Endogenous Capital User Costs and Fluctuations in Firm Productivity *

Poorya Kabir†, Eugene Tan‡, and Ia Vardishvili§

December 30, 2023

Abstract

We provide direct evidence, using unique data from India, that endogenous capital utilization and maintenance—measures of capital user cost—drive fluctuations in firm total factor productivity (TFP) and average revenue product of capital (ARPK). Drawing inference from a calibrated neoclassical investment model with capital adjustment frictions and endogenous capital user costs, we uncover three novel results: In India, (i) almost one-third of fluctuations in firm TFP arises from endogenous utilization, (ii) user cost variations amplify dispersion in ARPK but are associated with a 4.0% gain to aggregate productivity, and (iii) elimination of maintenance inefficiencies increases aggregate productivity by 13.6% .

*We would like to thank Miguel H. Ferreira and Gorkem Bostanci for insightful discussions of the paper. We would also like to thank Diego Restuccia, Julieta Caunedo, Matthias Kehrig, Andrea Lanteri, Andy Abel, Víctor Ríos Rull, Harold Cole, Urban Jermann, and Gideon Bornstein for their insightful comments. We are grateful to various conference and seminar participants. Kabir acknowledges financial support from the Singapore Ministry of Education AcRF Tier 1 Research Grant No. A-8000758-00-00.

†National University of Singapore: poorya.kabir@nus.edu.sg

‡University of Toronto: eugene.tan@rotman.utoronto.ca

§Auburn University: izv0013@auburn.edu

1 Introduction

It is well documented that firm total factor productivity (TFP)—that is, the firm-level Solow residual—and average revenue product of capital (ARPK) is highly volatile. Moreover, firm TFP and ARPK are also highly correlated. By and large, the interpretation of these statistical facts is that fluctuations in TFP arises from innovations to TFP (i.e., uncertainty), which, along with a myriad of investment, market, or size-dependent policy frictions, propagate into fluctuations in ARPK (e.g., David and Venkateswaran (2019)). In turn, volatility in ARPK, and a positive covariation with TFP, is often seen as synonymous with losses to aggregate productivity (e.g., Restuccia and Rogerson (2008); Hsieh and Klenow (2009)); that is capital "misallocation" arising from uncertainty, investment, market, or size-dependent policy frictions.

However, fluctuations and co-movement in firm TFP and ARPK might also reflect endogenous variations in capital utilization and endogenous capital user costs, as in the literature on business cycles (e.g., Burnside and Eichenbaum (1996); Basu, Fernald, and Kimball (2006)). Importantly, such variations can reflect a firm's ability to exploit variable utilization as a strategy to mitigate the capital constraints that arises when capital adjustment frictions render the capital stock quasi-fixed, a standard mechanism in workhorse business cycle models. In other words, fluctuations and co-movement in TFP and APRK need not reflect a drag on aggregate productivity. On the contrary, they might reflect gains to aggregate productivity if it reflects flexibility of firms to mitigate capital frictions. However, despite the potentially important role of endogenous user costs in measuring capital misallocation, unlike the business cycle literature, there is virtually no research in the firm dynamics literature that investigates the role of endogenous user cost as a source of fluctuations in firm productivity.

In this paper, we study the degree to which endogenous user cost drives fluctuations in firm TFP and ARPK and in turn, quantify their implications for inference on investment frictions and capital misallocation. Using a combination of novel data on capital user costs and a structural model of firm investment and endogenous user costs, we argue that more than one-third of fluctuations in TFP and ARPK are attributed to fluctuations in user costs. In turn, we show that the dispersion in ARPK accounted for by endogenous user cost is associated with aggregate productivity *gains*—our calibrated model economy predicts that aggregate productivity is around 4.0% higher relative to an otherwise equivalent model with fixed user cost. This stands in contrast to the conventional argument (e.g., Hsieh and Klenow (2009)), based on assumptions of fixed capital user costs, that higher dispersion in ARPK is associated with aggregate productivity *losses*.

To provide empirical support for our hypothesis, we bring in new evidence using firm-level accounting data from companies in India, where besides standard balance sheet and expense items, we also have information on both capital maintenance expenses and capacity utilization rates. In particular, we argue that maintenance rates, the amount of maintenance expenses relative to the capital stock, are a literal measure of capital user costs, since they are the literal costs firms pay to keep the capital stock usable. Furthermore, maintenance rates are large and highly volatile: The median maintenance rate is 3.5% of capital stock, comparable to typically assumed values for capital depreciation rates of between 6% to 10%; and has a within-firm standard deviation of 0.83 , similar in magnitude to the within-firm standard deviation of ARPK (around 0.66).

Our evidence can be iterated as follows. First, we show that when capacity utilization is corrected for measurement error using an instrumental variables approach, a 1% increase in utilization rates corresponds to a 1% increase in firm TFP. This is consistent with our hypothesis that a fraction of TFP is simply due to variations in capacity utilization. Second, we show that plausibly exogenous increases in capacity utilization is associated with an increase in capital user costs in the form of higher maintenance rates. Depending on the specific instrumental variable, a 1% increase in utilization implies a 1.67% to 1.82% increase in maintenance rates. Finally, we show that consistent with standard Jorgensonian investment theory, ARPK is indeed highly correlated with capital user cost: A 1% increase in maintenance rates is associated with a 0.60% increase in ARPK. In combination, our results show that there is a direct relationship between higher utilization rates, higher TFP, higher user costs, and finally, higher ARPK.

Motivated by these stylized facts, we construct and calibrate a quantitative firm investment model to quantify the contribution of endogenous user costs to fluctuations in firm TFP and ARPK. Our model is an otherwise standard neoclassical investment model with partial investment irreversibility, modified to include endogenous depreciation due to varying capital utilization and maintenance rates (e.g., Greenwood, Hercowitz, and Huffman (1988); McGrattan and Schmitz (1999)), a flow cost in maintenance, as well as maintenance disruption costs. Our formulation of endogenous depreciation follows Caunedo and Keller (2021), whereby utilization and maintenance are complements. In turn, while capital depreciation rates are relatively stable in the model, total user costs are highly volatile due to fluctuations in maintenance expenses in response to variations in capital utilization rates. Finally, maintenance disruption costs—a channel novel our framework—dampen the responsiveness of maintenance to productivity shocks, allowing us to replicate the dynamics of maintenance as in the data. While our calibrated model matches our targeted moments well, we also conduct a series of additional empirical analysis for external validation. For instance, in our model, our formulation of maintenance disruption costs imply that the elasticity of maintenance to firm TFP and ARPK is decreasing in utilization rates. We show that this result is replicated in the data. Using the calibrated model, we infer that around 36.5% of volatility in firm productivity, and 42.5% of volatility in ARPK, can be explained by endogenous fluctuations in capital utilization rates and user costs.

We examine three implications of our model, given the quantitative relevance of endogenous user costs in driving fluctuations in firm productivity. First, we compare aggregate total factor productivity and dispersion of ARPK predicted by our baseline model relative to an otherwise equivalent model with fixed user costs. We find that aggregate productivity is about 4.0% higher in our baseline model relative to the fixed user cost counterfactual, even though the dispersion in ARPK is almost 42.5% lower in the counterfactual. We show that the key mechanism is that variable capital utilization effectively converts the capital stock from a quasi-fixed factor to a flexible factor. Specifically, it allows firms that have "too little" ("too much") capital relative to productivity to overcome their short-run constraint by utilizing capital more (less) intensely. In turn, endogenous utilization "undoes" the negative effects of uncertainty and capital adjustment frictions, but amplifies the dispersion of ARPK.

Second, we examine the role of maintenance disruption costs for aggregate total factor productivity. We find that, by muting the presence of maintenance disruption costs, aggregate TFP increases by almost 13.6% . This gain is much larger than TFP losses attributable to conventional capital adjustment frictions (e.g., David and Venkateswaran (2019) find that capital adjustment frictions lower productivity by about 1% in China). We find that the lion's share of the gains to aggregate TFP (approximately 87%) comes through higher aggregate utilization—by lowering the cost of maintenance, aggregate utilization in India rises from around 69% to 78%. Importantly, we find that these increases are empirically plausible, with the baseline utilization rate consistent with our data, and the upper bound of gains consistent with estimates in the United States.

Finally, we examine the implications for inference on the size of investment irreversibility. Our baseline model predicts a capital resale loss of 8.7% , consistent with standard estimates using data from the United States (e.g., Lanteri (2018)). Conversely, we find that a re-calibrated model with fixed user costs would require a 40% loss in capital to match the same set of targeted moments, similar to estimates using data from developing economies with much larger estimated volatility of firm TFP (e.g., Lanteri, Medina, and Tan (2023)). In our framework, these different estimates arise because in the fixed user cost counterfactual, all variations in firm TFP is mechanically loaded into volatility in actual productivity. Consequently, holding all else constant, the counterfactual model predicts a much more volatile investment process, which necessitates larger investment frictions to dampen the investment process. In this context, our paper provides a potential resolution to the large range of estimates in the literature, and in particular, could be informative for future researchers who are interested in studying the dynamics of capital reallocation in India.

The rest of our paper is as follows. After the literature review, we present our data in Section 2, and main stylized facts that motivate our analysis of endogenous capital utilization and user costs in Section 3. Then, in Section 4, we present our model and discuss our model mechanisms in relation to our stylized facts. In Section 5, we present additional reduced-form evidence to prove support for our model mechanism, as well as present our calibration strategy and results. In Section 6, we present the three counterfactual exercises we discussed, and we conclude in Section 7.

A Literature Review

Our paper contributes to the literature on firm investment dynamics and capital misallocation in three broad ways.

First, our paper is related to the voluminous literature, building on the seminal contributions of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009), that studies the sources of the dispersion of average revenue products of capital (ARPK), and its relevance for the allocative efficiency of capital in the economy. For instance, papers such as David, Hopenhayn, and Venkateswaran (2016); Midrigan and Xu (2014); Tan (2022); David, Schmid, and Zeke (2022); Edmond, Midrigan, and Xu (2023); Tan and Zeida (2023) focus, respectively, on the role of specific frictions such as information frictions, investment frictions, disinvestment and insurance market frictions, risk premia, output market frictions, and racial animosity in relating the dispersion of ARPK, as well as the correlation of ARPK to TFP, to losses in aggregate productivity. On a similar vein, David and Venkateswaran (2019) provides a comprehensive framework to decompose the sources of capital misallocation into information, capital adjustment, and size-dependent policy frictions. A single commonality in these framework is the assumption of fixed capital user costs, and in turn, the result that dispersion in ARPK is associated with *lower* aggregate productivity. In contrast, we emphasize that dispersion in ARPK and correlation of ARPK with TFP arises naturally in a model of capital user costs, absent any other frictions, and we provide direct empirical evidence regarding this channel. Importantly, we argue that when the dispersion of ARPK is associated with endogenous capital user costs, this dispersion is associated with *higher* aggregate productivity. To some degree, our notation of "good" and "bad" dispersion is similar in spirit to Kehrig and Vincent (2023), although the authors are focused on the dispersion of ARPK within-firm-across-plants, while we focus on the dispersion within-firm-across-time.

Second, our paper fundamentally builds on a long literature in macroeconomics emphasizing the role of endogenous user costs as drivers of fluctuations in aggregate productivity and returns to capital (e.g., Greenwood, Hercowitz, and Huffman (1988); Burnside and Eichenbaum (1996); Basu, Fernald, and Kimball (2006); Fernald (2014); Chodorow-Reich, Karabarbounis, and Kekre (2023)). While relatively new to the literature on heterogeneous firms in the tradition of Hopenhayn (1992), endogenous user costs are a workhorse assumption in business cycle models with nominal rigidities (e.g., Christiano, Eichenbaum, and Evans (2005)). However, unlike the business cycle literature, our paper's goal is to study the role of firm-level heterogeneity in capital utilization as a source of differences in aggregate productivity.

Finally, our paper is also relevant for the literature on estimating (e.g., Cooper and Haltiwanger (2006) and quantifying the role capital adjustment frictions, and in particular, capital reallocation frictions, for aggregate productivity (e.g., David and Venkateswaran (2019)). Since the seminal paper by Ramey and Shapiro (2001), there is now increasing empirical and theoretical evidence that capital reallocation is costly and a source of friction for growth and investment (e.g., Bloom, Floetotto, Jaimovich, Saporta-eksten, and Terry (2018); Lanteri (2018); Kermani and Ma (2022); Baley and Blanco (2022); Lanteri, Medina, and Tan (2023)). We show that estimates of the degree of capital reallocation frictions can depend heavily on the assumptions of user cost dynamics—muting the role of endogenous user costs leads to estimates of investment irreversibility that are five times as large as our baseline estimate.

Furthermore, our paper emphasizes a new interpretation of "capital adjustment friction" in the form of endogenous user cost and, importantly, maintenance disruption costs; where convexity in the cost of using capital sharply curtails the efficiency of the installed capital stock. In turn, we show through our quantitative model that inefficiencies in maintenance have a much larger impact on aggregate productivity than standard adjustment frictions; in our model, aggregate productivity increases by almost 13.6% if we eliminate maintenance frictions in India, whereas standard estimates regarding the cost of capital adjustment frictions range from 1% to 2% (e.g., David and Venkateswaran (2019)).

2 Data and Measurement

In this section, we present our data source, measurement strategy, and report descriptive statistics.

A Data

We use data from Prowess, a database collected by the Center for Monitoring Indian Economy (CMIE). The dataset covers an unbalanced panel of firms from 1989 to 2019 and includes the typical income statement and balance sheet data available in other firm-level datasets such as COMPUSTAT. Two unique features of Prowess are crucial for our study. First, information on the repair and maintenance expenses for a large sample of firms over time.[1] Second, the dataset includes product-level capacity for a subset of firms.

For our analysis, we keep firms with the fiscal month ending on March 31 (more than 95% of firm-year observations) to ensure sample consistency. As we are solely interested in studying within-firm variations, our sample of analysis only includes firms with more than 20 years of data in Prowess. We also require firms to have positive capital, sales, assets, and maintenance. To reduce the effect of outliers in our analysis, we winsorize all variables at a 1% level. Throughout the paper, we use industry to refer to 6-digit NIC codes unless otherwise stated.[2] Finally, we also use GDP deflator data from the World Bank database to convert variables to 2019 values and use the average daily exchange rate of INR to USD in year 2019 (USD = 70.51 INR) to convert INR to USD. Next, we discuss how we measure our variables of interest.

B Measurement

Our analysis in the following sections will require measuring firm total factor productivity, capital stock, investment, maintenance, and capacity utilization. We briefly summarize our measurement strategy in this section and present the relevant summary statistics in Table 1.

[1] Except for McGrattan and Schmitz (1999), which uses the Capital and Repair Expenditures survey from Canada that has information on repair and maintenance cost, we are not aware of any other dataset that includes information on repair and maintenance cost for a large sample of firms.

[2] Our results are robust to alternative industry definitions.

B.1 Capital, Investment, and Total Factor Productivity

We use data on sales and capital inputs to recover productivity and investment. Following Hsieh and Klenow (2009), we consider a "value-added" production function, with capital and labor as the two factors of input. Value added is computed as the difference between sales and materials cost, and capital is measured as the book value of "Plant, Property, and Equipment" (i.e., PPE). The log average revenue product of capital (ARPK) is simply the difference in log value-added and log capital, that is,

$$\log ARPK = \log py - \log k. \tag{1}$$

To compute firm productivity, we assume that firms face a revenue-generating function of the form,

$$py = \left(z\,(uk)^{\alpha}\,l^{1-\alpha}\right)^{\nu}, \tag{2}$$

where z is physical productivity, k and l are capital and labor respectively, and u is the utilization rate of capital. The revenue-generating function can be microfounded by a Cobb-Douglas production function with capital output elasticity α, and an isoelastic demand function with elasticity $\frac{1}{1-\nu}$. In turn, following David and Venkateswaran (2019), we compute firm total factor productivity (ω) using the following formula,

$$\begin{aligned}\log\omega &\equiv \log z + \alpha\log u \\ &= \frac{1-(1-\alpha)\nu}{\nu}\left(\log py - \frac{\alpha\nu}{1-(1-\alpha)\nu}\log k\right),\end{aligned} \tag{3}$$

where we impose that $\alpha = 1/3$ and $\nu = \frac{3}{4}$.[3] Hence, TFP ($\log\omega$) is the sum of true productivity ($\log z$) and utilization rate ($\alpha\log u$).

For investment, we focus on computing statistics associated with net (rather than gross) investment to avoid taking a stand on a value for the depreciation rate. In other words, investment is simply computed as the first-difference in capital stocks; similarly, investment rates are computed net investment normalized by the capital stock.

Table 1 reports the summary statistics. The first and second rows report statistics for the capital stock and value-added. The typical firm is relatively large, with median capital stock and value-added of US$5.7 million and US$10.7 million, respectively.

Rows three and four report net investment and investment rates. Like other firm-level

[3] This formula can be microfounded by a model where labor is a flexible input, and is consistent with our model which we present in Section 4.

datasets, investment features high dispersion and a large excess kurtosis in the right tail. The next two rows show summary statistics for TFP (using equation 3) and ARPK.

B.2 Maintenance Rates

Repair and maintenance expenses are reported in Prowess's income statements in line item "repair and maintenance". For brevity, we refer to this item as maintenance expenses and define maintenance rates as maintenance expenses divided by the capital stock ($\frac{m}{k}$).

Maintenance expenditures in accounting are categorized into two types: expensed and capitalized. Expenditures expensed are recognized in the income statement in the period they are incurred. This approach is commonly applied to routine maintenance activities that neither enhance the asset's value nor prolong its useful life. Examples of such activities include regular servicing, minor repairs, and replacing small parts.

In contrast, capitalizing maintenance costs involves treating these expenses as additions to the value of an asset, consequently reflecting them under the Property, Plant, and Equipment (PPE). This method is typically employed for maintenance activities that substantially increase the asset's value, improve its performance, or extend its useful life. Activities falling under this category include major overhauls, significant upgrades, and enhancements that boost the asset's efficiency or capacity.

Our analysis focuses on the "repair and maintenance" line item from the income statement. This item accounts for maintenance expenditures when they are expensed. It is most relevant to our study as it aligns with our conceptualization of maintenance.

Summary statistics for maintenance are reported in the seventh to ninth rows of Table 1. Maintenance spending is large; on average, capital maintenance rates are about 6.2%, comparable to capital depreciation rates which are typically assumed to be between 6% to 10%. This suggests that maintenance, as a literal user cost, is on the same magnitude as conventional estimates for capital user cost. More importantly, maintenance also exhibits large variability, suggesting that there is substantial volatility in capital user costs.[4]

B.3 Utilization Rates

We use Prowess to construct a measure of capacity utilization at the firm-year level. For a subset of firms, Prowess provides data on product-level installed capacity, defined as "the maximum level of goods and services a company can achieve." We also observe the quantity of products sold. We define product level capacity utilization as the quantity of

[4]This finding is also consistent with Kabir and Mansouri (2021), who document, using the same data, substantial variation in unit cost of capital across different firms within narrowly defined industries.

products sold divided by the installed capacity and drop observations where the measurement units of products sold and installed capacity differ. Finally, we take the average of product-level capacity utilization across different product categories to get capacity utilization at the firm-year level. We denote this measure by $\tilde{u}$, which corresponds to u^{α} in our model. Summary statistics are reported in the last row of Table 1. The average and median capacity utilization is 66.4%, and 54.1%.

Our measure of capacity utilization likely suffers from a classic measurement error problem. First, a firm's reported installed capacity might differ from an economist's notion of maximum capacity. Second, unlike revenue and cost, which directly affect profitability and taxable net income and are audited for accuracy, firms might have fewer incentives to report capacity accurately. Finally, and crucially, the measure of utilization rates is at the product level, and we use simple averages to get firm-level measures. Since we do not have an explicit model of multi-product firms, this rather arbitrary aggregation would also introduce additional measurement error. In Section 3, we use several instrument variable strategies to address this issue.

Table OA.1 in the appendix provides the full summary statistics for the sub-sample of firms with non-missing capacity utilization data. The rest of the rows are the same as table 1. In Figure OA.1 in the Appendix, we do a formal t-test of mean differences across all the variables in the summary statistics table. Overall, we find that firms in the capacity utilization sample are slightly larger, but the normalized variables are very similar and the point estimates for the standardized differences are almost equal to zero and are well below the threshold of 0.20 recommended by Imbens and Rubin (2015). In our analysis, when we are not studying utilization, we focus on the broader sample – those with and without utilization – to have more observations. Nevertheless, our main findings are similar if we only focus on the sub-sample of firms with utilization data.

3 Empirical Evidence

In this section, we present evidence that fluctuations in firm total factor productivity (TFP) and returns to capital (as proxied by the average revenue product of capital, or ARPK), are driven by variations in capital utilization and capital user costs, the latter of which we proxy by maintenance rates. In turn, we argue that the high correlation of TFP and ARPK, a consistent fact documented in the literature and often interpreted as due to correlated distortions, arises because of co-movement with fluctuations in capital utilization and user costs.

To that end, we first present empirical evidence that a component of the variation in within-firm total factor productivity (TFP) comes directly from variations in capacity utilization rates. Next, we show that capital maintenance rates, a component of capital user costs, are highly tied to capacity utilization. Finally, we show that firm investment decisions and ARPK are highly tied to variations in capital maintenance rates. Jointly, these facts explain why firm ARPK is positively correlated with TFP.

A Fact 0: Firm TFP and ARPK are highly volatile and positively correlated

It is well documented that firm TFP and ARPK are highly dispersed and positively correlated (e.g.,Asker, Collard-wexler, and Loecker (2014)). As our starting point, we document the size of TFP and ARPK volatility in our data, as well as their correlations.

In the first two rows of Table 2, we report the TFP and ARPK standard deviation in our dataset. In the third row, we report their correlations; in the fourth row, we report the correlation of ARPK with lagged TFP. Column 1 reports these statistics using measures of TFP and ARPK that have been residualized of industry-by-year fixed effects. Like the prior literature, we find that TFP and ARPK are highly dispersed and positively correlated.

Because dispersion in TFP and ARPK can arise purely from permanent differences across firms, and do not necessarily reflect *fluctuations* in productivity (the focus of our paper), we further compute the dispersion of TFP and ARPK for the *within* component of these two measures (Column 2). To construct these measures, we further residualize TFP and ARPK using firm-fixed effects. More than half of the dispersion in TFP and ARPK can be classified as arising from volatility rather than permanent differences. In the third and fourth rows, we also find a substantial component of the joint co-movement that occurs within the firm.

Absent any frictions, the standard neoclassical model predicts that productivity and ARPK are uncorrelated, because ARPK is uniformly proportional to the user cost of capital (e.g., Hsieh and Klenow (2009)). Similarly, in a model with investment under uncertainty, but otherwise absent other frictions, lagged productivity is uncorrelated with ARPK, because any dispersion in ARPK is driven by innovations to lagged productivity. As such, a common interpretation of these facts (e.g., Restuccia and Rogerson (2008); David and Venkateswaran (2019)) is that firms face "correlated distortions", arising from investment frictions or policy-induced frictions that are associated with firm productivity. For instance, financial frictions induce positive co-movement between TFP and ARPK in standard models (e.g., Midrigan and Xu (2014)) because high-productivity firms also exhibit high returns to capital.

The key argument of our paper is that such correlations do not have to derive from distortions or frictions. Instead, we postulate a fraction of firm productivity fluctuations simply derive from variations in capital utilization, similar to the argument in the real business cycle literature. Because greater capital utilization implies higher rates of wear and tear, this in turn implies greater expenditure on maintenance per unit of capital, which raises the literal user cost of capital. In turn, using standard Jorgensonian intuition, the higher user cost implies a higher returns to capital. This implies that the co-movement between TFP and ARPK could simply be due to co-movement between utilization and maintenance.

The rest of Table 2 provides suggestive evidence of our hypothesis. From the fifth and sixth row, we see that the volatility of the within-firm component of capital utilization and maintenance rates are similar in magnitude to TFP and ARPK. Finally, in the last row, we also see that utilization and maintenance rates co-move as we hypothesize.

Our stylized facts here do not provide definitive evidence supporting our hypothesis. As such, for the remainder of this section, we will provide a sequence of analysis to buttress our claim that variations in utilization casually drive variations in TFP and ARPK.

B Fact 1: Utilization drives changes in firm TFP

We use the following equation to estimate the relationship between TFP and utilization:

$$\log \omega_{i,t} = \beta \log \tilde{u}_{i,t} + \gamma \log k_{i,t} + \alpha_i + \alpha_{j,t} + \varepsilon_{i,t}, \tag{4}$$

where α_i refers to fixed effects, $\alpha_{j,t}$ refers to industry $\times$ year fixed effects, and k refers to capital. Our hypothesis is that, controlling for physical productivity (i.e., z in our notation from Section A), utilization and ω should co-move one-for-one, reflecting equation 3. We use capital here as a proxy for physical productivity since we cannot directly observe productivity.

In Column 1 of Table 3, we report the OLS estimate of β using equation 4. We find that a 1% increase in utilization implies a 0.17% increase in TFP. While this result is promising, it is not fully supportive of our hypothesis; in particular, β is much lower than unity, which stands in contrast to equation 3, which would imply an elasticity of one.[5]

[5]To see this, equation 3 implies $\log \omega = \log z + \log \tilde{u}$; that is, holding fixed productivity, a 1% increase in utilization implies a 1% increase in productivity.

We postulate that this low elasticity is an artifact of attenuation bias driven by measurement error in capacity utilization, as we described in Section 2. As such, we next use an instrumental variables (IV) approach to address measurement error. The logic of the IV approach is that the component of correctly measured utilization is correlated with the instrument of interest, but the measurement errors themselves are uncorrelated with the instrument. As such, this would allow us to address the attenuation bias.

We consider two instruments. First, we use energy intensity, namely energy expenditure relative to capital, as an instrument. The idea here is that energy expenditures are correlated with utilization, since a firm that has higher capacity utilization must necessarily use energy more intensely in the short-run. Furthermore, energy expenditures are reported in audited income statements and are much less prone to measurement error. In principle, our strategy is similar to recent literature that has used measures of electricity usage and/or night-time ambient light emissions to correct for measurement error in self-reported GDP (e.g., Martínez (2022)).

Second, we utilize a "Hausman instrument" approach (Hausman, Leonard, and Zona, 1994; Hausman, 1996). Our instrument is the average of sales in the focal firm's industry-year cell, excluding the focal firm. The identification assumption is that changes in other firms' sales are correlated with the focal firm's utilization decisions but are unrelated to measurement error in the focal firm's reported utilization. Intuitively, the underlying assumption for this identification strategy is that demand across different firms is correlated, but measurement error is not. Appendix A.ii summarizes in greater detail our approach and construction of this instrument.

Our results, using the IV approach, are reported in Columns 2 to 4, where Column 2 reports the results using energy intensity as an IV, Column 3 reports results using the Hausman instrument, and Column 4 reports results using both instruments. We find estimates for β ranging from 0.91 to 1.15, and consistent with attenuation bias in the presence of measurement error, these estimates are all larger than Column 1. Furthermore, the estimates are also much closer to one. Even though we use two different IVs – likely to use differing variations in the data – it is reassuring that our estimates are very similar.

C Fact 2: Utilization drive variations in maintenance

We estimate the relationship between maintenance and utilization using the following equation:

$$\log\left(\frac{m_{i,t}}{k_{i,t}}\right) = \beta \log \tilde{u}_{i,t} + \gamma \log k_{i,t} + \alpha_i + \alpha_t + \varepsilon_{i,t}, \tag{5}$$

Here, our goal is to establish that exogenous increases in utilization rates increase the maintenance rates, with the underlying assumption being that more intense utilization leads to great wear-and-tear of capital, therefore necessitating more maintenance and repairs.

Similar to the earlier subsection, we first estimate the equation above via OLS. In Column 1 of Table 4, we find that a 1% increase in utilization rates imply a 0.199% increase in maintenance rates. While again supportive of our hypothesis, this simple correlation does not reflect causality. As such, we now turn to an IV approach.

We rely on three IVs, all building on the logic of a "Hausman instrument" as discussed in the previous section. Our instruments are the average of sales in a focal firm's industry-year cell, but excluding a subset of firms associated with the focal firm. Our first instrument excludes only the focal firm, while our second instrument excludes all firms within the focal firm's district. Our last instrument excludes firms within the focal firm's state.[6] Appendix A.ii provides more details on the construction of these IVs.

The relevance condition for the IV relies on firms in narrowly defined industries facing correlated demand in the same year. The exclusion restriction for the IV assumes that the only channel through which changes in other firm's sales affect the focal firm's maintenance decisions is through changes in the focal firm's utilization decisions. A violation of exclusion restriction can happen if a change in sales of other firms increases their demand for labor with maintenance expertise and drives up wages for labor with maintenance expertise. Then, the focal firm pays more for maintenance expenses not because of increased utilization but simply because maintenance is more expensive. If this is the case, we expect the effect to weaken substantially as we increase the excluded set of firms since labor markets are local.

Columns 2 to 4 report the respective results, showing that the elasticity of utilization to maintenance is large; a 1% increase in utilization rates imply between 1.674% to 1.817% increase in maintenance rates. Furthermore, estimates are similar when excluding geographic units with differing coverages, suggesting that the exclusion restriction is likely to hold.

[6]Firms in our dataset are in 218 districts and 27 states.

D Fact 3: ARPK and maintenance rates are positively correlated

To formally estimate the relationship between maintenance rates and ARPK, we estimate the following equation

$$\log ARPK_{i,t} = \beta \log(\frac{m_{i,t}}{k_{i,t}}) + \alpha_i + \alpha_{j,t} + \varepsilon_{i,t}. \tag{6}$$

Here, our goal is to establish that variations in maintenance behave like variations in user cost, which in turn drive variations in returns to capital.

Table 5 reports our estimation results using OLS. Column 1 reports our estimate with no fixed effects, column 2 with industry-by-year fixed effects, column 3 additionally with firm-fixed effects, and column 4 with state-by-year fixed effects. By and large, we find estimates of the elasticity of ARPK to maintenance rate to be around 0.58 to 0.61, and the estimates are all statistically significant at conventional levels.

While our results support our hypothesis — that is, maintenance proxies for capital user costs because it is a literal user cost — we note two points. First, our estimates could suffer from omitted variables bias. Nevertheless, because we include tight fixed effects, many alternative explanations do not apply: by including industry $\times$ year, time-varying industry level variation, such as allocation of credit subsidies to an industry, or time-varying state-level variation, such as changes in state tax rates, can not explain our findings.

Second, if we took maintenance rates as a literal measure of user cost, one would expect the elasticity of ARPK to maintenance to be 1%, given standard Jorgensonian intuition. The fact that our estimated elasticity is less than 1 suggests that maintenance only partially captures variation in user costs from the firm's perspective. In our model, we will rationalize this result by introducing a maintenance disruption cost, which raises the user cost of capital above its observed cost.

4 Model

In this section, we present our model of firm investment with endogenous capital maintenance and utilization. We use this framework to present key insights into how endogenous capital user costs—within an otherwise standard model of firm investment—can be used to jointly explain the four facts we presented.

A Households

Time is discrete and infinite. An infinitely lived representative household ranks streams of consumption and labor effort according to the following utility function:

$$U_0 \equiv \sum_{t=0}^{\infty} \beta^t \left(\log C_t - \chi N_t\right), \tag{7}$$

where C_t is an aggregate consumption bundle, N_t is labor supply, and $\beta \in (0,1)$ is the discount factor, and $\chi > 0$ is the labor disutlity parameter. C_t is a constant elasticity of substitution (CES) aggregator of a continuum of different varieties of goods

$$C_t = \left(\int_0^1 c_{it}^{\nu}\right)^{\frac{1}{\nu}}, \tag{8}$$

where i denotes a generic variety, and $\frac{1}{1-\nu}$ is the elasticity of substitution across goods. The budget constraint of the household is

$$\int_0^1 p_{it} c_{it} dj = w_t N_t + D_t, \tag{9}$$

where p_{it} is the price of the good i, w_t are wages, and D_t are aggregate dividends from ownership of all the firms in the economy. The usual cost-minimizing bundle gives the standard demand schedule

$$p_{it} = \left(\frac{c_{it}}{C_t}\right)^{-\frac{1}{1-\nu}} P_t, \tag{10}$$

where for the rest of the paper, we normalize the price index P_t to 1 (that is, the final output good is the numeraire).

Finally, the optimality condition for labor supply is given by

$$\chi C_t = w_t. \tag{11}$$

B Firms

Firms live forever, and discount future profits with a real interest rate of r. Output is given by a constant return to scale production function $y_{it} = z\left(u_{it} k_{it}\right)^{\alpha} l_{it}^{1-\alpha}$. k and l refer to capital and labor inputs respectively; u is the utilization rate of capital, with $u \in [0,1]$ (e.g., Greenwood, Hercowitz, and Huffman (1988); Burnside and Eichenbaum (1996)); z is physical productivity; and α is the capital share parameter. Firms internalize the demand schedule given by equation (10); therefore, the revenue function of firm i is given by

$$p_{it}y_{it} = C_t^{1-\nu}\left(z_{it}\left(u_{it}k_{it}\right)^{\alpha} l_{it}^{1-\alpha}\right)^{\nu}. \tag{12}$$

Finally, we assume that z evolves stochastically following an AR(1) process in logs:

$$\log z_{it} = \rho \log z_{i,t-1} + \sigma\epsilon_{it}, \quad \epsilon' \sim N(0,1).$$

Timing of factor inputs We assume that capital is pre-determined last period prior to observing z, but labor is chosen after z is observed. Therefore, for a given wage rate w and utilization rate u, the gross profit function (net of labor cost) can be written as

$$\begin{aligned}\pi\left(z_{it}, u_{it}, k_{it}\right) &= p_{it}y_{it} - w_t l_{it} \\ &= A\left(w_t, C_t\right)\left(z_{it}u_{it}^{\alpha}\right)^{\Theta_z}\left(k_{it}\right)^{\Theta_k},\end{aligned} \tag{13}$$

where A is a function of wage w and aggregate output C, $\Theta_z \equiv \frac{\nu}{1-(1-\alpha)\nu}$, and $\Theta_k \equiv \alpha\Theta_z$.[7] We defer discussing the role of capital utilization until we have described the law of motion of capital. Notice that based on equation 13, we can back firm total factor productivity ω as

$$\log\omega \equiv \log z + \alpha\log u \propto \frac{1-(1-\alpha)\nu}{\nu}\left(\log py - \frac{\alpha\nu}{1-(1-\alpha)\nu}\log k\right), \tag{14}$$

which motivates our formula given by equation 3 in Section 2.

B.1 Capital Accumulation

Capital stock at the firm level evolves according to the accumulation equation

$$k_{i,t+1} = \left(1-\delta(u_{it}, \tilde{m}_{it})\right)k_{it} + i_{it}, \tag{15}$$

where i is gross investment, $\delta(u, \tilde{m})$ is the depreciation rate of capital, and u is utilization. $\tilde{m}$ is the maintenance rate of capital, which we define as the ratio of maintenance services to capital stock (i.e., $\tilde{m} \equiv \frac{m}{k}$, where m denotes maintenance services in levels). As we explicitly spell out, capital depreciation is endogenous to the utilization and maintenance rate of capital; as such, the next period's capital stock is determined by the investment a

[7]The derivation for labor demand is presented in Appendix OA.3

firm makes in the current period and how intensely a firm utilizes and maintains its capital stock. As is standard, we assume that i, m, and u are chosen with perfect information over z_{it}, but before $z_{i,t+1}$ is observed.

Following Caunedo and Keller (2021), we parameterize the depreciation function as

$$\delta\left(u_{it}, \tilde{m}_{it}\right) = u_{it} \exp\left(-\frac{\tilde{m}_{it}^{\zeta^m}}{u_{it}^{\zeta^u}}\right), \tag{16}$$

where ζ^m and ζ^u separately control the degree to which capital depreciation is affected by maintenance and utilization. Our formulation implies that capital depreciates faster if utilized relatively more intensively than maintained.

Next, we present the costs of investment and maintenance that will determine i, m, and u.

B.2 Maintenance Costs

We assume that every unit of maintenance costs p^m units of labor; therefore, for m unit of maintenance services purchased, the firm pays $wp^m m$ in units of the output good. In addition, we assume that conducting maintenance is disruptive to firm production. For instance, a farm mill might shut for a few days for maintenance, disrupting its milling activities. We parameterize these disruption costs as follows

$$\mathcal{C}\left(\tilde{m}_{it}\right) u_{it} k_{it} = \frac{\theta}{2} \tilde{m}_{it}^2 u_{it} k_{it}, \tag{17}$$

where the parameter θ captures the severity of the disruption cost, and the cost scales proportionally with utilized capital. Like with the flow cost of maintenance services, we define the disruption costs in units of labor.

We interpret θ as determining the firm's flexibility in doing maintenance; therfore, the higher the value of θ, the less flexible it is for the firm to do maintenance as the disruption cost is larger. From an empirical perspective, this allows us to rationalize why maintenance rates do not necessarily co-move one-for-one with arpk.

B.3 Capital adjustment friction

We assume that firms only face investment irreversibility as a capital adjustment friction. To model investment irreversibility, we assume that for a firm purchasing capital, the price of capital is one unit of the output good, whereas, for a dis-investing firm, the price of capital is $\lambda \in [0, 1]$ units of the output good. For notation, we denote this adjustment cost by

$$Q(i_{it}) = \begin{cases} 1 & i_{it} \geq 0 \\ 1-\lambda & i_{it} < 0 \end{cases}$$

such that the total cost of investment (or benefit from disinvestment) is $Q(i)i$.

While one can, in principle, model a complete set of standard adjustment frictions (e.g, convex adjustment costs, fixed costs), we find in practice, as well as in our external validation exercises, that investment irreversibility is the dominant adjustment friction.[8]

C Bellman equations

We now summarize the firm's dynamic programming problem. Let $V(z,k)$ denote the value function of a firm with productivity z and physical capital stock k, and let "prime" variables denote next-period variables. The Bellman equation is,

$$\begin{aligned} V(z,k) = \max_{i,m,u}\ & \pi(z,u,k;w) - Q(k',k)\, i - wp^m m - w\frac{\theta}{2}\tilde{m}^2 uk + \frac{1}{1+r}\mathbb{E}\left[V(z',k')\,|z\right] \\ & s.t. \\ & k' = i + (1-\delta(u,\tilde{m}))\, k. \end{aligned}$$

Note that because we focus on a stationary equilibrium, the risk-free rate is constant and equal to $\frac{1}{\beta} - 1$.

D Equilibrium

A stationary equilibrium in this economy is (i) a set of value and policy functions for the firm (V,$\tilde{m}$,u,k',l), (ii) a wage rate w, and (iii) a joint distribution over z and k such that taking wages as given, (a) the value and policy functions solve the firm's optimization problem, (b) the output and labor markets clear, and (c) the distribution is invariant over time.

E Model Intuition: Understanding Facts 0 to 3

We now briefly discuss some key model intuition into how our model rationalizes the stylized facts in Section 3. To most succinctly present the intuition, consider the simpler

[8]This is also consistent with earlier work by Lanteri, Medina, and Tan (2023) who study investment frictions in a similar economic environment.

model where firms face no capital adjustment cost (i.e., $\lambda = 0$) nor maintenance disruption costs (i.e., $\theta = 0$). In that case, optimal capital utilization implies the following first order condition,

$$\Theta_k \frac{Az^{\Theta_z} (uk)^{\Theta_k}}{k} = u \frac{\partial \delta}{\partial u}. \tag{18}$$

Notice then that this implies that u is an increasing function of z because $u\frac{\partial \delta}{\partial u}$ is increasing in u given our parametric assumption of δ. As such, since firm total factor productivity $\omega \equiv zu^{\alpha}$, trivially, we obtain Fact 1.

For Fact 2, note that optimal maintenance implies the following first-order condition,

$$p^m = -\frac{\partial \delta}{\partial \tilde{m}}. \tag{19}$$

Since our parametric assumption on δ implies $-\frac{\partial^2 \delta}{\partial \tilde{m} \partial u} > 0$, equation (19) further imply that maintenance and utilization are complements. Consequently, increases in utilization rates are associated with increases in maintenance.

Finally, for Fact 3, notice that the term on the left-hand side in equation (18) is proportional to the average revenue product of capital. Trivially, then, ARPK is increasing in u. Because maintenance is also increasing in u, we obtain that ARPK and maintenance rates are positively correlated.

Combining the discussion, we see that firm TFP and ARPK are correlated due to their respective association with capital utilization and maintenance (i.e., Fact 0). Importantly, notice that our argument so far has not required the presence of *any* frictions. However, if capital utilization and user costs are fixed (as is standard), then this logic breaks down because capital is fixed in the short-run and equation (18) does not hold. Consequently, absent frictions, firm ARPK is not auto-correlated, and neither is it correlated with lagged productivity.

A key point we therefore make here is that, in a model with endogenous user costs, the ability to vary the intensity of capital utilization effectively converts capital from a quasi-fixed factor to a flexible factor. This effect allows the model to replicate the facts we report without needing to additionally model investment or capital adjustment frictions (e.g., Asker, Collard-wexler, and Loecker (2014)), nor the need to model correlated distortions (e.g., Restuccia and Rogerson (2008); David and Venkateswaran (2019)). Crucially, this is not simply a casual point of interest. The conversion of capital to a flexible input effectively negates the negative impact of uncertainty and capital adjustment frictions, but simultaneously increases the dispersion of ARPK through equation 18 because utilization becomes more volatile. In other words, our model allows for the possibility that increases in the dispersion of ARPK is associated with a simultaneous increase in aggregate productivity. A key part of our quantitative exercise will be to explore this result.

5 Quantifying the Role of Endogenous Use Costs

In this section, we first present additional empirical evidence as support for our key model mechanisms. In particular, we show additional evidence highlighting the role of utilization as a driver in the dispersion of ARPK, and the role of investment irreversibility for our quantitative model. Then, we present our calibration and strategy, and validate the properties of the model.

A Additional supporting empirical evidence

A.1 The volatility and autocorrelation of "utilization-adjusted" ARPK is smaller than uncorrected ARPK

In our model, one can show that the volatility and autocorrelation of "utilization-adjusted" ARPK, which we define as log ARPK net of log utilization, is smaller than uncorrected ARPK.

To see this, equation (18) shows that volatility in ARPK in entirely driven by volatility in the term $u\frac{\partial \delta}{\partial u}$. In turn, utilization-adjusted ARPK is simply the term $\frac{\partial \delta}{\partial u}$. Since $\frac{\partial \delta}{\partial u}$ is increasing in u, it must be that $var\left(\frac{\partial \delta}{\partial u}\right) < var\left(u\frac{\partial \delta}{\partial u}\right)$; consequently, it must be that the volatility of utilization-adjusted ARPK is lower than that of unadjusted ARPK.

A similar logic explains why utilization-adjusted ARPK has lower autocorrelation: ARPK is autocorrelated because the term $u\frac{\partial \delta}{\partial u}$ is directly correlated with z, which itself is autocorrelated. Therefore, by adjusting ARPK with u, we mechanically reduce a fraction of the autocorrelation in ARPK.

However, operationalizing this adjustment in the data is difficult because our direct measure of capacity utilization suffers from severe measurement error. We surmount this challenge in three ways. First, we use the logic of our IV and use the predicted value of utilization by regressing utilization on energy intensity and using the predicted values to generate "corrected" utilization rates. Second, we proxy for utilization rates using maintenance rates; in other words, we simply subtract the log of maintenance rates from log ARPK, instead of utilization. This follows from our model that maintenance and utilization are positively correlated, and would be an exact adjustment if maintenance and utilization were log-linearly related. Finally, we proxy for utilization rates using energy intensity, following the same logic.

Table 6 reports our results, showing that our prediction is supported by the data. In general, the dispersion of ARPK is lower by between 12% to 27%, while the autocorrelation is lower by between 1% to 17%.[9]

A.2 The joint dynamics of maintenance and investment

In our framework, we modeled investment irreversibility but abstracted away from vintage effects in the capital stock. Because vintage effects are a potentially important driver of maintenance and investment decisions and capital productivity (e.g., Caunedo and Keller (2021)), we now present a discussion of this assumption, and show that the empirical evidence is supportive of our assumption.

In Panel A of Figure 1, we plot the average investment rate of a firm in a window of -5/+2 years around an investment spike. We define an investment "spike" as the largest investment rate recorded for a firm across all years in the sample. Our results show that investment is "lumpy", and consistent with a wide array of earlier research (e.g., Doms and Dunne (1998)). Lumpy investment, in turn, implies non-convexities in the capital adjustment cost function. This could arise because of investment irreversibility (as in our model), or vintage upgrading (where spikes correspond to a large replacement of equipment).[10]

In Panel B of Figure 1, we now plot the average maintenance rates around the same window of investment spike. Here, we see that the maintenance rate rises in the years before the spike, peaks at the spike, and then falls with investment after the spike.

We argue that our parsimonious model is consistent with this evidence, whereas an alternative model of *only* vintage effects is not. For ease of exposition, assume that firms do not face any maintenance disruption costs (i.e., $\theta = 0$). Then, the maintenance indifference condition can be written as

$$-q^{marg}\frac{\partial \delta}{\partial \tilde{m}} = p^m, \tag{20}$$

where $q^{marg} \equiv \frac{1}{1+r}\frac{\partial \mathbb{E}[V(z',k')|z,k]}{\partial k'}$ denotes marginal q.[11] Furthermore, note that q^{marg} satis-

[9] In Appendix E, we report in Table OA.3 a comparison of our calibrated model's prediction of this exercise relative to the data. We also note that attenuation in the estimated autocorrelation is not strictly support for our mechanism, since measurement errors alone in utilization, maintenance, or energy would mechanically attenuate the estimate for the AR(1) parameter.

[10] Trivially, this would also reject the role of convex adjustment costs.

[11] Appendix B presents a derivation.

fies the following properties: (i) $q^{marg} = 1$ when the firm is investing or indifferent between investing and inaction, (ii) $q^{marg} = 1 - \lambda$ when the firm is disinvesting or indifferent between disinvesting and inaction, and (iii) q^{marg} varies smoothly and monotonically between 1 and $1 - \lambda$ when the firm is in the inaction region.[12] Since the depreciation rate is decreasing in maintenance, maintenance rates are highest when investing, decreasing in capital when in the inaction region, and finally lowest when disinvesting. Figure OA.2 in Appendix E plots an illustrative plot using parameters from our calibrated model.

Importantly, this implies that for as many years as a firm is in the inaction region (i.e., prior to the investment spike), it must be that maintenance rates are rising. Furthermore, because investment spikes in models of investment irreversibility happen solely due to unexpected positive productivity shocks, this implies that marginal q jumps to 1 during a spike; consequently, maintenance rates peak during an investment spike. Figure OA.3 plots an illustrative plot using our calibrated model.

In contrast, models with vintage effects typically imply that the scrap value of the current vintage is lower than the value of new capital (we present a simple model of vintage upgrading in Appendix C to buttress this point). As such, this implies that maintenance rates will fall when investment spikes, which is inconsistent with our evidence.

We emphasize that our findings do not necessarily reject a role for vintage effects. Instead, our argument is that a relatively standard capital adjustment cost, in the form of investment irreversibility, is able to rationalize the empirical results; and as such, we choose a parsimonious approach to our model.

Finally, while the discussion so far focuses on vintage models with endogenous maintenance, a competing story is that maintenance can be exogenous to the vintage. For instance, the maintenance cost of a capital stock could be increasing in age; in turn, maintenance expenses would peak the year new replacement capital is purchased, but not yet installed, thus generating a path of maintenance and investment similar to our event study.

Indeed, our event study does not reject this hypothesis, and exogenous increases in maintenance could indeed be driving this finding. However, purely exogenous movements in maintenance would be counterfactual to the empirical evidence we documented in Section 3; namely, that exogenous variations in utilization, driven by demand shocks, drive maintenance decisions. That said, we recognize that our event study is, more likely than not, arising from both endogenous and exogenous maintenance. As such, in our calibration strategy, we focus on matching the elasticity of maintenance to various pro-

[12] Caballero (1999) presents an intuitive proof of this result, while Baley and Blanco (2022) presents a proof in continuous time. We re-derive the proof in the context of our model in Appendix B.ii.

ductivity measures—akin to how one calibrates capital adjustment frictions by matching the elasticity of investment to productivity—rather than simply matching the total variation in maintenance rates.

B Model calibration

Building on our evidence, we now describe our calibration strategy. We follow a standard approach, by externally fixing some parameters to standard values used in the literature, and internally calibrating the rest to informative statistics from the data.

B.1 Externally calibrated parameters.

We consider a calibration with an annual frequency, consistent with our data. We set β to 0.96, consistent with an interest rate of 4%. The labor disutility is set to 1.45, to give us an average of 8 hours worked per day. Finally, the elasticity of substitution and capital intensity parameters, which we set to 3/4 and 1/3 respectively, following standard assumption. The values are summarized in Panel A of Table (7).

B.2 Internally calibrated parameters.

For the remaining parameters, we calibrate them to match salient moments of our data. We discretize the productivity process with 41 nodes using the method of Tauchen (1986). Then, we use a method of moments procedure to internally calibrate the remaining seven parameters ρ, σ, λ, p^m, ζ^m, ζ^u, and θ. For our calibration exercise, we minimize the sum of squared deviations of the seven moments reported in Panel B of Table 7.

We briefly discuss here our choice of moments, as some of these depart from conventional assumptions. Note that while we associate each parameter to a specific moment, we report this mapping with the understanding that the parameters are all jointly calibrated and that varying any single parameter will lead to changes in all other moments.

First, for ρ and σ, we target the autocorrelation and volatility of measured firm TFP from equation 3 after accounting for firm and industry-by-year fixed effects. As we discussed earlier, unlike prior literature where these statistics would be a direct input into the model (e.g., Asker, Collard-wexler, and Loecker (2014); David and Venkateswaran (2019)), our explicit treatment of capital utilization implies that the TFP conflates both actual physical productivity and utilization, and thus necessitates an indirect inference approach.

Second, for λ, we target the net disinvestment rate. There is a natural mapping where the median net disinvestment rate is increasing in λ, since the more irreversible investment is, the smaller the rate of disinvestment. That said, this departs from the conventional approach of targeting the average fraction of firms that report a gross disinvestment (e.g., Cooper and Haltiwanger (2006); Lanteri, Medina, and Tan (2023)). Our rationale here is similar to that for the calibration of ρ and σ—because depreciation rates are endogenous, there is no clear definition of gross disinvestment. As such, we choose a moment that does not force us to pick a specific value for economic depreciation.

Finally, for p^m, ζ^m, ζ^u, and θ, we target the median maintenance rate, median utilization rate, and the elasticity of maintenance rates to ω and $ARPK$. Our focus on matching the elasticity of maintenance rates to ω and $ARPK$ follows from our discussion earlier, where we emphasize that our goal is to replicate the variation of maintenance that are associated with productivity.

Table 7 shows that our model predicts a fit that is reasonably close to the data. One important outcome of the calibration exercise is that we find that the volatility of physical productivity is about 36.5% lower than the volatility of measured firm TFP, implying that firms face much lower uncertainty in our model compared to a alternative model that directly uses the volatility of firm TFP as a model primitive. We will explore the implications of this result, in more detail, in Section 6.

Finally, Panel (C) of Table 7 presents a set of untargeted moments for external validation. Given our emphasis on explaining jointly maintenance and investment behavior, we report in the first two rows the elasticity of net investment rates to maintenance rate and firm TFP respectively. We find that our model predictions are reasonably close to the data; it slightly overestimates the association of investment with maintenance, but underestimates the association of investment with productivity.

We further examine the model's fit with respect to the share of maintenance as a fraction of firm costs and output. In our model, we assume that maintenance costs, both the flow cost and disruption cost, are paid in terms of labor units. We examine our model's prediction in the context relative to the data, and find a reasonably close fit (third row); in the data, the average maintenance share of labor cost is approximately 15.9%, while it is 13.5% in the model. Along a similar vein, we report in the fourth role the average maintenance share of value added. We find that this is approximately 3.6% in the data but around 7.4% in the model.

In the third last row, we report our model's prediction with respective to the median depreciation rate. In our model, we find that the median firm has a depreciation rate of 7%, which is within the typical range of values used in calibrated macroeconomic models (typically between 6% to 10%). In the data, we find that the median accounting depreciation rate is around 9.7%. While the numbers are similar and within the typical range of values, it is important to keep in mind that accounting depreciation does not necessarily map to economic depreciation.

Next, in the second last row, we report our model's prediction of the dispersion of (log) maintenance rates.[13] Relative to the data, our model predicts about one-third of the total dispersion of maintenance rates. As we hypothesized, exogenous increases in maintenance due to capital age is likely to be a simultaneous driver of maintenance decisions, which we do not model in this paper. This would rationalize the much larger variation of maintenance rates; in addition, it would explain why investment rates co-vary less with maintenance in the data relative to the model.

Finally, in the last row, we report our model's prediction for the dispersion in ARPK. We find that model generates a dispersion in ARPK that is approximately 75.5% of the dispersion in the data. We note that this is a much closer match to the data than typical models of firm investment dynamics that abstract from endogenous user costs. While not a main point of our paper, this suggests a smaller role for correlated distortions in driving the dispersion of ARPK relative to prior estimates.

6 Quantitative Results

We now explore the quantitative implications of our model. First, because the allocation of capital and distribution of ARPK is directly associated with the distribution of user costs, we explore the implications of our model for the dispersion of APRK and aggregate total factor productivity, relative to a fixed user cost model. Second, we explore the role of the maintenance disruption cost, a dimension of firm investment dynamics that has not been studied. Finally, we examine the implications of ignoring the role of endogenous utilization as a driver of fluctuations in firm productivity, in the context of calibrating an otherwise equivalent model of firm investment dynamics.

A Capital allocation in a model of endogenous user cost

We now begin by presenting our first result: In a model of endogenous capital user cost, increases in the dispersion of ARPK can be associated with increases in aggregate TFP. This arises because endogenous utilization effectively converts capital from a quasi-fixed asset to a flexible input.

[13] We report the dispersion in logs to avoid mechanical scaling effects from the price of maintenance.

To fix ideas, we first note that for any two comparison models in our framework, the log difference in aggregate TFP is can be expressed as

$$\Delta \log TFP = \frac{1-(1-\alpha)\,\nu}{\nu}\left\{\log\left(\mathbb{E}_b[u^{\Theta_k}]\mathbb{E}_b[z^{\Theta_z}k^{\Theta_k}] + cov_b\left(u^{\Theta_k}, z^{\Theta_z}k^{\Theta_k}\right)\right) ..\right.$$
$$\left. .. - \log\left(\mathbb{E}_c[u^{\Theta_k}]\mathbb{E}_c[z^{\Theta_z}k^{\Theta_k}] + cov_c\left(u^{\Theta_k}, z^{\Theta_z}k^{\Theta_k}\right)\right)\right\}, \tag{21}$$

where we use the subscripts b and c to denote "baseline" and "counterfactual", respectively. The subscripts are also appended to the expectation and covariance operators to denote the idea that the joint distribution of productivity, capital, and utilization are different in both models.[14]

This expression reveals that differences in aggregate TFP is driven by three components. First, it is directly affected by aggregate utilization rates (i.e., $\mathbb{E}[u^{\Theta_k}]$), as in Greenwood, Hercowitz, and Huffman (1988) or Burnside and Eichenbaum (1996). Second, it is affected by the joint distribution of productivity and physical capital ($\mathbb{E}[z^{\Theta_z}k^{\Theta_k}]$), as in Restuccia and Rogerson (2008) and Hsieh and Klenow (2009).

Finally, it is also affected by the degree to which utilization introduces flexibility into the capital stock, an otherwise quasi-fixed asset. This effect is captured in the last term $cov\left(u^{\Theta_k}, z^{\Theta_z}k^{\Theta_k}\right)$: A positive covariance implies that utilization allows firms that have "too little" ("too much") capital relative to productivity to overcome their short-run constraint by utilizing capital more (less) intensely. It is important to point out that the covariance term is endogenous to the model structure. For instance, one can show that if capital was a purely flexible input, then the covariance term is zero; if capital is already a flexible input, there is simply no room for "more" flexibility.[15]

This expression implies that endogenous user costs have two direct effects on aggregate productivity through a pure allocation effect: (1) It can reshape the distribution of physical capital, and (2) it can reshape the distribution of utilized capital.

We decompose these channels as follows. First, we solve an otherwise equivalent model where user costs are fixed and all firms utilize capital at the same average utilization as in the baseline model. Specifically, we impose that firms always face a depreciation rate and maintenance cost (per unit of capital) equivalent to the average depreciation and maintenance cost in the model, and utilize capital at the rate $\left(\int u^{\Theta_k} d\Lambda(z,k)\right)^{\frac{1}{\Theta_k}}$). This counterfactual allows us to compute the total effect of endogenous user costs on the allocation of capital and thus aggregate productivity, while muting the mechanical effect of

[14] Appendix E.iii derives this decomposition.
[15] Appendix D derives this result.

average utilization driving aggregate TFP due to Jensen's inequality.[16]

Second, starting from the baseline model, we compute aggregate productivity using the same distribution of capital and average utilization, but imposing that the covariance term is 0. This isolates the effect utilization has in converting capital from a quasi-fixed input to a flexible input. Using the notation of equation 21, we essentially impose that $\mathbb{E}_c[\cdot] = \mathbb{E}_b[\cdot]$ and $cov_c(\cdot) = 0$.

Table 8 reports our decomposition. The first row reports aggregate TFP in each counterfactual relative to our baseline model, in log differences. In the second column, we see that muting the role of endogenous user costs leads to an economy with TFP of around 4.0% lower than our baseline; while in the third column, we see that imposing a zero covariance term while keeping the distribution of capital fixed leads to aggregate TFP that is around 5.2% lower than our baseline.

These results are revealing of two important roles endogenous user costs and utilization play in shaping aggregate productivity and the distribution of capital. First, it indicates that the conversion of capital from a quasi-fixed factor to a flexible input can greatly increase aggregate productivity, mitigating the effects of short-run constraints in physical capacity. However, it also shows that allowing for endogenous user costs in fact *worsens* the distribution of physical capital; absent the covariance effect, the baseline economy would have aggregate productivity that is 1.2% lower relative to an economy with fixed user cost (i.e., the difference between the second and third columns).

Why does allowing for endogenous user costs give rise to a worse allocation of physical capital? Panel (A) of Figure 2, which plots the investment policy function for a low- and a high-productivity firm respectively, is immediately revealing. Compared to a firm facing fixed user costs, the low productivity firm tends to hold too much capital, and is less like to disinvest. Conversely, a high productivity firm buys less capital, and is less willing to invest.

These results arises because endogenous user costs reinforce the effect of real options coming from investment irreversibility. As is standard, investment irreversibility implies that low productivity firms disinvest at lower rates, while high productivity firms invest at lower rates. In our case, endogenous user costs reinforce this effect in two ways. For low productivity firms, the lower expected future user cost (coming from lower utilization and maintenance) imply that firms are even less willing to disinvest; while for high productivity firms, the higher expected user cost implies that they are less willing to invest. Panel (B) of Figure 2 illustrates the expected user costs for a low- and high-

[16]Specifically, if we had simply impose a utilization rate of $\int u d\Lambda$, Jensen's inequality would imply that $\left(\int u d\Lambda\right)^{\Theta_k} > \int u^{\Theta_k} d\Lambda$ and the counterfactual would mechanically exhibit higher TFP.

productivity firm. In Figure 3, we further see this effect translated to the aggregate distribution of capital: Relative to the fixed user cost model, the distribution of capital is less dispersed. These results imply that endogenous user costs amplify the misallocation effect of investment irreversibility, which also tends to compress the distribution of capital.

We now examine the implications of endogenous user costs for the dispersion of ARPK. Conventional insight building on models of fixed user cost (e.g., Hsieh and Klenow (2009); David and Venkateswaran (2019)) argue that increases in the dispersion of ARPK is strictly associated with losses in aggregate productivity. However, looking to the second row of Table 8, we see a stark result: The dispersion of ARPK in our baseline model is substantially larger than that in either counterfactual models, even though aggregate productivity is highest in the baseline model. For instance, the standard deviation of ARPK in our model is around 0.496 , whereas muting the user cost channel reduces this by over 40% to 0.285.

Further analysis reveal the following: Looking to the third column, we see that the bulk of the amplification in the dispersion of ARPK comes from the covariance term; simply muting the covariance channel reduces the standard deviation of ARPK by 35%. In other words, while flexibility in adjusting the effective capital stock improves aggregate productivity, it also simultaneously increases the dispersion of ARPK. In fact, this outcome arises directly from the indifference condition for utilization (equation 18): ARPK is directly tied to the dispersion of utilization rates.

In contrast, comparing the second and third columns, we find the more conventional result that higher dispersion in ARPK is associated with capital misallocation; in this case, mirroring our result that the distribution of physical capital is actually worse relative to the fixed user cost counterfactual. Taken in totality, our results point towards a key conclusion—that dispersion in ARPK is "good" when it is associated with flexibility in static "reallocation" of a quasi-fixed asset, but bad when it directly changes the distribution of physical assets itself. Our quantitative analysis suggests that the "good" effect strictly dominates the "bad" effect by orders of magnitudes.

B The role of maintenance disruption costs

In this subsection, we study the role of maintenance disruption costs in shaping aggregate productivity, a dimension that has not been considered in prior literature. In our model, maintenance disruption costs lower aggregate productivity because they effectively make utilization more costly, while also reducing firms' flexibility to "adjust" the effective capital stock. We examine the quantitative importance of these disruption costs.

We solve a counterfactual that imposes $\theta = 0$ but keeps all other parameters the same as our baseline values. Column 1 of Table 9 reports the main result: If India faced no maintenance disruption cost, aggregate TFP would be 13.6% higher. This is much larger relative to estimates of the contribution of capital adjustment frictions to aggregate productivity losses, both in our context and the broader literature. For instance, in our calibration, we find that investment irreversibility alone lowers aggregate TFP, relative to our baseline economy, by about 0.1%. Similarly, recent research (e.g.,David and Venkateswaran (2019)) find that capital adjustment frictions lower aggregate TFP by about 1% in China and 2% in the US.[17]

What drives this large gain to aggregate TFP? Using the logic of equation 21, we decompose the contribution of average utilization rates, flexibility (i.e., the covariance term), and the distribution of capital in generating the difference in aggregate TFP between the baseline model and the counterfactual.

To construct our decomposition, we utilize the logic of Shapely decomposition, by starting from the $\theta = 0$ model, and then computing the partial contribution of each channel. This necessitates a framework to "shut down" each channel. To do that, we utilize equation 21 to guide our decomposition. To "shut down" the covariance channel, we replace the covariance term in the $\theta = 0$ model with the covariance term from our baseline model; to "shut down' the capital distribution channel, we replace the distribution in the $\theta = 0$ model with that from the baseline model; and finally, to "shut down" the average utilization channel, we simply replace the utilization term with average utilization from the baseline model. The full decomposition then uses a permutation of various "submodels" to compute the partial effect of each channel. More details of the decomposition are provided in Appendix E.iii.2.

The rest of the columns in Table 9 summarizes the decomposition. In Column 2, we see that 11.8 percentage points (i.e., around 87% of the gains) come purely from an increase in average utilization rates. Indeed, in the counterfactual, we find that average firm level capacity utilization rises from around 69% in the baseline to 78% in the counterfactual. Along a similar vein, aggregate capacity utilization, computed as the actual total physical output aggregated using the CES index, relative to the output aggregated using the CES

[17]To be precise, we simply solve our model while setting $\lambda = 0$, and compared the predicted aggregated TFP between this and our baseline model. This is different from the approach of David and Venkateswaran (2019), who compute TFP losses relative to a neoclassical model with fixed user costs. The very low losses attributable to investment frictions in our framework of endogenous user costs reflects the insight from the earlier subsection; namely, that utilization allows firms to "bypass" investment frictions.

index if utilization was uniformly 1, increases from 72.5% to 83.1%.[18]

Next, in Column 3, we see that flexibility contributes about 1.65 percentage point. While relatively small in the context of our model, the magnitude is similar to the contribution of capital adjustment frictions to capital misallocation. Finally, in Column 4, we see that capital allocation is slightly better in the economy with no maintenance disruption; it contributes 0.11 percentage points to aggregate productivity.

Taken together, our results suggest that a large gain for productivity in the Indian economy could arise solely from addressing the role of maintenance disruption, where the key channel comes through an increased utilization rate. For comparison, using data from FRED, we find that aggregate capacity utilization in the US is approximately 79% on average for a comparable sample period, similar to the 83% predicted by our model when $\theta = 0$. Taking our results at face value, this suggests that a large fraction of the gap in productivity between the US and India could be closed by simply addressing inefficiencies in maintenance.[19]

From a practical perspective, what might be a source of such inefficiencies? While this is outside the scope of our paper, we propose two possibilities. First, to some degree, our findings mirror that of Bloom, Eifert, Mahajan, McKenzie, and Roberts (2012), who contend that improving management practices to that of standards in the United States could improve average productivity by around 17%; in particular, one key management practice they propose is the "regular maintenance of machines...".[20] Our reduced form estimate of θ could in part be capturing this dimension of difference.

Second, we also contend that firm size differences in US and India could be a leading driver. One potential reason why firms in the US might suffer from less inefficiencies in maintenance might simply be because they are substantially larger; maintenance is less disruptive because they can simply rotate their maintenance schedules around multiple plants. This logic is similar to the hypothesis by Kehrig and Vincent (2023), who argue that multi-plant firms are relatively more productive because they can rotate production around their plants. In contrast, maintenance becomes extremely disruptive if a firm is small and only has a small number of machines.

[18] To be precise, let $y(u)$ denote physical output when utilization is endogenous, and $\tilde{y} = y(1)$ when utilization is set to 1. Then, we compute the aggregate capacity utilization as $\frac{(\int y(u)^{\nu})^{\frac{1}{\nu}}}{(\int \tilde{y}^{\nu})^{\frac{1}{\nu}}}$.

[19] For context, Hsieh and Klenow (2009) contend that *fully* switching out the distribution of policy distortions in India to that of the US would increase aggregate TFP in manufacturing by around 40% to 60%. Our results suggest that simply closing the gap in one single dimension, maintenance decisions, could generate a quarter of total possible gains.

[20] See Section III.B of their paper, which delineates five key areas of management practices, of which maintenance is one key area.

C Implications for Capital Reallocation Frictions

Our last exercise examines the role of endogenous user cost in estimating the size of investment irreversibility, that is, the degree to which capital reallocation is frictional. While it is now well-documented that substantial capital reallocation frictions exists, and have important impact on aggregate productivity in developing economies, the literature often sees a large range in the estimates of capital reallocation frictions. For instance, earlier research by Lanteri (2018) propose a resale loss of around 6.7% for firms in the United States, whereas Lanteri, Medina, and Tan (2023) propose a resale loss of almost 41% for firms in Peru. We examine whether these differences can be partly reconciled by accounting for endogenous user costs.

To that end, we calibrate a model of fixed user cost with investment irreversibility, targeting a subset of moments that are relevant to this model. This amounts to simply setting ρ and σ directly to the estimated autocorrelation and volatility of firm TFP (consistent with standard approaches), and then varying λ to re-target the median disinvestment rate. We find that when we abstract from endogenous user costs, the resale loss rises to almost 40%. In contrast, recall that our baseline value for resale loss was around 8.7% . In other words, accounting for endogenous user costs leads to estimates for resale losses in India that are remarkably similar to that for the United States.

This large increase in inferred resale loss is not surprising, and arises simply because the volatility of productivity is almost 36.5% higher in this counterfactual. Because the dispersion of investment rates is mechanically larger, a correspondingly larger degree of irreversibility is needed to match the same target. Indeed, our finding is consistent with prior literature such as Lanteri, Medina, and Tan (2023), who calibrate a similar firm TFP process and obtain resale losses of around 41%. To some degree, our results suggests that, to the extent one wishes to quantify the degree of capital reallocation frictions in India, accounting for endogenous user costs might be quantitatively important.

7 Conclusion

In this paper, we direct empirical evidence that a component of the variation in firm-level total factor productivity is explained by variations in capital utilization rates. Moreover, we show that variations in firm-level ARPK is highly associated with variations in capital user costs in the form of maintenance rates. Furthermore, we argue that our stylized facts support a model with endogenous capital utilization, user costs, and maintenance disruption costs that weaken a firm's response to positive productivity shocks.

We examine the quantitative implications of endogenous capital utilization and user costs in an otherwise standard model of firm investment dynamics with investment irreversibility. Through our calibrated model, we find that endogenous user costs amplify the dispersion of ARPK while simultaneously increasing aggregate total factor productivity, where both outcomes arise jointly due to the role of endogenous capital utilization, converting capital from an otherwise quasi-fixed factor to a flexible factor. We also find that maintenance disruption costs have a significant negative drag on the Indian economy; a complete removal of maintenance frictions could lead to a substantial 13.6% improvement in Indian aggregate total factor productivity. Finally, we also showed that in the context of India, abstracting from endogenous user costs would have led to an over-estimate of the size of capital resale frictions by almost five-fold.

8 Figures and Tables

A Tables

Table 1: **Summary Statistics**

	Mean	Std. Dev.	p25	Median	p75
capital (k)	43.932	146.001	1.778	5.695	21.379
value added (y)	67.378	204.858	3.184	10.684	37.535
net investment $(i = k' - k)$	1.394	15.371	-0.709	-0.099	0.461
net investment rate $(\frac{i}{k})$	0.130	0.661	-0.111	-0.049	0.092
log measured productivity $(\log \omega)$	0.992	0.861	0.509	1.039	1.546
ARPK $(\frac{y}{k})$	3.503	5.570	0.911	1.774	3.644
maintenance (m)	1.405	4.478	0.051	0.188	0.732
maintenance rate $(\frac{m}{k})$	0.062	0.081	0.016	0.035	0.073
maintenance to value added $(\frac{m}{y})$	0.030	0.036	0.010	0.020	0.037
capacity utilization* $(\tilde{u})$	0.663	0.695	0.325	0.548	0.801
Number of firm-year observations	61,803				

This table presents the summary statistics for the main sample from 1989 to 2019. Capital (k) is measured as the value of PPE. Value added (y) is the difference between sales and cost of material input. Investment (i) is net investment and is measured as the differences in PPE. Maintenance (m) is the repair and maintenance item from the income statement. $\log \omega$ is measured logarithm of TFP and is defined as a Solow residual $\log \omega = \frac{1}{\Theta_z}(\log py - \Theta_k \log k)$, and $\log ARPK = \log y - \log k$. Utilization is reported as physical capacity utilization, and is only available for a subset of the sample. Where relevant, the values are reported in Millions of US dollars in 2019.

Table 2: **Volatilities and Autocorrelations Within and Across Firms**

Variable	Total	Within	Fraction Within
$\log \omega$	0.754	0.405	53.719
$\log ARPK$	1.038	0.695	66.974
$\beta(\log ARPK, \log \omega)$	0.978	1.489	.
$\beta(\log ARPK_t, \log \omega_{t-1})$	0.810	0.995	.
$\log u$	1.035	0.676	65.284
$\log \frac{m}{k}$	1.079	0.714	66.162
$\beta(\log \tilde{u}, \log \frac{m}{k})$	0.166	0.182	.

For the rows with variables $\log \omega$, $\log ARPK$, $\log u$, and $\log \frac{m}{k}$, "Total" refers to the standard deviation of the residualized variable wrt. to Industry $\times$ Year FEs. "Within" refers to the standard deviation of the residualized variable wrt. to Industry $\times$ Year FEs and *Firm FE*. "Fraction Within" is Within/Total $\times$ 100. For rows with $\beta(Y, X)$, "Total" refers to the regression coefficient of Y on X with Industry $\times$ Year FEs. "Within" refers to the regression coefficient of X on Y with Industry $\times$ Year FEs and *Firm FE*.

Table 3: **The Effect of Measurement Error Adjusted Capacity Utilization on TFP**

Dependent variable	$\log(\omega)$			
	(1)	(2)	(3)	(4)
$\log \tilde{u}$	0.18***	0.91***	1.15***	0.98***
	(0.0092)	(0.054)	(0.063)	(0.048)
Fixed Effects				
Firm	✓	✓	✓	✓
Industry$\times$Year	✓	✓	✓	✓
IV	—	$\log(\frac{e}{k})$	Leave out focal firm	Both IVs
Observations	27,749	27,749	27,749	27,749

This table reports the effect of capacity utilization ($\log(u)$) on TFP ($\log(\omega)$) using equation 4. In column (1), we report the OLS without adjusting for measurement errors in capacity utilization. Column (2) adjusts for measurement error in capacity utilization using the firm's energy expenditures by capital ($\log(\frac{e}{k})$ as n IV. Column (3) adjusts for measurement error in capacity utilization using the average sales of other firms in the same industry – leaving the focal firm out – as an IV. Column (4) adjusts for measurement error using both IVs. Standard errors are clustered at the firm level. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 4: **The Effect of Capacity Utilization on Maintenance Rate**

Dependent variable	$\log(m/k)$			
	(1)	(2)	(3)	(4)
$\log \tilde{u}$	0.199***	1.813***	1.671***	1.730***
	(0.0118)	(0.421)	(0.386)	(0.422)
Fixed Effects				
Firm	✓	✓	✓	✓
Year	✓	✓	✓	✓
IV		Leave out focal firm	Leave out focal firm's district	Leave out focal firm's state
Observations	27,749	27,749	27,749	27,749

This table reports the effect of capacity utilization ($\log(u)$) on maintenance rates ($\log(m/k)$) using equation 5. Column (1) shows the OLS estimates. In column (2), we use the average sales of other firms in the same industry – leaving the focal firm out – as an IV. In columns (3) and (4), we use the average sales of other firms in the same industry, leaving out the focal firm's district and state, respectively. Standard errors are clustered at the firm level. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: **Correlation of Maintenance Rate and ARPK**

Dependent variable	Log(ARPK)			
	(1)	(2)	(3)	(4)
$\log(m/k)$	0.61***	0.58***	0.60***	0.60***
	(0.0089)	(0.0095)	(0.0080)	(0.0080)
Fixed Effects				
Industry $\times$ Year	—	✓	✓	✓
Firm	—	—	✓	✓
State $\times$ Year	—	—	—	✓
Observations	61,803	61,803	61,803	61,803

This table reports the correlation of maintenance rate ($\log(m/k)$) and ARPK ($\log ARPK$) using equation 6. Column (1) is the OLS without any FEs. Column (2) includes Industry $\times$ Year FEs. Column (3) adds Firm FEs. Column (4) adds State $\times$ Year FEs. Standard errors are clustered at the firm level. ***, **, * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 6: **Volatility and Autocorrelation of ARPK Adjusted for Capacity Utilization**

Adjustment by	—	$\log(\frac{m}{k})$	$\log(\tilde{u})$	$\log(\frac{e}{k})$
	(1)	(2)	(3)	(4)
σ	0.70	0.61	0.51	0.58
ρ	0.69	0.57	0.59	0.68

This table reports the standard deviation and autocorrelation of ARPK ($\log ARPK$) after adjusting for various utilization measures. Column (1) shows the statistics for ARPK without any adjustment. For columns (2), (3), and (4), we report standard deviation and autocorrelation for maintenance-rate adjusted ARPK ($\log ARPK - \log(m/k)$), utilization-adjusted ARPK ($\log ARPK - log\tilde{u}$), and energy expenditure adjusted ARPK ($\log ARPK - \log(e/k)$). We correct for measurement error in utilization using $\log(\frac{e}{k})$ as an IV. For both standard deviation and autocorrelation, we residualize each variable wrt. to Industry $\times$ Year and Firm FEs.

Table 7: **Calibration and Validation**

Panel A: Calibration Parameters

Parameter	Value	Description	Target/Source	Data	Model
		Panel A: Externally calibrated parameters			
r	4%	Interest rate	Standard	–	–
κ	1.45	Hours worked	Standard	1/3	1/3
ν	3/4	Returns to scale	Standard	–	–
α	1/3	Capital intensity	Standard	–	–
		Panel B: Internally calibrated parameters			
ρ	0.767	Autocorrelation of TFP	Autocorrelation of $\log \omega$	0.71	0.72
σ	0.165	Dispersion of TFP	Volatility of $\log \omega$	0.27	0.26
λ	0.087	Resale loss	$\mathbb{E}(i/k \mid i/k < 0)$	-0.11	-0.06
p^m	0.530	Price of maintenance	Median maintenance rate	3.5%	4.4%
ζ^m	0.218	Elas. of δ to $\tilde{m}$	Elas. of $\tilde{m}$ to $ARPK$	0.63	0.63
ζ^u	0.760	Elas. of δ to u	Avg. utilization rate	66.3%	68.9%
θ	0.327	Maintenance disruption cost	Elas. of $\tilde{m}$ to ω	0.77	0.81

Panel C: Untargeted Moments: Data and Model

Description	**Data**	**Model**
Elas. of i/k to $\tilde{m}$	0.22	0.27
Elas. of i/k to ω	0.24	0.22
Avg. maintenance share of value added	3.0%	7.4%
Avg. maintenance share of labor cost	15.8%	13.5%
Std. dev. of $\log \tilde{m}$	0.71	0.32
Median depreciation	9.4%	7%
Std. dev. of $\log ARPK$	0.695	0.496

Panel (A) shows the externally calibrated parameters. Panel (B) reports internally calibrated parameters and the statistics we use for calibration. Panel (C) shows untargeted moments in the data and our model.

Table 8: **Decomposition of the Sources of Dispersion in ARPK**

	Baseline	Fixed user cost	No covariance only
Agg TFP	—	-4.0%	-5.2%
$\sigma(arpk)$	0.496	0.285	0.320
$\sigma(arpk)$ rel. to data (%)	75.5%	43.4%	48.7%

This table reports a decomposition of the sources of aggregate TFP and dispersion of ARPK in the baseline model, using the decomposition presented in page 26.

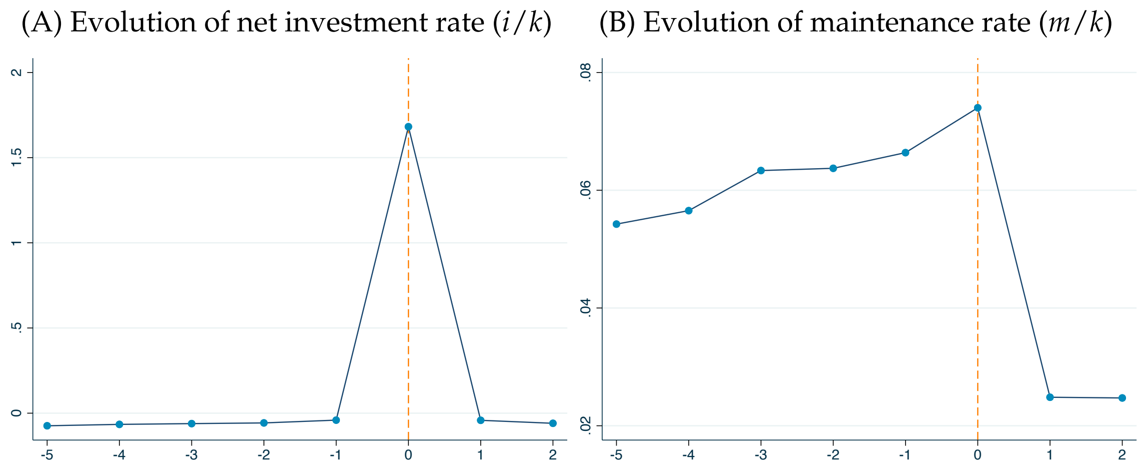
Table 9: θ **as a Source of Misallocation**

	Total	Avg. util	Cov only	Distribution only	$\lambda = 0$
	(1)	(2)	(3)	(4)	(5)
Agg TFP change	13.6%	11.8%	1.65%	0.11%	0.1%
Agg TFP (as fraction of total)	—	87.05%	12.16%	0.80%	—

Columns 1 to 5 report a decomposition of the sources of the gains in aggregate total factor productivity in the model without disruption cost relative to our baseline model, using the decomposition presented in page 30. Column 5 reports the TFP gain relative to our baseline model if we simply set $\lambda = 0$.

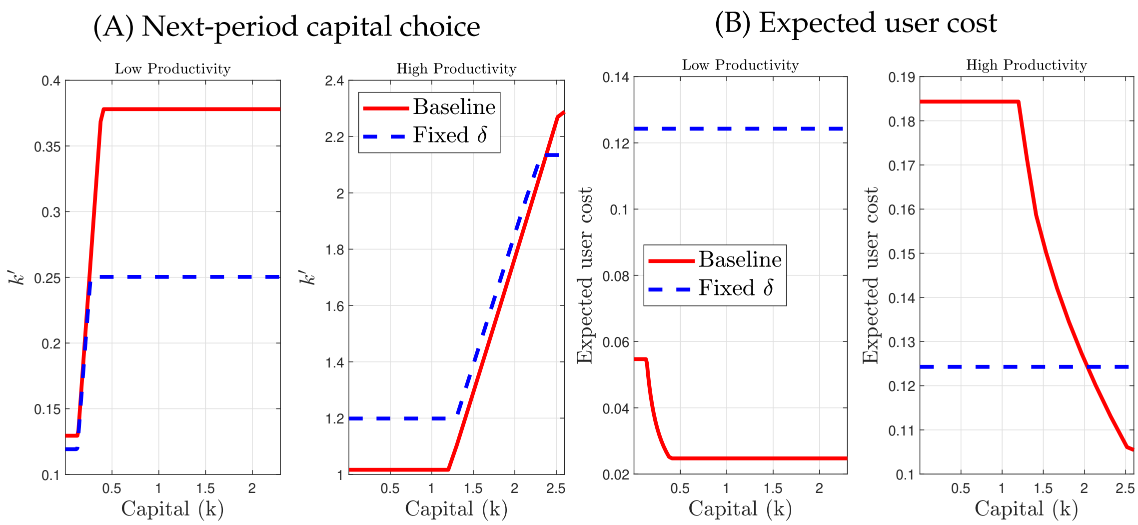
B Figures

Figure 1: **Evolution of Net Investment and Maintenance Rate Around the Year with the Highest Net Investment Rate: Data**



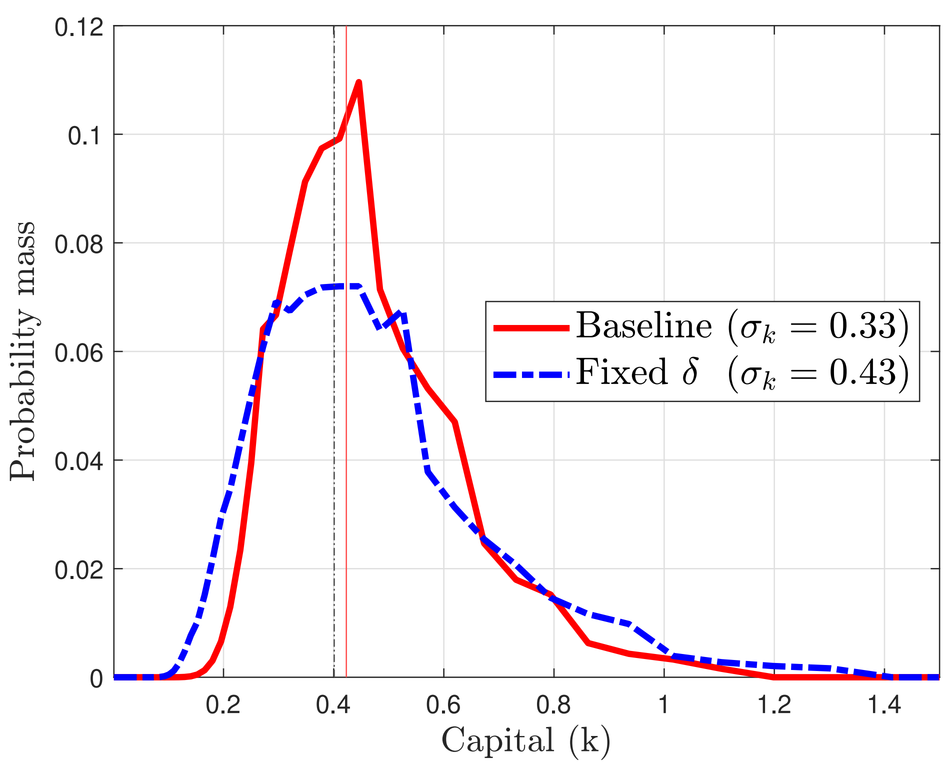
Panel (A) and (B) plot the median net investment (i/k) and maintenance rate (m/k) around the period with peak investment rate using the data. Event time = 0 is the period with the peak investment rate for a firm.

Figure 2: **Next-period capital choice and expected user costs**



Panel (A) plots the next-period capital choice for a low- (left sub-panel) and high- (right sub-panel) productivity firm, as a function of capital. Panel (B) plots the expected user cost for the same corresponding firm.

Figure 3: **Distribution of capital in baseline and fixed user cost counterfactual**



This figure plots the distribution of capital for our baseline model (red solid line) and the counterfactual with fixed user costs (blue dashed-dot line).

References

ASKER, J., A. COLLARD-WEXLER, AND J. D. LOECKER (2014): "Dynamic Inputs and Resource (Mis) Allocation," *Journal of Political Economy*, 122(5), 1013–1063.

BALEY, I., AND A. BLANCO (2022): "The Macroeconomics of Partial Irreversibility," Working Paper.

BASU, S., J. G. FERNALD, AND M. S. KIMBALL (2006): "Are Technology Improvements Contractionary?," *American Economic Review*, 96(5), 1418–1448.

BLOOM, N., B. EIFERT, A. MAHAJAN, D. MCKENZIE, AND J. ROBERTS (2012): " Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*, 128(1), 1–51.

BLOOM, N., M. FLOETOTTO, N. JAIMOVICH, I. SAPORTA-EKSTEN, AND S. J. TERRY (2018): "Really uncertain business cycles," *Econometrica*, 86(3), 1031–1065.

BURNSIDE, C., AND M. EICHENBAUM (1996): "Factor-Hoarding and the Propagation of Business-Cycle Shocks," *American Economic Review*, 86, 1154–1174.

CABALLERO, R. J. (1999): "Aggregate Investment," *Handbook of Macroeconomics*, 1.

CAUNEDO, J., AND E. KELLER (2021): "Capital obsolescence and agricultural productivity," *Quarterly Journal of Economics*, 136(1), 505–561.

CHODOROW-REICH, G., L. KARABARBOUNIS, AND R. KEKRE (2023): "The Macroeconomics of the Greek Depression," *American Economic Review*, 113(9), 2411–57.

CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113(1), 1–45.

COOPER, R., J. HALTIWANGER, AND L. POWER (1999): "Machine Replacement and the Business Cycle: Lumps and Bumps," *American Economic Review*, 89(4), 921–946.

COOPER, R. W., AND J. C. HALTIWANGER (2006): "On the Nature of Capital Adjustment Costs," *The Review of Economic Studies*, 73(3), 611–633.

DAVID, J. M., H. A. HOPENHAYN, AND V. VENKATESWARAN (2016): " Information, Misallocation, and Aggregate Productivity *," *The Quarterly Journal of Economics*, 131(2), 943–1005.

DAVID, J. M., L. SCHMID, AND D. ZEKE (2022): "Risk-adjusted capital allocation and misallocation," *Journal of Financial Economics*, 145(3), 684–705.

DAVID, J. M., AND V. VENKATESWARAN (2019): "The sources of capital misallocation," *American Economic Review*, 109, 2531–2567.

DOMS, M., AND T. DUNNE (1998): "Capital adjustment patterns in manufacturing plants," *Review of Economic Dynamics*, 1(2), 409–429.

EDMOND, C., V. MIDRIGAN, AND D. Y. XU (2023): "How Costly Are Markups?," *Journal of Political Economy*, 131(7), 1619–1675.

FERNALD, J. (2014): "A quarterly, utilization-adjusted series on total factor productivity," Working Paper.

GAVAZZA, A., AND A. LANTERI (2021): "Credit shocks and equilibrium dynamics in consumer durable goods markets," *The Review of Economic Studies*, 88(6), 2935–2969.

GREENWOOD, J., Z. HERCOWITZ, AND G. W. HUFFMAN (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *The American Economic Review*, 78, 402–417.

HAUSMAN, J., G. LEONARD, AND J. D. ZONA (1994): "Competitive analysis with differenciated products," *Annales d'Economie et de Statistique*, pp. 159–180.

HAUSMAN, J. A. (1996): "Valuation of New Goods under Perfect and Imperfect Competition," in *The Economics of New Goods*, ed. by T. F. Bresnahan, and R. J. Gordon, pp. 209–247. National Bureau of Economic Research, Chicago.

HOPENHAYN, H. A. (1992): "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, 60(5), 1127–1150.

HSIEH, C.-T., AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 127(4), 1403–1448.

IMBENS, G., AND D. RUBIN (2015): *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press, Cambridge.

KABIR, P., AND S. M. MANSOURI (2021): "Capital quality, productivity, and financial constraints: Evidence from India," *Productivity, and Financial Constraints: Evidence from India (November 5, 2021)*.

KEHRIG, M., AND N. VINCENT (2023): "Good Dispersion, Bad Dispersion," Working Paper.

KERMANI, A., AND Y. MA (2022): "Asset Specificity of Non-Financial Firms," *The Quarterly Journal of Economics*, 138(1), 205–264.

LANTERI, A. (2018): "The market for used capital: Endogenous irreversibility and reallocation over the business cycle," *American Economic Review*, 108(9), 2383–2419.

LANTERI, A., P. MEDINA, AND E. TAN (2023): "Capital-reallocation frictions and trade shocks," *American Economic Journal: Macroeconomics*, 15(2), 190–228.

MARTÍNEZ, L. R. (2022): "How Much Should We Trust the Dictator's GDP Growth Estimates?," *Journal of Political Economy*, 130(10), 2731–2769.

MCGRATTAN, E., AND J. SCHMITZ (1999): "Maintenance and repair: Too big to ignore," *Federal Reserve Bank of Minneapolis Quarterly Review*, 23(4), 2–13.

MIDRIGAN, V., AND D. Y. XU (2014): "Finance and misallocation: Evidence from plant-level data," *American Economic Review*, 104(2), 422–458.

RAMEY, V. A., AND M. D. SHAPIRO (2001): "Displaced Capital : A Study of Aerospace Plant Closings," *Journal of Political Economy*, 109(5), 958–992.

RESTUCCIA, D., AND R. ROGERSON (2008): "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 11(4), 707–720.

TAN, E. (2022): "Entrepreneurial investment dynamics and the wealth distribution," Working Paper.

TAN, E., AND T. H. ZEIDA (2023): "Consumer demand and credit supply as barriers to growth for Black-owned startups," *Journal of Monetary Economics*.

TAUCHEN, G. (1986): "Finite state markov-chain approximations to univariate and vector autoregressions," *Economics Letters*, 20(2), 177–181.

Appendix

This appendix contains supplementary material, tables, and figures.

A Additional Data Details

A.i Utilization sub-sample

We present summary statistics for the sub-sample of firm-years with non-missing utilization. Only about 45% of the sample reports utilization.

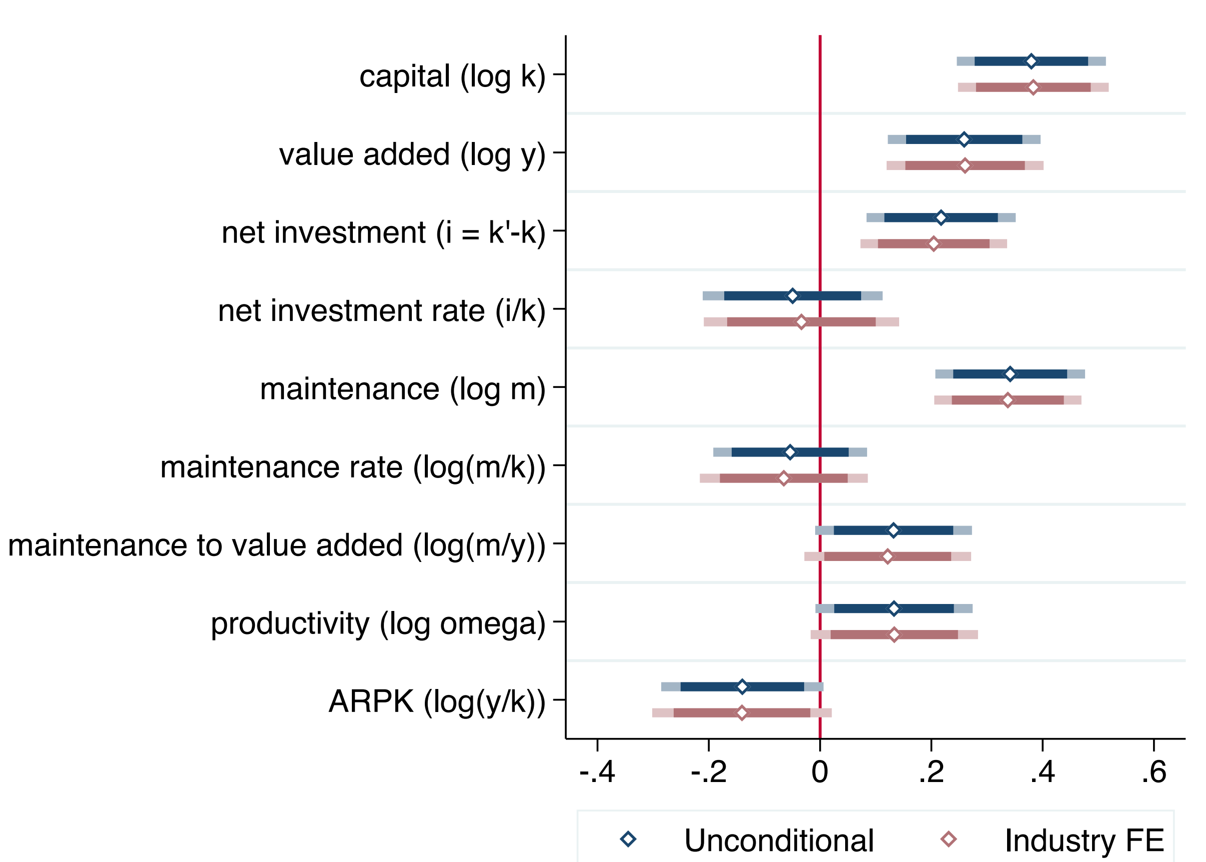
Table OA.1: **Summary statistics capacity utilization sub-sample**

	Mean	Std. Dev.	p25	Median	p75
capacity utilization $(\tilde{u})$	0.663	0.695	0.325	0.548	0.801
capital (k)	46.944	149.159	2.184	6.291	23.545
value added (y)	70.283	209.503	3.747	11.736	41.590
net investment $(i = k' - k)$	1.921	16.755	-0.712	-0.105	0.719
net investment rate $(\frac{i}{k})$	0.161	0.709	-0.108	-0.045	0.114
log measured productivity $(\log \omega)$	1.039	0.810	0.563	1.072	1.578
ARPK $(\frac{y}{k})$	3.139	4.439	0.938	1.768	3.506
maintenance (m)	1.500	4.615	0.063	0.211	0.777
maintenance rate $(\frac{m}{k})$	0.061	0.080	0.015	0.034	0.072
maintenance to value added $(\frac{m}{y})$	0.028	0.030	0.010	0.020	0.035
Number of firm-year observations	27,749				

This table presents the summary statistics for the sub-sample of firms with non-missing utilization data. Product-level capacity utilization is defined as the quantity of products produced divided by product-level installed capacity. Capacity utilization is the average of product-level capacity utilization across different products within a firm. Capital (k) is measured as the value of PPE. Value added (y) is the difference between sales and cost of material input. Investment (i) is net investment and is measured as the differences in PPE. Maintenance (m) is the repair and maintenance item from the income statement. $\log \omega$ is measured logarithm of TFP and is defined as a Solow residual $\log \omega = \frac{1}{\Theta_z}(\log py - \Theta_k \log k)$, and $\log ARPK = \log y - \log k$. Where relevant, the values are reported in Millions of US dollars in 2019.

The second and third rows report statistics for the capital stock and value-added. These firms are slightly larger than the full sample. The median capital stock and value-added are US\$6.29 million (relative to US\$5.7 million for the full sample) and US\$11.73 million (relative to US\$10.7 million for the full sample), respectively. Despite the differences in size, the ratio of normalized variables is much closer. For instance, the median maintenance rate is 3.5% in the full sample and 3.4% in the sub-sample. Figure OA.1 plots the covariate balance between the full sample and the sub-sample with utilization. The figures show that while firms in the utilization sub-sample are slightly larger, the normalized variables look similar in both sub-samples. To maximize the sample size, we use the full sample, as discussed in the main text, unless the specific statistic (e.g., a regression specification) requires observing utilization directly.

Figure OA.1: **Covariate Balance**



This figure shows coefficient estimates and 90% (lighter bars) and 95% (darker bars) error bands of the difference between firms with and without capacity utilization data for different variables. All variables are normalized to have a mean of zero and a standard deviation of one. "Unconditional" refers to the sample where we compare the samples without conditioning on any fixed effects. "Industry FE" refers to the sample where we compare the samples with conditioning on industry fixed effects.

A.ii Hausman instrument

In this subsection, we provide more details on constructing the Hausman instruments. For brevity, we only explain the details for "leave out focal firm's state". The rest follow the same logic. Let $Sales_{i,j,t}$ denote the sales of the focal firm i in year t operating in industry j. First, we regress $Sales_{i,j,t}$ on firm and year FEs and take the residuals to construct the IV:

$$IV_{i,j,t} = \sum_{i \notin S} Sales_{i,j,t} \tag{OA.1}$$

where S denotes the set of firms operating in the same state as the focal firm. The underlying assumption behind the instrument is that demand across firms operating in the same narrowly defined industry and year is correlated.

A.iii FRED Capacity Utilization Data

We retrieve capacity utilization data from FRED.[21] Capacity utilization is computed as an output index divided by a capacity index. The capacity index "captures the concept of sustainable maximum output - the greatest level of output a plant can maintain within the framework of a realistic work schedule, after factoring in normal downtime and assuming sufficient availability of inputs to operate the capital in place." Table OA.2 below presents summary statistics of FRED's monthly capacity utilization index for 1989-2019 (the same period as our main sample).

Table OA.2: **Summary Statistics of US Capacity Utilization**

	Mean	Std. Dev.	p25	Median	p75
Total Capacity Utilization (TCU)	79.10	3.67	76.63	79.73	81.99

This table presents the summary statistics for FRED's monthly capacity utilization index for 1989-2019 (the same period as our main sample).

B Additional Model Details

In this section, we derive in greater detail key model equations referenced to in the main text.

[21] The data was accessed on Dec 18 2023.

B.i Labor demand and profits

Static profit maximization of labor gives us the first order condition,

$$(1-\alpha)\nu l^{(1-\alpha)\nu-1}C^{1-\nu}\left(z(uk)^{\alpha}\right)^{\nu} = w \tag{OA.2}$$

Rearranging terms, we obtain for labor demand,

$$l = \left(\frac{(1-\alpha)\nu}{w}\right)^{\frac{1}{1-(1-\alpha)\nu}} C^{\frac{1-\nu}{1-(1-\alpha)\nu}} \left(z(uk)^{\alpha}\right)^{\frac{\nu}{1-(1-\alpha)\nu}} \tag{OA.3}$$

Substituting labor demand into the profit function gives us the formulation presented in the main text.

B.ii Derivation of results in Section A.2

As in the main text, for the purposes of exposition, we assume that $\theta = 0$. Then, for a firm that is investing, the first order condition for maintenance is given by

$$p^m = -\frac{\partial \delta}{\partial \tilde{m}}, \tag{OA.4}$$

and the first order condition for investment is,

$$1 = \frac{1}{1+r}\frac{\partial \mathbb{E}[V(z',k')]}{\partial k'}. \tag{OA.5}$$

Combining both equations give us

$$p^m = -\frac{1}{1+r}\frac{\partial \mathbb{E}[V(z',k')]}{\partial k'}\frac{\partial \delta}{\partial \tilde{m}} \tag{OA.6}$$

For a firm that is disinvesting, the first order condition for maintenance is given by

$$\frac{p^m}{1-\lambda} = -\frac{\partial \delta}{\partial \tilde{m}}, \tag{OA.7}$$

and the first order condition for investment is,

$$1-\lambda = \frac{1}{1+r}\frac{\partial \mathbb{E}[V(z',k')]}{\partial k'}. \tag{OA.8}$$

Combining both equations give us

$$p^m = -\frac{1}{1+r}\frac{\partial \mathbb{E}[V(z',k')]}{\partial k'}\frac{\partial \delta}{\partial \tilde{m}} \tag{OA.9}$$

Finally, for a firm that is in the inaction region (i.e., $k' = (1-\delta)k$), the first order condition for maintenance is

$$p^m = -\frac{1}{1+r}\frac{\partial \mathbb{E}[V(z',k')]}{\partial k'}\frac{\partial \delta}{\partial \tilde{m}}. \tag{OA.10}$$

Therefore, across the state space, (z, k), the first-order condition for maintenance can always be summarized by

$$p^m = -q^{marg}\frac{\partial \delta}{\partial \tilde{m}}, \tag{OA.11}$$

which is equation 20 in the main text. Furthermore, it is clear that $q^{marg} = 1$ for a firm that is indifferent between investing and inaction, and $q^{marg} = 1-\lambda$ if it is indifferent between disinvesting and inaction; and that from the concavity of the value function, q^{marg} is strictly decreasing in capital for firms that are in the inaction region. Therefore, we see that $q^{marg} \in [1-\lambda, 1]$, as discussed in the main text.

C A Model of Capital Vintages and Endogenous Maintenance

We consider a framework similar to Cooper, Haltiwanger, and Power (1999) in the context of our model, while dispensing with the assumption of investment irreversibility. Specifically, Cooper, Haltiwanger, and Power (1999) propose a model where lumpy investment is associated with vintage upgrading; importantly, because they consider a model where a fixed cost is paid whenever new investment is purchased, all positive investment in their framework is associated with vintage upgrading. Our goal is to show that, in the context of our model with endogenous maintenance, such a model predicts a joint dynamics of maintenance and investment around investment spikes that is inconsistent with the data.

To summarize, one can write the firm's dynamic programming problem as

$$V(z,k) = \max_{h,k'} \pi + (1-\delta(\tilde{m}))k - p^m m - hfk - k' + \frac{1}{1+r}\mathbb{E}[V(z',k')|z] \tag{OA.12}$$

where $h = 1$ if the firm chooses $k' > (1-\delta)k$, and $h = 0$ otherwise. Following the assumption of Cooper, Haltiwanger, and Power (1999), we assume that all vintages of capital are identical in productivity. Otherwise, following the assumptions of our model, we assume that the depreciation rate is endogenous to maintenance.

Notice that the first order condition for maintenance rate, regardless of investment decision, can also be expressed in terms of OA.11. In turn, conditional on $k' \neq (1-\delta)k$, we see that $q^{marg} = 1$; while for $k' = (1-\delta)k$, $q^{marg} > 1$ and satisfies $\frac{\partial q^{marg}}{\partial k} < 0$.[22] This implies that for a firm in periods of inaction, we anticipate seeing a rise in maintenance rates, similar to our model with investment irreversibility; but maintenance rates *fall* conditional on an investment spike, which is counterfactual to our empirical finding.

A similar counterfactual result can be obtain if we modify the vintage model to incorporate scrappage costs (e.g., Gavazza and Lanteri (2021)). To summarize, one can write the firm's dynamic programming problem as

$$V(z,k) = \max_{h,k'} \pi + (1-hf)(1-\delta(\tilde{m}))k - p^m m - k' + \frac{1}{1+r}\mathbb{E}[V(z',k')|z] \qquad \text{(OA.13)}$$

where the only modification is that the fixed cost is paid in terms of the depreciated capital stock. In other words, capital exhibits irreversibility (like in our model), but irreversibility shows up because firms need to scrap their entire capital stock before buying a new vintage.

In this model, q^{marg} has the same characteristics as the previous model, but the firm's first order condition for maintenance depends on investment. To be precise, for a firm in inaction or disinvestment, maintenance is given by equation OA.11, but for a firm that is upgrading, maintenance solves

$$p^m = (1-f)\frac{\partial \delta}{\partial \tilde{m}}. \qquad \text{(OA.14)}$$

Because $f > 0$, this implies that maintenance rates fall even lower than the previous model, conditional on an investment spike. In other words, like the earlier model, for a firm in periods of inaction, we anticipate seeing a rise in maintenance rates, but maintenance rates *fall* conditional on an investment spike, which is counterfactual to our empirical finding.

[22]This is a standard result; see, for example, the handbook chapter in Caballero (1999) for a derivation.

D A Model of Utilization and Flexible Capital

Here, we sketch out a model of endogenous utilization with a fully flexible capital stock. The goal is to show that utilization plays no role in amplifying aggregate productivity through the "covariance" term in the main text.

Consider a firm that faces the following static profit-maximization problem,

$$\max_{u,k} z\,(uk)^{\alpha} - (r + \delta(u))\,k, \tag{OA.15}$$

where the physical capital user cost $\delta(u)$ is increasing in u.

The solution to the model is characterized by the pair of first-order conditions:

$$\alpha z u\,(uk)^{\alpha-1} = r + \delta(u), \tag{OA.16}$$

which is the first-order condition for capital, and

$$\alpha z\,(uk)^{\alpha-1} = \frac{\partial \delta(u)}{\partial u}, \tag{OA.17}$$

which is the first-order condition for utilization.

Dividing the first and second equations, we obtain,

$$u = (r + \delta(u))\,/\frac{\partial \delta(u)}{\partial u}. \tag{OA.18}$$

which tells us that the solution to u is independent of capital or productivity. In turn, from equation OA.16, we see that capital itself is varying only due to productivity. Consequently, the covariance term would be zero if capital was fully flexible.

E Additional Quantitative Model Results and Derivations

E.i Evolution of Net Investment and Maintenance Rate Around the Period with the Highest Net Investment Rate: Calibrated Model

Figure OA.2 plots the policy function for maintenance (solid blue line; left axis) and net investment (dashed orange line; right axis) for one value of productivity, as a function of capital. The dashed vertical threshold separate the state space in regions of investment, inaction, and disinvestment.

Figure OA.2: **Policy function for maintenance**

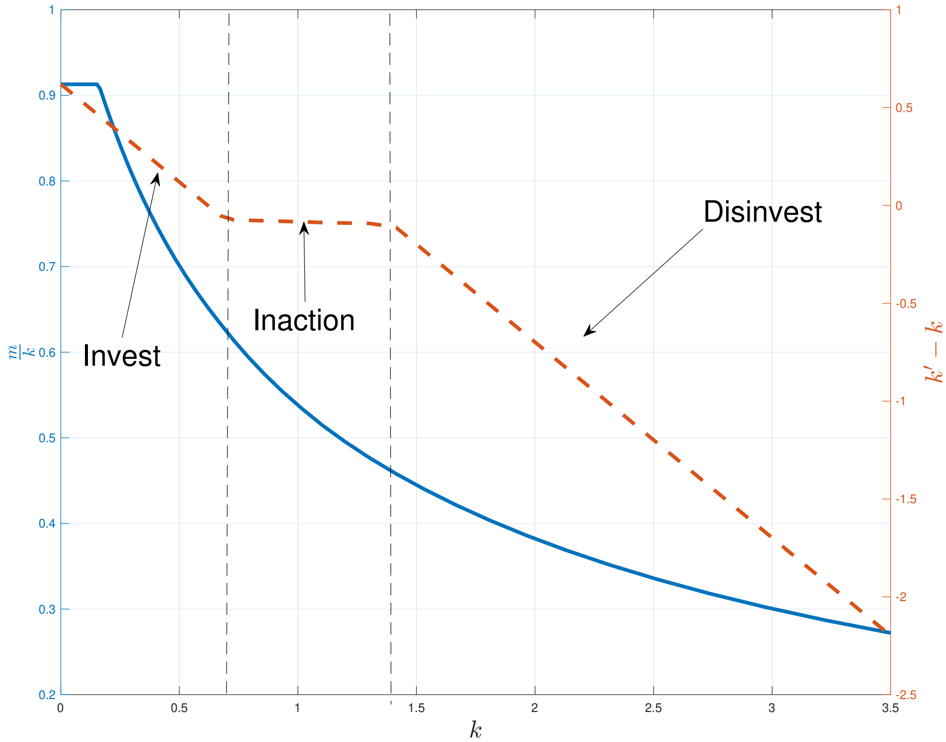
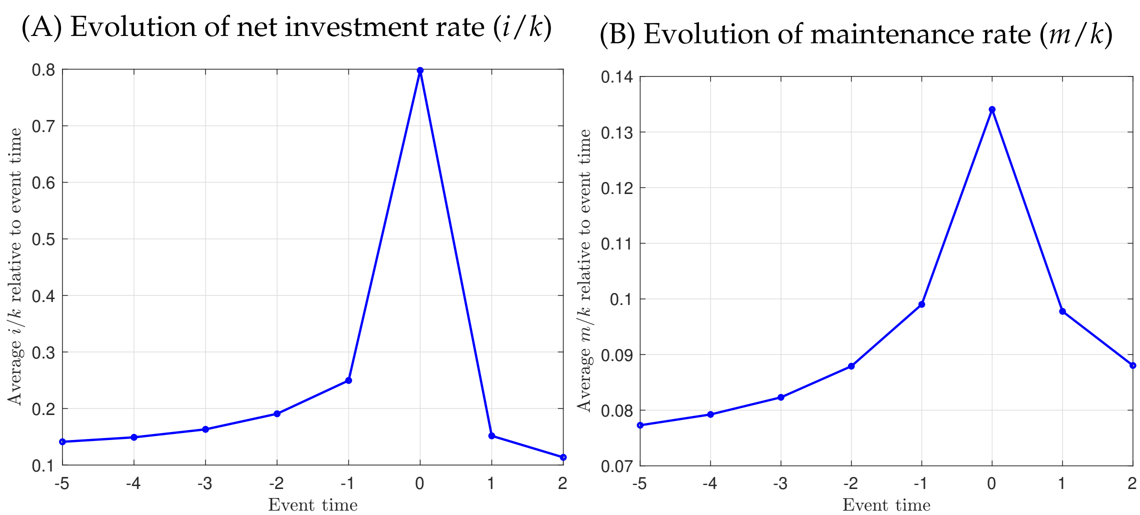


Figure OA.3 plots the model-implied dynamics of maintenance and investment around the period of investment spikes.

Figure OA.3: **Evolution of Net Investment and Maintenance Rate Around the Period with the Highest Net Investment Rate: Calibrated Model**



Panel (A) and (B) plot the median net investment (i/k) and maintenance rate (m/k) around the period with peak investment rate using the calibrated model. Event time = 0 is the period with the peak investment rate for a firm.

E.ii Statistics of Adjusted ARPK

We report here the dispersion and autocorrelation of ARPK adjusted for maintenance and utilization in the data and our model. In the context of utilization adjustment, we consider two variations. First, we adjust ARPK using the data concept of utilization, namely output capacity utilization. As previously discussed, this amounts to u^{α} in our model. Second, we adjust ARPK using actual capital utilization rates u. Note that from the perspective of the model, u is the correct measure of adjustment, whereas adjusting by u^{α} leads to a smaller attenuation in the dispersion of ARPK (specifically, by $1-\alpha$), and should have no impact on the autocorrelation of ARPK.

Table OA.3: Statistics of Adjusted ARPK

		ARPK	u-adj ARPK (data)	m-adj ARPK	u-adj ARPK (true)
Std. Dev.	data	0.695	0.51	0.61	—
	data (rel. to raw)	—	0.734	0.878	—
	model	0.496	0.379	0.187	0.144
	model (rel. to raw)	—	0.763	0.377	0.290
Autocorr	data	0.69	0.59	0.57	—
	data (rel. to raw)	—	0.855	0.826	—
	model	0.614	0.613	0.556	0.601
	model (rel. to raw)	—	0.997	0.904	0.978

Persistence and dispersion of ARPK after adjustment for utilization and maintenance. Energy adjustment is not available for the model.

E.iii Details of the TFP Decomposition

E.iii.1 Formula for TFP Decomposition

Aggregate total factor productivity, as by convention, is defined as

$$\log TFP \equiv \log \frac{\mathbb{E}[y^{\nu}]^{\frac{1}{\nu}}}{\mathbb{E}[k]^{\alpha}\mathbb{E}[l]^{1-\alpha}}, \tag{OA.19}$$

where the expectation operator $\mathbb{E}[\cdot]$ refers to the intergral over the distribution of firms over the state space z and k.

By substituting in a firm's labor decision (from equation OA.3), we can rewrite the definition of TFP within our model as

$$\log TFP = \frac{1-(1-\alpha)\,\nu}{\nu} \log \mathbb{E}[z^{\Theta_z} u^{\Theta_k} k^{\Theta_k}] - \alpha \log \mathbb{E}[k]. \tag{OA.20}$$

We now use this formula to further motivate our decomposition exercise. Using this formula, TFP differences between any two comparison models can be written as

$$\begin{aligned}\Delta \log TFP &= \log TFP_b - \log TFP_c \\ &= \frac{1-(1-\alpha)\,\nu}{\nu}\left(\log \mathbb{E}_b[z^{\Theta_z}u^{\Theta_k}k^{\Theta_k}] - \log \mathbb{E}_c[z^{\Theta_z}u^{\Theta_k}(xk)^{\Theta_k}]\right) \\ &\quad ... - \alpha\left(\log \mathbb{E}_b[k] - \log \mathbb{E}_c[xk]\right). \end{aligned} \tag{OA.21}$$

where we use the subscripts b and c to denote "baseline" and "counterfactual", respectively. The subscripts are also appended to the expectations operator to denote the idea that the joint distribution of productivity, capital, and utilization are different.

The term x in the formula is a proportionality constant. Specifically, x satisfies,

$$\mathbb{E}_c[xk] = K_c, \tag{OA.22}$$

where K_c is the aggregate capital stock in the counterfactual; and also satisfies,

$$\mathbb{E}_b[k] = \mathbb{E}_c[k]. \tag{OA.23}$$

In other words, x is a constant of proportionality that we use to re-express the counterfactual distribution of capital, such that absent the normalization by x, the aggregate capital stock is constant across both economies.

Trivially then, we see that

$$\log \mathbb{E}_b[k] - \log \mathbb{E}_c[xk] = -\log x. \tag{OA.24}$$

Similarly, we obtain,

$$\begin{aligned}\log \mathbb{E}_b[z^{\Theta_z}u^{\Theta_k}k^{\Theta_k}] - \log \mathbb{E}_c[z^{\Theta_z}u^{\Theta_k}(xk)^{\Theta_k}] &= \log \mathbb{E}_b[z^{\Theta_z}u^{\Theta_k}k^{\Theta_k}] - \log \mathbb{E}_c[z^{\Theta_z}u^{\Theta_k}k^{\Theta_k}] \\ &\quad ... - \Theta_k \log x \end{aligned} \tag{OA.25}$$

Since $\Theta_k = \frac{\alpha\nu}{1-(1-\alpha)\nu}$, equation OA.21 trivially reduces to

$$\begin{aligned}\Delta \log TFP &= \log TFP_b - \log TFP_c \\ &= \frac{1-(1-\alpha)\,\nu}{\nu}\left(\log \mathbb{E}_b[z^{\Theta_z}u^{\Theta_k}k^{\Theta_k}] - \log \mathbb{E}_c[z^{\Theta_z}u^{\Theta_k}k^{\Theta_k}]\right), \end{aligned} \tag{OA.26}$$

which is to say, differences in TFP across two comparison economies are not affected by the differences in scales of the economy itself (the x term subtracts out).

Finally, we can re-expressed equation OA.26 in terms of averages and covariances,

$$\Delta \log TFP = \frac{1-(1-\alpha)\nu}{\nu}\left\{\log\left(\mathbb{E}_b[u^{\Theta_k}]\mathbb{E}_b[z^{\Theta_z}k^{\Theta_k}] + cov_b\left(u^{\Theta_k}, z^{\Theta_z}k^{\Theta_k}\right)\right) ..\right.$$
$$\left... - \log\left(\mathbb{E}_c[u^{\Theta_k}]\mathbb{E}_c[z^{\Theta_z}k^{\Theta_k}] + cov_c\left(u^{\Theta_k}, z^{\Theta_z}k^{\Theta_k}\right)\right)\right\}, \tag{OA.27}$$

which is the expression in the main text. Therefore, this decomposition shows that the differences in aggregate TFP between two economies in our model is driven by differences in (1) aggregate utilization, (2) the joint distribution of productivity and capital, and (3) the covariance between utilization and physical output if utilization was uniformly 1.

E.iii.2 Details of the Shapely Decomposition

We are interested in decomposing how much of the change in aggregate TFP comes from each of the three channels discussed above. However, it is clear from equation OA.27 that the effects of each channel enter non-linearly into the computation of $\Delta \log TFP$. Therefore, we utilize the logic of Shapely decomposition to decompose these channels.

We now briefly summarize the logic of the Shapely decomposition, using as an example our computation for the partial contribution of the utilization channel. For notation, denote $U_i \equiv \mathbb{E}_i[u^{\Theta_k}]$, $COV_i \equiv cov_i\left(u^{\Theta_k}, z^{\Theta_z}k^{\Theta_k}\right)$, and $Y_i \equiv \mathbb{E}_b[z^{\Theta_z}k^{\Theta_k}]$. Like before, denote $i \in \{b, c\}$ the baseline and counterfactual models respectively. Finally, for notation, for a given statistic $X_{k,s}$, let k and s denote the k sub-model of permutation s.

To compute the total partial effect of utilization, we first compute the individual partial effects,

- $\Delta \log TFP_{0,1} = \log(Y_0 U_1 + COV_0) - \log(Y_0 U_0 + COV_0)$
- $\Delta \log TFP_{1,1} = \log(Y_1 U_1 + COV_0) - \log(Y_1 U_0 + COV_0)$
- $\Delta \log TFP_{1,1} = \log(Y_0 U_1 + COV_1) - \log(Y_0 U_0 + COV_1)$
- $\Delta \log TFP_{2,1} = \log(Y_1 U_1 + COV_1) - \log(Y_1 U_0 + COV_1)$

And we obtain the average partial contribution as

$$\frac{1}{3}\Delta \log TFP_{0,1} + \frac{1}{6}\Delta \log TFP_{1,1} + \frac{1}{6}\Delta \log TFP_{1,2} + \frac{1}{6}\Delta \log TFP_{2,1} \tag{OA.28}$$

In other words, in computing the partial effect of U, we consider all possible permutations where we could "shut down" the utilization channel, and then averaging out the effects. A similar logic follows for decomposing the partial effects of the other channels. Note further that, unlike total TFP change, which as we showed is scale invariant, the Shapely decomposition procedure in our case, if carried out naively, is not scale-invariant. Therefore, to maintain a decomposition procedure that is also scale-invariant, we normalized the Y_1 and COV_1 terms by x, the constant of proportionality, when conducting the decomposition.