}{2}\gamma_t^2\sigma_x^2 - \frac{1}{2}SIGN_S^2\gamma_s^2\sigma_s^2. \tag{OA.5.3}$$

Here, γ_t follows identical dynamics to those described in Section 2.1 (recall equation (11)), γ_s is the absolute value of the market price of investment efficiency risk, and $SIGN_S \in \{-1, +1\}$ determines whether investment efficiency shocks carry a negative or a positive market price of risk.

Optimality conditions and model intuition. To see why the utilization and investment-related spreads are distinct in this two-shock model, it is useful to examine the first-order conditions associated with a firm's choice of investment and utilization. The first-order condition for investment is given by the Euler equation

$$q_t = \mathbb{E}_t\left[M_{t,t+1}\left(\frac{\partial Y_{i,t+1}}{\partial K_{i,t+1}} - \frac{I_{i,t+1}}{K_{i,t+1}} - \frac{\phi - s_{t+1}}{2}\left(\frac{I_{i,t+1}}{K_{i,t+1}} - \delta_{i,t}\right)^2 + q_{t+1}\left(1 - \delta_{i,t+1} + \frac{I_{i,t+1}}{K_{i,t+1}}\right)\right)\right], \tag{OA.5.4}$$

where

$$q_t = 1 + (\phi - s_t)\left[\frac{I_{i,t}}{K_{i,t}} - \delta_{i,t}\right]. \tag{OA.5.5}$$

The first-order condition for utilization is

$$0 = \frac{\partial Y_{i,t}}{\partial u_{i,t}} + (\phi - s_t)\left[\frac{I_{i,t}}{K_{i,t}} - \delta_{i,t}\right]\frac{\partial \delta\left(u_{i,t}\right))}{\partial u_{i,t}} - q_t\frac{\partial \delta\left(u_{i,t}\right))}{\partial u_{i,t}}K_{i,t}, \tag{OA.5.6}$$

which can be simplified to

$$\frac{\partial Y_{i,t}}{\partial u_{i,t}} = \frac{\partial \delta\left(u_{i,t}\right))}{\partial u_{i,t}}K_{i,t}. \tag{OA.5.7}$$

Thus, equation (OA.5.7) shows that the firm's choice of utilization is independent of investment efficiency as both $\frac{\partial Y_{i,t}}{\partial u_{i,t}}$ and $\frac{\partial \delta(u_{i,t}))}{\partial u_{i,t}}$ are independent of s_t and K_t is predetermined.

In contrast, equation (OA.5.4) highlights that the firm's choice of investment critically and directly depends on the degree of investment efficiency s_{t+1}, and indirectly depends on s_t since $\rho_s > 0$ (i.e., shocks to investment efficiency are persistent). When investment efficiency *rises* (that is, when s_t increases), the cost of installing new capital *falls*. Thus, a positive investment efficiency shock can be interpreted as a technological improvement that induces the firm to invest (disinvest) more than the firm otherwise would when subject to a positive (negative) productivity shock.

Because the choice of utilization does not depend on shocks to s_t, the expected returns of the utilization-sorted portfolios remain tightly related to their exposures to aggregate productivity shocks. Thus, the utilization premium in this two-shock model will arise for the same reasons that we obtain a utilization premium in our single-shock model (see Section 3.1). In contrast, because a firm's investment policy depends on the degree of investment efficiency s_t, the expected returns of portfolios sorted on the basis of firms' investment-related policies (e.g., the value premium) will depend more heavily on the degree to which these portfolios are exposed to the investment efficiency shocks. To see this, we can fix a firm's investment policy and write the exposure of the ex-dividend firm value to s_t shocks as

$$\beta_{i,s}^{ex-div} = \frac{\partial q_{i,t}}{\partial s_t} = -\frac{I_{i,t}}{K_{i,t}}. \tag{OA.5.8}$$

Equation (OA.5.8) show that a firm with a positive investment rate that is expanding its capital stock has a *negative* exposure to the investment efficiency shock s_t. Intuitively, a firm that is expanding its capital stock has a relatively larger growth option value in the presence of capital adjustment costs because shocks are absorbed in firm value rather than investment. Consequently, a positive shock to s_t lowers capital adjustment costs (recall equation OA.5.1) and, in turn, induces a negative valuation effect by decreasing the growth option value of the firm (in fact, there is no growth option value in the limit where $\phi = s_t$ and $q_{i,t} = 1$). The opposite logic applies to a firm that is disinvesting, whose exposure to s_t shocks is negative. Moreover, a positive s_t shock makes an investing firm have a higher I/K and a disinvesting firm have a lower I/K, all else equal. This interaction effect enhances the exposure to investment efficiency shocks (in absolute value).

As growth (value) firms invest (disinvest) more heavily, they have a negative (positive) exposure to the investment efficiency shocks. This creates a strong dependence of the value premium (and other investment-related spreads) on shocks to s_t.

Calibrating the two-shock model. Adding the "investment efficiency" shock introduces three new parameters to the model. In the spirit of Li (2018), who proposes a unified explanation for the value premium and the moment effect, we set the volatility of these investment efficiency shocks to 7.25% per annum and the autocorrelation of these shocks to 0.90. We also set $SIGN_s$ to +1, indicating that investment efficiency shocks carry a positive market price of risk. A positive market price of risk for investment efficiency shocks is not only consistent with Li (2018), but also mirrors an important result from Garlappi and Song (2017). Namely, investment-specific technology (IST) shocks carry a *negative* price of risk if utilization is fixed, but a *positive* price of risk if utilization is sufficiently flexible.[36] The latter case fits our framework with time-varying utilization. We calibrate the magnitude of the market price of this investment efficiency shock (γ_s) to match the book-to-market spread in the data. Since our primary objective are to (i) establish an economically sizable utilization premium, and (ii) establish that the utilization premium is independent from the value premium, we emphasize that neither this choice nor the choice of any other calibration parameter targets the utilization spread directly. Thus, the utilization premium remains a testable prediction of our two-shock model.

In the two-shock model we opt to set the fixed cost of disinvestment (f) to zero. Recall that in our single-shock model, part of the difference between a firm's choice of utilization and investment arises because of this fixed cost of disinvestment. This cost induces periods in which firms are "waiting to see" if aggregate productivity improves before exercising their costly options to sell capital. In the two-shock framework, the investment efficiency shock introduces a sufficient degree of heterogeneity between firms' investment and utilization policies. By setting the fixed cost of disinvestment to zero, we narrow the economic distinction between the utilization and the value premia to (i) hinge on the investment efficiency shocks (rather than on real option nature of investment), and (ii) limit the degrees of freedom in the model. Because the model no longer features a real option, the model is smooth and differentiable with respect to the state and control variables. We therefore solve the model using a fourth order perturbation method in Dynare++.

We report the remaining parameters used to calibrate the two-shock model in Table OA.5.1. These remaining parameters are either identical to, or in the close vicinity of, the parameters used in our baseline one-shock model that are reported in Table 6.

Table OA.5.2 reports the key moments implied by the two-shock model when the investment efficiency shocks carry either a positive or a negative market price of risk. We compare these moments to the same moments implied by our baseline one-shock model and the data. The table shows that the two-shock model provides a close match to the data along many dimensions related to firm-level investment rates, aggregate utilization rates, aggregate sales growth rates, and aggregate asset-pricing data. In particular, and compared to the single-shock model, the two-shock model yields an improvement in the model's fit for the second-order autocorrelation and the inter-decile

[36]Garlappi and Song (2017) show that a positive IST shock increases the relative productivity of the investment sector compared to the consumption sector, and hence diverts labor away from consumption firms and into investment firms. This will, on the one hand, drop the supply of labor in the consumption sector and reduce the amount of the consumption good that is available, increasing the marginal utility of households. On the other hand, if utilization is flexible, then the consumption sector can offset the decline in labor by increasing the utilization of capital (which becomes relatively cheaper), thereby producing more of the consumption good. This decreases the marginal utility of the household. If capital utilization is sufficiently flexible, the second effect more than offsets the negative effect of labor diverting to the investment sector and results in a positive price of risk.

Table OA.5.1: **A model with aggregate investment efficiency shocks: calibration**

Symbol	Description	Value
Stochastic processes		
ρ_x	Persistence of aggregate productivity	0.922
σ_x	Conditional volatility of aggregate productivity	0.014
ρ_z	Persistence of idiosyncratic productivity	0.700
σ_z	Conditional volatility of idiosyncratic productivity	0.150
ρ_s	Persistence of capital efficiency	0.900
σ_s	Conditional volatility of capital efficiency	0.073
β	Time discount factor	0.988
γ_0	Constant price of aggregate productivity risk	2.175
γ_1	Time-varying price of aggregate productivity risk	-89.250
γ_s	Time-varying price of capital efficiency risk	2.000
Technology		
α_k	Capital share	0.333
α_l	Labor share	0.667
θ	Returns to scale of production	0.900
δ_k	Fixed capital depreciation rate	0.080
λ	Elasticity of marginal depreciation	2.350
ω	Sensitivity of wages to aggregate productivity	0.200
ϕ	Adjustment cost parameter	0.653

The table reports the model parameters for the model extension that features two aggregate priced sources of risk: aggregate productivity shocks and investment efficiency shocks.

range of investment, and the volatility of market returns. Most moments are largely identical when $SIGN_s = 1$ compared to when $SIGN_s = -1$, although the equity premium is about 0.2% per annum larger in the former case (which is our benchmark calibration of this two-shock model).

Model-implied spreads. As a first step, we check whether the model produces both a sizable utilization premium and a sizable value premium. Using model-simulated data, we sort the cross-section of firms into three univariate portfolios on the basis of either capacity utilization rates or book-to-market ratios. Here, our portfolio formation procedure follows the benchmark approach used in Section 1.2 of the main text. That is, in each period t, we sort firms into portfolios on the basis of the cross-sectional distribution of the characteristic of interest in period $t-1$. We then hold each of these portfolios for one period, at which point in time all portfolios are rebalanced.

Table OA.5.3 reports the results of these univariate portfolio sorts in the two-shock model, and shows that the model produces both a sizable utilization premium and a sizable book-to-market spread. The utilization (book-to-market) spread is 6.40% (3.82%) per annum when $SIGN_s = 1$, while the utilization (book-to-market) spread is 6.52% (3.05%) per annum when $SIGN_s = -1$.

Two differences between the case of $SIGN_s = 1$ (our benchmark) and the case of $SIGN_s = -1$ are worth noting. First, the value premium is sensitive to the price of risk of the investment efficiency shocks. If these shocks carry a positive price of risk ($SIGN_s = 1$), then the value premium matches the data. In contrast, if these shocks carry a negative price of risk ($SIGN_s = -1$), then the premium shrinks by 20% (in relative terms) from 3.82% per annum to 3.05% per annum. This findings is consistent with the model's optimality condition for investment in equation (OA.5.4) and the ensuing discussion that shows that growth (value) firms have a negative (positive) exposure to s_t shocks. When $SIGN_s = 1$, growth firms hedge the investment efficiency risk and command a lower risk premium.

Table OA.5.2: **A model with aggregate investment efficiency shocks: model-implied moments versus data**

Variable	Data	Benchmark Model	$SIGN_s = 1$	$SIGN_s = -1$
Vol. of investment rate (time-series)	0.14	0.14	0.14	0.15
Vol. of investment rate (cross-sectional)	0.16	0.11	0.12	0.12
AC(1) of investment rate	0.52	0.58	0.56	0.57
AC(2) of investment rate	0.26	0.38	0.28	0.29
Inter-decile range of investment rate	0.32	0.22	0.30	0.30
Vol. of aggregate capacity utilization level	4.09	4.13	4.02	4.19
AC(1) of aggregate capacity utilization level	0.65	0.92	0.93	0.94
Vol. of aggregate sales growth	6.58	7.51	7.03	7.12
AC(1) of aggregate sales growth	0.46	0.40	0.40	0.41
Real risk-free rate	1.19	1.21	1.15	1.15
Excess market return	6.28	5.39	4.05	3.84
Volatility of excess market return	17.20	20.88	17.11	17.83

The table shows model-implied moments alongside their empirical counterparts, computed using data from 1967 to 2015. The column "Benchmark model" reports the moments implied by our single-shock benchmark model from Section 2 that does not include aggregate investment efficiency shocks. In contrast, the column $SIGN_s = 1$ ($SIGN_s = -1$) refers to moments implied by the two-shock model in which the investment efficiency shocks carry a positive (negative) market price of risk.

Second, the utilization premium is materially unchanged when the market price of risk of the s_t shocks turns negative. Specifically, the utilization premium with $SIGN_s = 1$ is only 1% smaller (in relative terms) that the utilization premium when $SIGN_s = -1$. This finding is once again consistent with the model's optimality conditions that show that the choice of utilization is independent from the investment efficiency shocks (recall equation OA.5.7). The fact changing $SIGN_s$ to a negative value has a considerable effect on the value premium but only a negligible effect on the utilization premium provides an early indication that the two spreads are relatively distinct. Comparing the drivers of the two spreads also shows that the utilization premium is more closely tied to aggregate productivity risk, whereas the value premium (and other investment-related spreads) is more closely tied to investment efficiency risk. We verify this intuition below.

Distinction between the utilization premium and investment-related spreads. To demonstrate the distinction between the value and utilization premia we begin by implementing portfolio double sorts using model-simulated returns. Table OA.5.4 confirms that we continue to obtain a sizable utilization premium even if we control for book-to-market ratios. We establish this fact in three steps. First, in each period t, we control for book-to-market ratios by sorting the cross-section of model-implied firms into three portfolios on the basis of their book-to-market ratios in period $t-1$. Here, we use the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of book-to-market ratios as portfolio breakpoints. Next, *within* each book-to-market sorted portfolio, we further sort the cross-section of firms on the basis of their capacity utilization rates in period $t-1$ using the same portfolio breakpoints. Finally, each portfolio is held for one period, at which point in time all portfolios are rebalanced.

The table shows that the capacity utilization premium remains almost unchanged at about 6% to 7% per annum after controlling for differences in book-to-market ratios across firms. This

Table OA.5.3: **A model with aggregate investment efficiency shocks: univariate spreads**

Portfolio	Panel A: Utilization premium $SIGN_s = 1$	Panel A: Utilization premium $SIGN_s = -1$	Panel B: B/M spread $SIGN_s = 1$	Panel B: B/M spread $SIGN_s = -1$
Low (L)	7.80	7.60	3.35	3.51
Medium	3.99	3.84	3.53	3.56
High (H)	1.40	1.08	7.17	6.56
Spread	6.40	6.52	3.82	3.05

The table reports the average annual value-weighted returns of portfolios sorted on capacity utilization rates (Panel A) and book-to-market ratios (Panel B), as implied by the model with investment efficiency shocks. As in the empirical analysis, a firm is sorted into the high (low) utilization portfolio if its level of capacity utilization is above (below) the 90^{th} (10^{th}) percentile of the cross-sectional distribution of capacity utilization rates in the previous period. Similarly, a firm is sorted into the high (low) book-to-market portfolio if book-to-market ratio is above (below) the 90^{th} (10^{th}) percentile of the cross-sectional distribution of book-to-market ratios in the previous period. The columns $SIGN_s = 1$ ($SIGN_s = -1$) refer to the cases in which the market price of risk of the aggregate investment efficiency shocks is positive (negative).

distinction between the value and utilization spreads in this two-shock model is quantitatively more pronounced than the same distinction between these spreads in the single-shock model. This is because the sort on capacity utilization in the two-shock model is akin to a sort on each firm's exposure to aggregate productivity shocks, while the sort on book-to-market ratios is akin to a sort on each firm's exposure to investment efficiency shocks.

Table OA.5.4: **A model with aggregate investment efficiency shocks: double sorts**

	Utilization premium within B/M portfolios			
	Low CU	Medium CU	High CU	Spread (L-H)
Low B/M	6.61	3.17	0.33	6.28
Medium B/M	7.35	3.27	0.18	7.19
High B/M	10.48	5.72	2.78	7.70

The table shows the returns obtained from a conditional portfolio double sort procedure implied by the model with investment efficiency shocks. Specifically, the control variable, or the first dimension sorting variable, is a firm's book-to-market ratio, and the second dimension sorting variable is the capacity utilization rate. The portfolios are constructed as follows. First, in each period t, firms are first sorted into three portfolios on the basis of the cross-sectional distribution of the control variable in period $t-1$. Here, we use the 20^{th} and 80^{th} percentiles of the cross-sectional distribution as portfolio breakpoints. Next, within each portfolio, firms are further sorted into three additional portfolios on the basis of the second stage sorting variable in period $t-1$. Here, we once again use the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of this variable as portfolio breakpoints. This procedure produces nine portfolios that are held for one period, at which point in time all portfolios are rebalanced.

Projection evidence. To further highlight the distinction between the two spreads we project the utilization premium onto the value premium via the following time-series regression

$$\text{Utilization premium}_t = \alpha + \beta \text{B/M spread}_t + \varepsilon_t. \qquad \text{(OA.5.9)}$$

If the utilization premium is merely a linear transformation of the value premium (i.e., if both spreads are linearly driven by the same fundamentals), then the constant in this regression will be zero. In contrast, the results from a population-sample simulation show that α in the regression above is 4.90% per annum, an economically sizable quantity. This alpha is consistent with the empirical counterpart of this estimate that is 5.11% per annum (see Panel A of Table OA.3.20).

OA.5.2 A model with time-varying firm-level markups

Firm-level TFP is a combination of (i) technology, (ii) markups, and (iii) utilization. Our benchmark model is somewhat silent on the underpinnings of this decomposition. However, the empirical analyses in Section OA.3.6 suggest that the distinction between the TFP premium and the utilization premium originates from the fact that the former spread reflects a sort on the first two components of TFP (technology and markups) while the latter spread reflects a sort on the third component of TFP (utilization).

To mirror this type of result in the model, we introduce markups to the two-shock model described in Section OA.5.1. Because the model is a partial equilibrium, we cannot directly model monopolistic competition to generate endogenous markups. Therefore, we assume that markups fluctuate exogenously. These fluctuations may reflect time-varying degrees of price adjustment (menu) costs or substitutability between the firm's product and competitors products. Specifically, we introduce these time-varying markups by replacing equation (4) with

$$Y_{i,t} = \exp(x_t + z_{i,t}) \left[u_{i,t}^{(\theta\alpha_k \cdot \xi)} + \mu_{i,t}^{\xi} \right]^{\frac{1}{\xi}} K_{i,t}^{\theta\alpha_K} L_{i,t}^{\theta\alpha_L}, \tag{OA.5.10}$$

where the markup process $\mu_{i,t}$ follows the dynamics

$$\mu_{i,t+1} = \rho_\mu \mu_{i,t} + \varepsilon_{i,t+1}^{\mu}, \text{ where } \varepsilon_{t+1}^{\mu} \overset{i.i.d.}{\sim} \mathcal{N}\left(0, \sigma_\mu^2\right). \tag{OA.5.11}$$

To keep the setup generic, the augmented production function in equation (OA.5.10) introduces a constant elasticity of substitution (CES) aggregator in which ξ represents the elasticity of substitution between utilization and markups for output. We model the augmented output function in this way for two reasons. First, to emphasize that utilization is a factor of production that is related to the *intensity* of capital usage while markups simply induce a wedge between price and marginal cost, we raise $u_{i,t}$ but not $\mu_{i,t}$ to the power of α_k. This means that markups exist even if the firm relies entirely on labor to produce output (i.e., if the share of capital in output $\alpha_k \to 0$). Second, setting σ_μ^2 in equation (OA.5.11) to zero reduces the augmented output function in equation (OA.5.10) to our baseline production function in equation (4). Thus, this augmented production function nests our baseline production function.

Since our goal is to *numerically* demonstrate that markups further boost the independence between the utilization premium and the productivity premium, we set the dynamics of the firm-level markup process $\mu_{i,t}$ to follow an autoregressive progress with $\sigma_\mu = 0.012$, and $\rho_\mu = 0.48$. This suggests that markups are moderately persistent (in line with Basu et al. (2006)), but not very volatile (i.e., the volatility of firm-level technology shocks is much larger the volatility of markup shocks). To make the quantitative distinction the starkest, we let $\xi \to 1$ so that the output function is the simple linear aggregation of the effects of utilization and markups. In particular, when $\xi \to 1$, the first-order condition for utilization in unchanged compared to equation (OA.5.7). Since $\partial Y_{i,t}/\partial u_{i,t}$ is independent of $\mu_{i,t}$, the firm's choice of utilization does not depend on the current markup. However, as $\partial Y_{i,t}/\partial K_{i,t}$ depends on $\mu_{i,t}$, firms' investment policies depend on markups and technology via the Euler equation for investment. We solve this augmented model using the same method as that described in Section OA.5.1.

Estimating firm-level TFP. Before evaluating the degree to which the utilization and productivity premia are related, we estimate firm-level TFP in the model. We do this in a similar fashion to the empirical approach of Imrohoroglu and Tuzel (2014). While Imrohoroglu and Tuzel (2014) apply the Olley and Pakes (1996) estimator to elicit the firm-specific component of productivity, we employ a simplified OLS-based estimator since our production-based model is stripped of many of the underlying identification issues that the Olley and Pakes (1996) estimator is designed to overcome. For instance, there are no entry and exit decisions in our setup so we do not need to account for the probability that a firm survives until the next period. As such, we elicit firm-level TFP as the residual from the following panel regression estimated using model-simulated data:

$$\ln\left(Y_{i,t}\right)=\delta_t+\beta_K \ln\left(K_{i,t}\right)+\beta_L \ln\left(L_{i,t}\right)+\epsilon_{i,t}. \tag{OA.5.12}$$

Here, $Y_{i,t}$ represents the output of firm i, $K_{i,t}$ represents the firm's (predetermined) capital stock, and $L_{i,t}$ represents the firm's labor at time t, respectively. δ_t is a time fixed effect that subsume the effects of aggregate productivity shocks that hit all firms at the same time. By mapping equation (OA.5.10) to projection (OA.5.12), the residual $\epsilon_{i,t}$ captures technology ($z_{i,t}$), markup ($\mu_{i,t}$), and endogenous utilization ($u_{i,t}$).

Distinction between the utilization and productivity premia. As described above, a firm's choice of utilization is (i) independent of the degree of investment efficiency s_t (whereas the firm's investment decision depends on these shocks) and (ii) independent of markups. This creates two differences between the utilization and TFP premia. First, even without markups (i.e., if $\sigma_\mu = 0$), low TFP firms disinvest more and consequently face larger exposure to s_t shocks. In contrast, the choice of utilization is independent of these s_t shocks. Second, with markups (i.e., if $\sigma_\mu > 0$), firm-level TFP is a combination of $z_{i,t}$ and $\mu_{i,t}$, while utilization depends only on $z_{i,t}$. This creates another separation: controlling for TFP, the utilization premium in the model is driven by technology and capital, while controlling for utilization, the TFP premium is mostly driven by markups. As each state variable affects the conditional exposure of firms to priced shocks, the model produces a sufficient amount of dispersion in the conditional betas of the TFP-sorted versus the utilization-sorted portfolios. In the extreme case, a low utilization firm can also be a high markup firm, rendering the firm to have high TFP in spite of having a low utilization rate. Thus, a low utilization firm is not necessarily synonymous with a low productivity firm.

Before showing the distinction between the TFP and utilization premia, we first verify that this model produces a sizable TFP spread. To do this, we construct the TFP premium by sorting the cross-section of model-implied firms on their values of $\epsilon_{i,t}$ from equation (OA.5.12), which captures the firm-level productivity in the model. In line with the results in Imrohoroglu and Tuzel (2014), we find that the low (high) productivity portfolio earns an average annual return of 5.87% (0.11%), resulting in a model-implied productivity premium of 5.76% per annum.

To show that the utilization premium is quantitatively distinct from the productivity premium we first conduct a portfolio double sort using model-simulated returns. The first stage of this procedure sorts the cross-section of model-implied firms at each time t into three portfolios on the basis of firm-level productivity at time $t-1$. Here, we use the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of firm-level productivity to form these portfolios. The second stage of this procedure then constructs the utilization premium *within* each of the productivity-sorted portfolios using the same percentiles of the cross-sectional distribution of utilization rates at time $t-1$. We hold each portfolio for one period, at which point in time all portfolios are rebalanced.

The results of this double sort procedure are reported in Table OA.5.5. Panel A of the table implements these double sorts in the model that *excludes* firm-specific markups (i.e., using the model from Section OA.5.1 in which $\sigma_\mu = 0$). We obtain an economically small but distinct utilization premium *within* each productivity-sorted portfolio. That is, controlling for firm-level productivity in the model without markups results in a utilization premium of about 1.5% per annum. As mentioned above, the two-shock model in Section OA.5.1 produces a utilization premium that is distinct from spreads that are related to investment policies, including TFP.

Table OA.5.5: **Double sorts controlling for TFP: A model with and without markups**

	Low CU	Medium CU	High CU	Spread (L-H)
Panel A: Including markups				
Low TFP	10.27	8.01	6.16	4.11
Medium TFP	5.60	3.44	1.44	4.16
High TFP	2.72	0.77	-1.04	3.76
Panel B: Excluding markups				
	Low CU	Medium CU	High CU	Spread (L-H)
Low TFP	1.31	0.21	-0.58	1.88
Medium TFP	4.65	3.49	2.96	1.69
High TFP	9.41	8.58	9.12	0.28

The table reports the results from a double sort procedure within a model that features investment efficiency shocks, and either includes time-varying markups (Panel A) or excludes these markups (Panel B). In each panel, we first sort the cross-section of model-implied firms at each time t into three portfolios on the basis of firm-level productivity at time $t-1$, estimated using equation (OA.5.12). Here, we use the 20^{th} and 80^{th} percentiles of the cross-sectional distribution of firm-level productivity (TFP) to form these portfolios. Next, we construct the utilization premium *within* each of the three productivity-sorted portfolios by further sorting the cross-section of firms into three portfolios on the basis of their capacity utilization rates at time $t-1$. Here, we one again use the same percentiles of the cross-sectional distribution of utilization rates as portfolio breakpoints. This procedure produces nine portfolios that are held for one period, at which point in time all portfolios are rebalanced.

Moreover, Panel B of Table OA.5.5 shows that adding firm-specific markups to the model significantly boosts the distinction between the utilization and TFP premia. Including markups in the model results in a utilization premium of around 4% per annum even when we condition on firm-level productivity. This highlights a significant degree of distinction between the two spreads.

Lastly, in this extended model, the unconditional utilization premium is above 6% (similar to the former subsection). Projecting the model-implied utilization premium on the model-implied productivity premium through a time-series regression similar to equation (OA.5.9) results in an alpha that is very close to the unconditional spread of about 6%. This model-implied alpha is in line with the empirical estimate of the same alpha reported in Panel A of Table OA.3.20. This time-series approach once again affirms the takeaways from the portfolio double sorts and Fama and MacBeth (1973) regressions: the utilization and productivity premia are distinct spreads.

OA.5.3 A model with depreciation shocks

We pursue an extension of our benchmark model motivated by the empirical evidence in Section OA.7.1. Specifically, while our baseline model assumes that depreciation rates correlate positively with utilization rates (recall equation (8)), Table OA.7.1 shows that these two quantities comove together, but not *perfectly*. As a result, we modify equation (8) to also feature an exogenous shock to the depreciation rate. This implies that a firm's depreciation rate becomes a combination of (i) its choice of utilization rate, and (ii) a stochastic shock. This augmented depreciation function is

$$\delta(u_{i,t}) = \delta_k + \delta_u \left[\frac{u_{i,t}^{1+\lambda} - 1}{1+\lambda} \right] + d_t, \quad \text{where} \quad d_{t+1} = \rho_d d_t + \sigma_d \varepsilon_{t+1}^{\delta}, \tag{OA.5.13}$$

and ε_t^{δ} is a standard normal i.i.d. shock to depreciation. Since depreciation shocks can be correlated with aggregate productivity, we consider two extreme cases: shocks that are both perfectly positively and perfectly negatively correlated with aggregate productivity.

We calibrate ρ_d and σ_d such that the largest value of $|d_t|$ in a discrete state space with a truncated support (i.e., two standard deviations from zero) causes the depreciation rate to change by up to 2.58%. As the unconditional depreciation rate in the model (δ_k) is 8%, the impact of the largest depreciation shock causes a firm's depreciation rate to change by roughly 33% compared to its unconditional mean. Given the large magnitude of these calibrated depreciation shocks, this exercises places an upper bound on the effect that these shocks can have on the utilization premium.

Table OA.5.6 shows that depreciation shocks that are perfectly positively (negatively) correlated with aggregate productivity decrease (increase) the model-implied utilization spread by 0.60% per annum. In either the case the model-implied spread, which is based on industry-level portfolio returns, falls within the confidence interval of the utilization premium in the data. Since the true correlation between depreciation and utilization rates is unknown, but must necessarily fall within the range of correlations we consider (i.e., $[-1, +1]$), exogenous shocks to depreciation rates do not materially impact the magnitude of the utilization premium.

Table OA.5.6: **Capacity utilization spread: sensitivity to depreciation shocks**

	Positively correlated $\varepsilon_{i,t}^{\delta}$		Negatively correlated $\varepsilon_{i,t}^{\delta}$	
Portfolio	$E\left[R^{CU}\right]$	β	$E\left[R^{CU}\right]$	β
Low (L)	7.70	1.06	10.21	1.64
Medium	5.94	1.01	7.79	1.51
High (H)	4.38	0.96	5.74	1.38
Spread (L-H)	3.32	0.09	4.46	0.26

The table reports the average model-implied annual value-weighted returns of portfolios sorted on capacity utilization, as well as the exposure of each utilization portfolio to market returns (β), at the industry level. As in the empirical analysis, an industry is sorted into the high (low) utilization portfolio if its level of capacity utilization is above (below) the 90^{th} (10^{th}) percentile of the cross-sectional distribution of capacity utilization rates in the previous period. Here, the model economy is identical to the benchmark case and calibration with one exception: the depreciation rate of each firm is subject to an exogenous shock, as represented by equation (OA.5.13). In the left (right) portion of the table these depreciation rate shocks are perfectly positively (negative) correlated with the aggregate productivity shocks. Industry-level returns are simulated using the procedure described in Section 3.1. Population moments are obtained from one simulation of 50 industries for 40,000 periods (years).

OA.6 Model implications for macro-finance modeling

The implications of flexible utilization for asset prices span beyond the utilization premium. We highlight the roles of flexible utilization for jointly targeting cross-sectional risk premia and investment moments in the presence of real options while relying on a parsimonious adjustment cost specification. We first show the failures of the model without utilization to target asset-pricing and production moments. We then explain how utilization provides a solution to these misses. We demonstrate that flexible utilization permits us to target key moments with a lower degree of adjustment costs vis-à-vis a model with fixed utilization.

OA.6.1 A fixed utilization model: the failures

The benchmark model's success in jointly fitting (i) the volatility and skewness of investment, both across time and across firms, and (ii) risk premia, crucially hinges on variable capacity utilization rates. To illustrate this point, row (1) of Table 9 shows model-implied moments in an economy without flexible utilization (i.e., $\lambda \to \infty$).

With fixed utilization, the distribution of investment rates exhibits far less variability and asymmetry compared to the data both in the time-series and the cross-section. The time-series skewness of firm-level investment turns to -0.27, at odds with its empirical sign and magnitude of 0.67. Investment's time-series volatility drops to only 11%, and its autocorrelation becomes slightly too high. Cross-sectional moments also become severely distorted. The dispersion of investment rates is about a half of its empirical counterpart (7% in the model versus 16% in the data). The cross-sectional skewness of investment rates is merely 0.07 in the model, whereas it is much higher in the data (about 1.9). The model-implied value and investment spreads are also about 1% per annum smaller in this model than the data.

The fixed utilization model fails to capture the aforementioned moments since (1) the fixed adjustment cost makes disinvestment a real option, and (2) without flexible utilization, a firm's *only* way to respond to a negative productivity shock is by exercising this option. As discussed in Section 2.2, if a drop in productivity at time t is not extremely severe, then a "wait and see" effect tends to dominate. Thus, declines in productivity typically lead to periods of investment-policy inaction in which many firms do not alter their capital stock. Each waiting firm j sets its investment rate, $i_{j,\tau}$, to the constant depreciation rate of δ_k for all $\tau \in [t, t+\hat{t})$, where $\hat{t}$ is the ending time of the endogenous inaction period. Since a mass of waiting firms are clustered around the center of investment's distribution (i.e., around δ_k), investment's dispersion and cross-sectional skewness both decrease.

Furthermore, if productivity remains persistently low, then at time $t+\hat{t}$ waiting firms pass a tipping point in which they are overly burdened with unproductive capital and choose to disinvest this capital sharply. This implies that $i_{j,t+\hat{t}} << \delta_k$.[37] Thus, these periods of inaction are often followed by negative investment spikes, producing the negative skewness of firm-level investment that is inconsistent with the data.[38]

The distorted distribution of investment rates in the model with fixed utilization also has an

[37] Put differently, when firms are close to the disinvestment threshold, then the investment policy is locally concave and the expected value of investment becomes negative.

[38] Importantly, the counterfactuals in the model with fixed utilization cannot be remedied simply via the aggregation level. As many real options operate at the plant level, one can argue that a collection of production units in the model comprise one firm. Aggregation of many units into a single firm does indeed smooth model-implied investment rates by shrinking periods of investment inaction, and lowering the size of disinvestment jumps. However, these model-implied moments remain unaligned with the data. We verify this in untabulated simulation by aggregating 100 production units into a firm. The resulting model-implied volatility of investment is smaller than the data. While the skewness of investment turns positive, this quantity is close to zero (remaining significantly lower than the data).

adverse impact on risk premia. Because investment's distribution features too little dispersion and asymmetry, there is too little heterogeneity between firms' risk exposures. Sorting firms into portfolios based on investment (or valuation ratios) implies that both the top and bottom quintiles (tails) contain fewer extreme outcomes compared to the benchmark with flexible utilization. As differences in cross-sectional risk premia are fundamentally driven by heterogeneity in investment, cross-sectional spreads get smaller with fixed utilization.

OA.6.2 Flexible utilization: a solution

Our benchmark model with flexible utilization overcomes the counterfactuals outlined in Section OA.6.1 by making the depreciation rate endogenously stochastic. This improved model fit is highlighted in row (2) of Table 9 by showing model-implied moments under the benchmark value of λ. When firms can choose utilization, they have an extra mechanism by which to scale down production in response to adverse productivity shocks, even as they "wait and see" if productivity recovers. That is, firms can respond to moderate drops in productivity by utilizing their existing machines less intensively rather than selling machines. As underutilized capital depreciates slower, more capital is preserved for more productive future periods (i.e., $\delta(u_{j,\tau}) < \delta_k$ if $u_{j,\tau} < 1\ \forall\tau \in [t, t+\hat{t})$).

Lower utilization reduces the natural investment rate needed to maintain the current capital stock. Thus, even as firms wait to sell capital, the investment required to maintain existing machines, $i_{j,\tau} = \delta(u_{j,\tau})$, becomes endogenously stochastic. This time-varying depreciation eliminates the long periods of constant investment. The time-series volatility of firm-level investment rises, and the cross-sectional dispersion of investment increases. To see the latter, note that firms' utilization rates depend on idiosyncratic productivity shocks. Since these shocks differ between waiting firms, $u_{j,\tau} \neq u_{k,\tau} \Rightarrow i_{j,\tau} \neq i_{k,\tau}$ for firms j and k.

Moreover, the positive correlation between productivity and utilization also implies that firms opt to raise utilization in times of high productivity. Utilizing capital more intensively in good times raises both depreciation and the natural rate of investment (i.e., $\delta(u_{j,\tau}) > \delta_k$), and means that larger investments are needed to expand capacity in future periods. To see this, suppose that at time τ a firm wants to expand capacity by $\delta_k K$. With fixed utilization, the required investment rate is $i_\tau = 2\delta_k$. However, with flexible utilization, the required investment rate rises to $i_\tau = \delta(u_{j,\tau}) + \delta_k > 2\delta_k$. Since investment becomes more procyclical, its time-series and cross-sectional skewness rise and turn positive (in line with the data).

The increases in the skewness and dispersion of investment under flexible utilization also boost risk premia, as seen by comparing the value premium between rows (2) and (1) of Table 9. A larger value premium under flexible utilization can be attributed to the fact that the cross-section of investment rates is more dispersed and almost 17 times as skewed in the model with flexible utilization.

Greater dispersion in investment leads to more heterogeneity in risk exposures to aggregate productivity, which increases return spreads.[39]

[39]The increase in the cross-sectional skewness of investment also has an impact on the value premium. With fixed utilization, and *symmetric* cross-sectional distribution, the portfolio of growth firms (bottom 20% of book-to-market) includes both firms with very high and moderately high investment rates (or Tobin's Q). With flexible utilization, and *asymmetric* cross-sectional distribution, the right tail of investment's distribution becomes thicker, and the portfolio of growth firms includes firms with only very high investment rates. As these firms expand capacity significantly (suggesting a much higher Q) precisely when the price of risk is high (bad states), their skewed investment behavior

More generally, as shown in rows (3) to (6) of Panel A in Table 9, the value of λ has a substantial quantitative impact on matching the data. As utilization becomes more flexible (i.e., λ decreases), the time-series/cross-sectional skewness and volatility of investment rise, the autocorrelation of investment slightly declines, and risk premia increase as well. This generally moves each moment towards its empirical counterpart when compared to the case of $\lambda \to \infty$. In particular, row (6) shows that when utilization becomes less flexible (i.e., λ is finite but high), investment's skewness and dispersion are too low. Rows (4) and (5) show that our results are only mildly affected by small perturbations of the benchmark value of λ. However, utilization cannot be overly flexible. When λ is very low, as in row (3), the volatility of utilization exceeds the 95% confidence interval of this quantity in the data.

OA.6.3 Required adjustment costs under flexible utilization

Without flexible utilization, the problem of matching investment's moments with the data is not simply alleviated by recalibrating the model. In this section we show that flexible utilization can reduce the magnitude of exogenous adjustment cost parameters required to target investment and risk premia jointly. We illustrate this role of utilization by (i) perturbing the capital adjustment cost parameter, and (ii) generalizing the adjustment cost function in a model without utilization.

Perturbing adjustment costs. In rows (7) to (10) of Panel B in Table 9 we alter the quadratic adjustment cost while keeping utilization fixed. Rows (7) and (8) consider the case of lower frictions compared to the benchmark. Sufficiently lower friction (row (7) can help turn the time-series skewness of investment to a positive value, but the cross-sectional skewness of investment is still too small. Lower frictions also cause risk premia to fall. The value premium, which is already too low in the model with fixed utilization, falls in row (7) to almost half of its empirical magnitude.

The diminished value premium in the model with fixed utilization can be boosted by increasing the quadratic capital adjustment cost. With higher adjustment costs, shocks are absorbed in asset prices rather than investment quantities. We demonstrate this in rows (9) and (10). We search for an adjustment cost parameter ϕ to match the value premium in the model with fixed utilization. Our structural search suggests that ϕ needs to be around 3.00 to match this spread (see row (10)). While this parameter is broadly consistent with existing literature, this value is *double* the value of ϕ under our benchmark model with flexible utilization. Moreover, doubling ϕ simultaneously distorts investment's distribution. Investment's dispersion becomes a quarter of its empirical magnitude. The time-series skewness of investment becomes even more counterfactually negative.[40]

Flexible utilization provides a channel that addresses the aforementioned concerns. Row (2) of Table 9 shows that flexible utilization allows our baseline model to feature adjustment costs that

provides an excellent hedge against bad states. This decreases the risk premium of the growth portfolio, and increases the magnitude of the value premium.

[40]In untabulated results, we further verify that under fixed utilization, the value premium cannot be targeted successfully in the model (jointly with investment's moments) by changing the fixed cost f. Specifically, we find that lowering the fixed cost f causes investment's skewness to rise and move closer to the data. However, even in the extreme case of $f = 0$, the cross-sectional skewness of investment is only a half of its empirical magnitude. Moreover, lowering f decreases the magnitude of the value premium (e.g., when $f = 0.01$, the implied value premium is merely 2.7%). Similarly, we find that while raising the fixed cost f to 0.06 allows us to obtain a value premium of 3.2% (roughly consistent with the data), this causes the model's investment mismatch to become even more severe. The time-series and cross-sectional skewness of investment become counterfactually negative (about -0.75), and investment's volatility falls to 10%.

are smaller than those required with fixed utilization. These smaller costs are sufficient to simultaneously produce sizable risk premia spreads, including many periods of depressed investment.[41] This happens because of the following key mechanisms.

The mechanism. The first mechanism is related to the impact of lower utilization on observed investment rates. For a given quadratic adjustment cost parameter, flexible utilization implies more observed disinvestment while keeping the amount of friction (risk) the same. To see this, suppose a firm wishes to drop its capital stock by $\delta_k K$. With fixed utilization, the firm chooses an investment rate of $i = 0$, and the quadratic cost is proportional to δ_k^2. However, with flexible utilization, a drop in productivity triggers a drop in utilization, which in turn lowers the firm's depreciation to $\delta(u_{i,t}) < \delta_k$. To shed $\delta_k K$ capital, the investment rate is set to the lower rate of $i = -\delta_k + \delta(u_{i,t}) < 0$. The quadratic cost will be unaltered, and remain proportional to $(-\delta_k + \delta(u_{i,t}) - \delta(u_{i,t}))^2 = \delta_k^2$. Thus, in a model with flexible utilization, one may see more disinvestment without compromising on the frictions that induce risk premia.

The second mechanism involves the changes in the cross-sectional distribution of investment rates that are caused by flexible utilization. As we outline in Section OA.6.2, utilization makes the cross-sectional distribution of investment more dispersed and skewed. By featuring more extreme observations in the tails of investment's distribution, one can obtain quantitatively large return spreads with moderated values of adjustment costs.

The third mechanism is utilization's ability to enhance payout cyclicality. Value firms have low idiosyncratic productivity and high capital, and desire to reduce excess capital in bad states. With fixed utilization, these firms are riskier because they need to pay large quadratic costs that reduce their payoff precisely when aggregate productivity is low. With flexible utilization, these firms also desire to lower utilization to conserve capital for future periods. The reduction in utilization implies today's output is even lower, contemporaneously with a bad aggregate state. Thus, the cyclicality of output is larger, amplifying risk.

Alternative adjustment costs. A model with fixed utilization can potentially match the aforementioned moments using a more elaborate adjustment cost function. Extra adjustment costs may include piecewise quadratic and linear terms (e.g., Belo and Lin (2012)).

While it is hard to rule out the possibility of an admissible adjustment cost function, the next subsection provides evidence that adjustment costs in the form of Cooper and Haltiwanger (2006) are unlikely to reconcile risk premia alongside the time-series and cross-sectional moments of investment. We augment the model to feature asymmetric quadratic and linear adjustment costs. The downside quadratic (linear) coefficient is ten times (two to three times) larger than the upside coefficient, in line with extant papers. We find that the asymmetric fixed cost has a negligible effect on all moments. By contrast, when the quadratic adjustment cost is asymmetric, and re-calibrated to successfully match the cross-sectional skewness of investment, the time-series skewness and volatility of investment are much larger than the data, and return spreads are still too low.

[41] Moreover, when utilization is flexible, this result is not very sensitive to small perturbations of ϕ. Rows (12) and (13) of Table 9 show that when the quadratic adjustment cost (ϕ) slightly increases (decreases), the volatility of investment drops (rises), the value premium slightly increases (decreases), but these moments are almost identical to the benchmark in row (2). Row (14) shows that if ϕ is set to 3.00 (which is required in the case of fixed utilization), then the implied value premium is almost 5%, about 1.5% above its value under fixed utilization. In this case, the cross-sectional skewness of investment aligns very well with the data, but the cross-sectional dispersion is too low.

OA.6.4 Alternative adjustment cost specification

We document that even if the capital adjustment cost function given by equation (9) is augmented to include piecewise linear and quadratic terms, the model with fixed utilization is not likely to match key investment-related moments, while generating high risk-premia spreads.

Consistent with the works of Cooper and Haltiwanger (2006), Belo and Lin (2012), and Belo et al. (2014), among others, we generalize the adjustment cost function employed by our benchmark analysis to include asymmetric linear and quadratic adjustment cost. Specifically, the adjustment cost function we employ is given by

$$G_{i,t} = \left[\frac{\phi^+}{2} (ni_{i,t})^2 \, \mathbf{1}_{\{ni_{i,t}>0\}} + \frac{\phi^-}{2} (ni_{i,t})^2 \, \mathbf{1}_{\{ni_{i,t}<0\}} + f^+ \mathbf{1}_{\{ni_{i,t}>0\}} + f^- \mathbf{1}_{\{ni_{i,t}<0\}} \right] K_{i,t}. \quad \text{(OA.6.1)}$$

Here, $n_{i,t} \equiv \frac{I_{i,t}}{K_{i,t}} - \delta_k$ denotes the net investment rate of firm i at time t, while the functions denoted by $\mathbf{1}_{\{ni_{i,t}>0\}}$ ($\mathbf{1}_{\{ni_{i,t}<0\}}$) are indicator variables that take on a value of one when the firm increases (reduces) its stock of capital. The adjustment parameter ϕ^+ (ϕ^-) capture the quadratic adjustment cost of increasing (decreasing) capacity, while f^+ (f^-) captures the fixed adjustment cost of increasing (decreasing) capacity. While our benchmark adjustment cost specification imposes the constraints that $\phi^+ = \phi^-$ and $f^+ = 0$, we relax these constraints below and consider more general specification in which $\phi^+ \neq \phi^-$ and $f^+ > 0$.

The model in Zhang (2005) suggests the quadratic adjustment cost for investment is around one tenth of the magnitude of that for disinvestment. Accordingly, we focus on the case that $\phi^+ = \phi^- \times \frac{1}{10}$. Likewise, the fixed cost of investment in Belo and Lin (2012) and Belo et al. (2014) is approximately one third to one half of the magnitude of that for disinvestment. Consequently, we consider the cases in which $f^+ = f^{-1} \times \frac{1}{3}$ or $f^+ = f^{-1} \times \frac{1}{2}$. For completeness, we also consider calibrations in which the ratios f^+/f^- or ϕ^+/ϕ^- differ from these prior studies.

Table OA.6.1 reports the model-implied time-series and cross-sectional moments of investment rates, as well as the magnitude of cross-sectional risk premia, in a fixed-utilization model featuring asymmetric capital adjustment costs. Rows (2) to (6) of the table consider calibrations in which $\phi^+ = \phi^-$ in equation (OA.6.1), identical to our baseline calibration reported in row (1), but $f^+ > 0$. We keep in f^- at a positive value, consistent with the reason outlined in Section 2.1. By and large, perturbing introducing an upside linear adjustment cost does not help to reconcile the data. For instance, when $f^+ = f^- \times \frac{1}{2}$ in row (6) the dispersion and volatility of investment rates, as well as the magnitudes of the risk premia, remain far lower than their counterparts in the data.

Rows (7) to (9) of the table keep $f^+ = 0$ (identical to our baseline model) but allow for asymmetry in the quadratic adjustment costs (i.e., $\phi^+ \neq \phi^-$). Focusing on row (8), a calibration in which $\phi^+ = \phi^- \times \frac{1}{10}$, shows that this asymmetry can reconcile *cross-sectional* moments of investment rates with the data (e.g., the cross-sectional skewness is about 1.8 in both the model and the data). However, the same calibration fails to match the *time-series* moments of investment rates and risk premia. Specifically, the model-implied time-series volatility (skewness) of investment rates is almost twice (five times) as high as its empirical counterpart. Moreover, risk premia in this model are too low compared to the data.

In rows (10) and (11) the adjustment cost function is calibrated to feature both asymmetric linear *and* quadratic adjustment costs. Here, the degree of asymmetry follows the degrees of asymmetry considered in the literature (see, e.g., Zhang (2005), Belo and Lin (2012), and Belo et al.

(2014)). The results indicate that even with the most general form of the asymmetric adjustment costs, the fixed-utilization model is only able to reconcile the cross-sectional moments at the expense of not matching the time-series of investment and risk premia. For instance, row (10) is quite similar to row (8) in terms of matches and mismatches.

To complement the above evidence, we also search for a value of ϕ^- that is able to match the value premium, similar to the exercise presented in Section OA.6.3. The search result yields two conclusions that are almost identical to Section OA.6.3. First, even though we feature a more elaborate adjustment cost, as in Cooper and Haltiwanger (2006), we still need to increase the exogenous amount of capital friction compared to the flexible utilization model. We find that the downside quadratic adjustment cost ϕ^- has to be *doubled*. That is, a calibration in which $\phi^- = 3$, $\phi^+ = \phi^-/10$ and $f^+ = f^-/3$ yields a value premium of 3.5%, which is close to the data. Second, this higher value of ϕ^- renders two prominent counterfactual moments for investment: the time-series skewness of investment is much larger than the data (about 3.5), and the cross-sectional dispersion is about a half of the data (about 0.08).

To conclude, the evidence points out that asymmetry in the linear cost has only a marginal effect on the results, while asymmetry in ϕ helps only for the cross-section of investment. In particular, row (8) shows that without deviating from standard values of α, δ_k, or α_l, a model with fixed utilization and real options cannot match the data. Importantly, even if such calibration was feasible, flexible utilization would still offer valuable merit: flexible utilization provides a way to rely on a lower dimensional adjustment cost function and endogonize the implications of additional exogenous calibration parameters using a micro-founded margin. The only additional model parameter needed to accommodate flexible utilization is λ.

OA.7 Utilization and depreciation: Empirical evidence

OA.7.1 Utilization and depreciation dynamics

Equation (8) suggests that firms' depreciation rates should correlate positively with their utilization rates. In this section we check this prediction and explore its implications for the accuracy and the frequency of depreciation's measurement.

We examine the relation between utilization and depreciation via the projection

$$\Delta\delta_{j,t} = d_j + d_u \Delta u_{j,t} + d_x X_{j,t} + \varepsilon_{j,t},$$

where j denotes an industry index, $\Delta\delta_{j,t}$ is the log growth of industry j's depreciation rate from BEA, $\Delta u_{j,t}$ is the log growth of industry j's utilization rate, and $X_{j,t}$ is a control variable. We use the log growth of depreciation and utilization to reduce persistence in these variables, and to account for a non-linear relation between the level of the two. All variables are standardized for the ease of interpretation. The results of the projection are reported in Table OA.7.1. In columns (1) and (2) of the table we run the projection without any controls. We show that the simple correlation between log depreciation growth and log utilization growth is 30%, and that this correlation is not affected by the inclusion of industry fixed effects.

Recent studies in production-based asset pricing show that BEA- and Compustat-implied depreciation rates are strikingly different. The use of one over the other can lead to economically sizable differences in the distribution of gross investment rates (e.g., Clementi and Palazzo (2019); Bai et al. (2019)). In line with these papers, columns (3) and (4) of Table OA.7.1 demonstrate the

Table OA.6.1: **Model-implied moments and asymmetric capital adjustment costs**

Row	Model	Time-series			Cross-sectional		Risk premia	
		$\sigma_{TS}(ik)$	$S_{TS}(ik)$	$\rho_1(ik)$	$\sigma_{CS}(ik)$	$S_{CS}(ik)$	$E\left[R^{bm}\right]$	$E\left[R^{ik}\right]$
	Data	0.14	0.67	0.52	0.16	1.89	3.71	3.70
	Baseline without utilization							
(1)	$(\lambda \to \infty)$	0.11	-0.27	0.63	0.07	0.07	3.00	2.29
	Unchanged ϕ and $f^+ > 0$							
(2)	$f^+ = f^- \times \frac{1}{20}$	0.11	-0.20	0.62	0.07	0.13	2.99	2.12
(3)	$f^+ = f^- \times \frac{1}{10}$	0.11	-0.14	0.61	0.07	0.19	2.99	2.03
(4)	$f^+ = f^- \times \frac{1}{5}$	0.11	-0.03	0.60	0.07	0.27	2.99	1.88
(5)	$f^+ = f^- \times \frac{1}{3}$	0.11	0.10	0.59	0.07	0.41	2.98	1.77
(6)	$f^+ = f^- \times \frac{1}{2}$	0.11	0.24	0.57	0.07	0.57	2.97	1.65
	Asymmetric ϕ and $f^+ = 0$							
(7)	$\phi^+ = \phi^- \times \frac{1}{20}$	0.29	4.60	0.33	0.21	2.52	2.37	2.24
(8)	$\phi^+ = \phi^- \times \frac{1}{10}$	0.23	3.41	0.42	0.17	1.81	2.54	2.28
(9)	$\phi^+ = \phi^- \times \frac{1}{5}$	0.18	2.28	0.50	0.13	1.18	2.70	2.27
	Asymmetric ϕ and $f^+ > 0$							
(10)	$f^+ = f^- \times \frac{1}{3}$, and $\phi^+ = \phi^- \times \frac{1}{10}$	0.24	3.50	0.38	0.17	2.07	2.51	1.99
(11)	$f^+ = f^- \times \frac{1}{2}$, and $\phi^+ = \phi^- \times \frac{1}{10}$	0.23	3.41	0.42	0.17	1.81	2.54	2.28

The table reports model-implied population moments related to the time-series and cross-section of investment rates, as well as risk premia, under various calibrations of the model featuring an asymmetric capital adjustment cost function. Specifically, the augmented adjustment cost function is

$$G_{i,t} = \left[\frac{\phi^+}{2}(ni_{i,t})^2 \mathbf{1}_{\{ni_{i,t}>0\}} + \frac{\phi^-}{2}(ni_{i,t})^2 \mathbf{1}_{\{ni_{i,t}<0\}} + f^+\mathbf{1}_{\{ni_{i,t}>0\}} + f^-\mathbf{1}_{\{ni_{i,t}<0\}}\right] K_{i,t},$$

where $ni_{i,t} \equiv \frac{I_{i,t}}{K_{i,t}} - \delta_k$ represents the net investment rate of the firm, and the functions denoted by $\mathbf{1}_{\{ni_{i,t}>0\}}$ ($\mathbf{1}_{\{ni_{i,t}<0\}}$) are indicator variables that take on a value of one when the firm increases (reduces) its capacity. Here, . We consider various specifications of the function above, whereby we alter ϕ^+ (f^+) to either be unchanged relative to the baseline model, or a fixed multiple of ϕ^- (f^-).The table reports the time-series volatility ($\sigma_{TS}(ik)$), skewness ($S_{TS}(ik)$), and the first-order autocorrelation ($\rho(ik)$) of firm-level investment rates, as well as the cross-sectional dispersion ($\sigma_{CS}(ik)$) and skewness ($S_{CS}(ik)$) of investment rates. In addition, the table also reports the value premium ($E\left[R^{bm}\right]$) and investment premium ($E\left[R^{ik}\right]$) obtained by sorting the cross-section of model-implied returns association with each calibration on book-to-market ratios and investment rates, respectively. These risk premia are expressed as an annualized percentage. Each alternative calibration is identical to the benchmark calibration in all ways except for the degree of asymmetry in the alternative adjustment cost function. All moments are based on a simulations of 1,000 firms over 40,0000 periods (years). Finally, the top row of the table also reports the empirical counterpart of each moment.

discrepancy between these two depreciation measures. We set $X_{j,t}$ to be the log growth of industry j's Compustat-based depreciation rate, and restrict d_u to zero. The correlation between the growth of these depreciation measures is only 3%.[42]

In columns (5) and (6) we do not restrict d_u to be zero. First, the positive correlation between BEA-implied depreciation growth and utilization growth remains positive and sizable when controlling for Compustat-implied depreciation. Second, the (partial) correlation between the growth rates of BEA- and Compustat-implied depreciation increases to 14%. Thus, utilization narrows the wedge between these two measures. While measurement error may exist in both measures, the fact that the correlation between the two increases when controlling for utilization suggests that

[42]We describe the measurement of the these depreciation rates in Section OA.1 of the Online Appendix.

Table OA.7.1: **Empirical relation between capacity utilization and depreciation rates**

	(1)	(2)	(3)	(4)	(5)	(6)
β_{UTIL}	0.30	0.30			0.28	0.28
	(3.11)	(3.14)			(2.80)	(2.83)
β_{COMP}			0.03	0.03	0.13	0.14
			(3.50)	(3.53)	(3.54)	(3.58)
Industry FE	No	Yes	No	Yes	No	Yes
R^2	0.09	0.09	0.02	0.02	0.10	0.10

The table reports the empirical relation between the industry-level capacity utilization rate, industry-level depreciation rate from the BEA, and industry-level depreciation rate from Compustat. In each specification considered in the table we run projections of the log-growth rate of BEA-implied depreciation on the log-growth rates of capacity utilization and Compustat-implied depreciation rates, and standardize all variables for ease of interpretation. Columns (1), (3), and (5) of the table estimated pooled-OLS regressions, whereas columns (2), (4), and (6) of the table estimate panel regressions including industry fixed effects. t-statistics, reported in parentheses, and computed using standard errors clustered at the industry level. Finally, the time span underlying the regressions is from January 1967 to December 2015.

utilization can be used to accurately filter the true depreciation rate. We briefly illustrate this point in the next subsection.

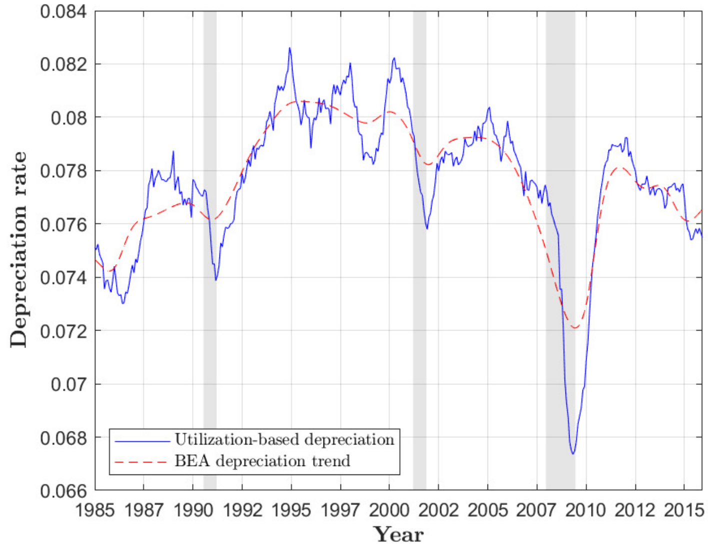
OA.7.2 High-frequency depreciation based on utilization data

As utilization data is available at the monthly frequency, the utilization-implied depreciation rate we propose is computed at a higher frequency than depreciation rates implied by either BEA or Compustat. First, to be consistent with the model, we adjust each industry's utilization rate to have a mean of one. Then, for each industry j, we obtain a utilization-implied depreciation rate, $\delta(u_j)$, by applying equation (8) to the industry's utilization data. Here, we use the model parameters in Table 6. Second, we average these depreciation rates across all industries to obtain an aggregate utilization-implied depreciation rate, $\delta(u_{agg})$. Third, we adjust $\delta(u_{agg})$ to share the same trend as the aggregate depreciation rate from the BEA. We do this by combining the business-cycle component of $\delta(u_{agg})$ with the stochastic trend component of the BEA's aggregate depreciation rate.[43] We obtain the components of each time-series using the Hodrick and Prescott (1997) filter.

In Figure OA.7.1 we plot the monthly time-series of $\delta(u_{agg})$ alongside the trend of the BEA's aggregate depreciation rate. By construction, the two time-series share the same trend, but $\delta(u_{agg})$ shows high-frequency business-cycle fluctuations around this common trend. These fluctuations could be important as they amplify the volatility of gross investment rates and can help to reconcile the dynamics of BEA- and Compustat-implied depreciation rates.

[43] The third step is optional and meant to ensure both series follow the same trend. An alternative measure of the utilization-implied depreciation rate involves only the first two steps, with similar results.

Figure OA.7.1: **Utilization-based high-frequency depreciation rate**



The figure shows a high-frequency measure of the aggregate depreciation rate. The figure reports the time series of the high-frequency depreciation rate from 1985, the beginning of the Great Moderation, to 2015.

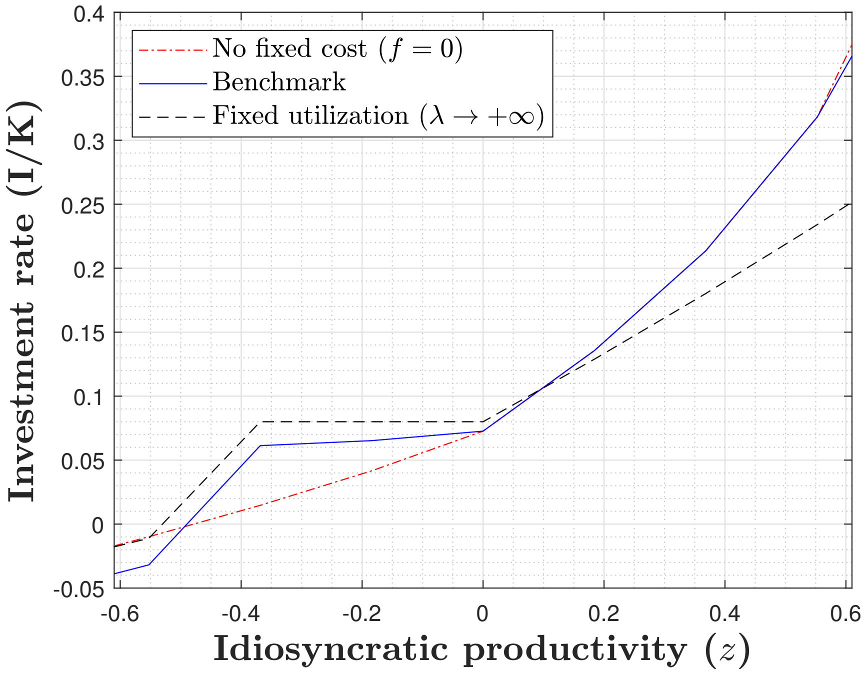
OA.8 Supplemental tables and figures

Table OA.8.1: **Model-implied CAPM alpha**

Portfolio	$[R^{CU}]$	α_{CAPM}
Low (L)	8.89 [4.71,15.16]	2.84 [0.78,4.88]
Medium	6.77 [3.40,12.37]	0.98 [0.47,1.49]
High (H)	4.96 [0.51,11.00]	-0.61 [-3.15,1.94]
Spread (L-H)	3.93 [0.67,7.45]	3.45 [-0.24,6.83]

The table reports the average annual value-weighted returns and CAPM alphas (α_{CAPM}) of portfolios sorted on capacity utilization at the industry-level across short-sample simulations of our model economy. As in the empirical analysis, an industry is sorted into the high (low) utilization portfolio if its level of capacity utilization is above (below) the 90^{th} (10^{th}) percentile of the cross-sectional distribution of capacity utilization rates in the previous period. Industry-level returns are simulated using the procedure described in Section 3.1, and short-sample moments are obtained by averaging moments across 500 simulations of 50 industries for 50 periods (years). Finally, square brackets report the 90% confidence interval related to each moment across the 500 Monte Carlo simulations of the economy.

Figure OA.8.1: **Model-implied investment policy**



The figure shows the optimal investment rate policy (I/K) as a function of idiosyncratic productivity (z). Capital and aggregate productivity are set at their stochastic steady-state values, and we let z varying between two standard deviations of its mean value. We consider the I/K policy under three versions of our model: (1) The benchmark model (solid blue line), (2) the model without fixed costs (i.e., $f = 0$) (the dashed red line), and (3) the model with fixed utilization (i.e., $\lambda \to +\infty$).