Capital Budgeting, Uncertainty, and Misallocation*

Ben Charoenwong†

Yosuke Kimura

Alan Kwan

Eugene Tan

April 8, 2022

Abstract

We develop an investment model with capital budgeting, where firms endogenously learn about firm fundamentals and make partially flexible investment plans. In the model, high-productivity firms have an incentive to acquire better information, giving rise to a novel channel of improved information allocation that reduces capital misallocation. However, net gains hinge on the cost of learning. We quantify the importance of this channel by calibrating our model to data on firms' expectations and planned investment. Misallocation is substantially mitigated when only learning incentives are accounted for, but these net gains are reduced because learning costs empirically increase in productivity.

*The authors thank Jose Maria Barrero, Laurent Cavenaile, Murat Celik, Guojun Chen, Russell Cooper, Xiang Fang, Zu Yao Hong, Alexander Kopytov, Francois Gourio, Cosmin Ilut, Yueran Ma, Stephen Terry, Xu Tian, Toni Whited, Jincheng Tong, Choongryul Yang, and discussants and participants at the 11th Rimini Centre for Economic Analysis Macro-Finance conference, the 2021 Asian Econometric Society annual meetings, the 2022 ASSA, the Asian Bureau of Finance and Economic Research (ABFER) 2021 conference, Asia Pacific Corporate Finance Online Workshop, Barcelona GSE Summer Forum, CICM 2021, NASMES 2021, ITAM Finance Conference, VMACS Junior conference, and seminar participants at Brock University, the University of Toronto and the National University of Singapore. The content in this article has been reviewed by the Ministry of Finance of Japan to ensure the anonymity of survey respondents. Any views or opinions expressed herein are not representative of those of the Policy Research Institute, Ministry of Finance, or the Cabinet Office of Japan. Ben Charoenwong acknowledges financial support from the NUS Start-Up Grant No. R-315-000-119-133 and the Singapore Ministry of Education AcRF Tier 1 Research Grant No. R-315-000-122-115. This paper was previously circulated as "Investment Plans, Uncertainty, and Misallocation." Any remaining errors are our own.

†Ben Charoenwong, Email: ben.charoenwong@nus.edu.sg, National University of Singapore Business School; Yosuke Kimura, Email: kimura.y.bq@m.titech.ac.jp, Tokyo Institute of Technology; Alan Kwan, Email: apkwan@hku.hk, The University of Hong Kong; Eugene Tan, Email: Eugene.Tan@rotman.utoronto.ca, Rotman School of Management, University of Toronto.

JEL Classification: E22, G31, O16

Keywords: Corporate Investment, Uncertainty, Expectations Formation, Misallocation, Enterprise Resource Planning

1 Introduction

Capital budgets and investment plans typically involve pre-commitments of significant financial, physical, organizational, or human capital, as well as costly external financing (Harris et al. 1982; Harris and Raviv 1996, 1998; Malenko 2019; Graham 2022). These pre-commitments in turn render deviations from capital budgets costly or even infeasible (Plambeck and Taylor 2007; Dyreng et al. 2021). Consequently, ex ante acquisition of information regarding the prospects of a project is valuable since it improves the capital budgeting process and minimizes the risk of costly adjustment ex post. However, making better forecasts is costly since it requires manager attention as well as internal and external information collection. Thus, managers face a trade-off between the costs of planning and the costs of adjusting these plans ex post. In this paper, we develop a theory in which such trade-offs in capital budget formation is a central feature, examine the interplay of these two forces, and study its theoretical and quantitative implications for the macroeconomy.

We develop a theoretical framework to study how managers endogenously choose to acquire information when forming capital budgets, and we use it, along with unique microdata on firm expectations and investment plans, to quantitatively evaluate how these decisions impact firm value and aggregate capital misallocation. The key contribution of our theoretical approach is that it allows us to parsimoniously decompose the cross-sectional distribution of firm information acquisition behavior into contributions from private benefits and cost; in turn, it allows us to quantify the contributions these individual channels have on capital misallocation across firms.[1] Our results emphasize the complementarity in average firm productivity and information acquistion for firm production, whereas our quantitative results unveil that the marginal cost of information acquisition is rising in productivity, leading most firms to not fully internalize these complementarities. Consequently, we find substantial

[1]Our approach towards quantifying the unobserved private value of information for capital budgeting mirrors recent advances in the macroeconomics and finance literature in using implicit values to measure private or unobserved values of intangible assets; see, for instance, the approach taken by Bhandari and McGrattan (2020) to estimate the sweat equity of private firms, or Farboodi et al. (2021) to estimate the private value of financial information.

potential gains to be made from information reallocation across firms.

To formulate our framework, we develop a general-equilibrium investment model with heterogeneous firms a la Hopenhayn (1992) and embed endogenous capital budgeting. We emphasize three key dimensions of capital budgeting. First, firm managers formulate an investment plan *prior* to observing their productivity for the current fiscal year, then consider final adjustments to investment decisions *after* observing their productivity. The information set the firm has for forecasting current productivity includes their previous period productivity, as well as new information purchased during the initial periods of the current fiscal year. For example, firms can purchase enterprise resource management software (such as SAP, which introduces an internal accounting system to track the performance of different business and products within a firm), or engage with consultants for market research. Second, it is costly to deviate from investment plans ex post (Harris and Raviv 1996). Consequently, managers do not fully react to new information in real time, leading to losses in firm value resulting from over- or under- investment. Taken together, these two dimensions of capital budgeting and investment in practice imply that firms have an incentive to acquire better information to formulate more precise forecasts of both revenue and the marginal product of investment, even though information acquisition is costly. Finally, we allow for heterogeneous comparative advantage with regards to information acquisition, capturing the idea that different types of firms might have access to different technologies for information acquisition.

A key theoretical result of the model is that the marginal benefit of an additional "unit" of information for a firm always increases in its previous period productivity, which reflects the intuition that a similarly-sized forecasting error would be costlier for a higher productivity firm. Our result implies that information and productivity are complements in production, at both the firm and aggregate levels. This finding is akin to the predictions of a standard model of purely physical factor inputs, leading us to interpret information as an implicit factor of production (i.e., an intangible asset). Importantly, we derive our result using standard assumptions on the firm's revenue-generating function. We believe that this implication of the standard investment model for information acquisition is new to the literature, and consequently potentially important for future quantitative work in pricing the value of information, or future studies which wish to further examine capital budgeting in this context.

Our result implies that heterogenous incentives for information acquisition arises endogenously in our model, giving rise to a distribution of information across firms. Moreover, due to the complementarity between productivity and information, firm values and aggregate productivity are higher in our economy compared to a standard model with a degenerate distribution of information (but otherwise identical economy). This effect implies that capital budgeting creates an "information allocation effect" whereby firms' endogenous choice of information leads to the allocation of scarce resources (information) to firms that benefit from it the most. That said, the extent to which this effect is realized depends on the relative comparative advantage of different firms in acquiring information. Quantifying these gains therefore requires access to a dataset that, at a minimum, allows us to observe firm investment plan formation and sales forecasts. While this combination of data is rare in empirical settings, it would allow us to measure how flexible investment is when firms make forecasting errors, and how forecasting errors relate to productivity.

To that end, we bring in a large, nationally representative firm-level panel dataset from the Japan Ministry of Finance (MoF). The data contain firms' internal forecasts of sales and profits, and investment plans, as well as their realized values and other standard balance sheet and income statement items. This provides us key data to calibrate our model, which we use to draw indirect inference of the aforementioned information allocation effect. In addition, we present empirical predictions that guide our identification strategy, and also help provide reduced-form evidence of the model mechanism.

Our calibrated model imply that high-productivity firms have a strong incentive to acquire better information. When we assume that there isn't comparative advantage in information acquisition costs, our model predicts that a 1% increase in initial firm-level total factor productivity (TFP) leads to an average of 0.065% decrease in the dispersion of forecast errors of current fiscal year TFP. This magnitude is economically large, given that the standard deviation of forecast errors of TFP, and TFP itself, are around 0.14 and 0.82 respectively. Quantitatively, information acqusition under these assumptions would eliminate almost all misallocation of capital due to uncertainty around capital budgeting. Specifically, when we compare this economy to an otherwise equivalent one with a degenerate distribution of information (but same average uncertainty), our economy sees only a 0.06% loss in measured TFP, whereas the comparison model predicts a 0.66% loss.

However, we find that the correlation in the data is substantially weaker. Consequently, when we calibrate our model to also replicate the allocation of information as implied by this statistic, we find TFP losses of around 0.53%, which is substantially closer to the economy with a degenerate information distribution. Specifically, our calibrated model reveals that high-productivity firms face a substantially higher marginal cost to reducing their uncertainty. We argue that this finding is intuitive: While our model describes all firms facing the same problem (i.e., reducing uncertainty), the types of uncertainty faced by different firms (and thus information required) are probably highly heterogeneous. For instance, smaller or less productive firms have relatively straightforward ways to improve productivity, such as streamlining production processes or upgrading technology. Meanwhile, larger and more sophisticated firms have more complex operations where searching for improvements may involve consolidating information across different processes or business lines.

We futher investigate the implications of this finding for firm value and the aggregate economy. To be precise, we consider an exercise where a hypothetical social planner can, through a series of subsidies and taxes, reallocate information towards higher productivity firms while holding the level of aggregate uncertainty fixed. Importantly, the social planner has the exact same information set as the firm, basing this reallocation entirely only on the intitial distribution of productivities. Holding wages fixed (i.e., in partial-equilibrium), we find substantial gains to firm values, ranging from 2.8% to 4.1% depending on initial conditions.

Moreover, we find that aggregate output rises by about 2.60% relative to our baseline, even though the implicit cost of such a reform is only 0.15% of pre-reform aggregate output. However, when we consider general-equilibrium effects, the gains are more muted, with aggregate output increasing by around 0.66%, although the reform cost remains essentially identical to the partial-equilibrium model. The key reason for this difference is because the improved allocation of information leads to a sharp increase in labor demand by high-productivity firms, which overwhelms the reduction in labor demand by low-productivity who see information reallocated away from them. This raises overall labor demand, substantially driving up wages, which in turn mutes the gains of information reallocation.

After discussing the related literature below, the rest of our paper proceeds as follows: Section 2 describes our model, and Section 3 presents our main theoretical results; Section 4 describes our data and presents our calibration strategy and results; Section 5 discusses the quantitative results of our calibrated model; and finally, Section 6 concludes.

Related Literature

Our results contribute to several strands of the macroeconomics and corporate finance literature. First, our paper builds on the large literature in corporate finance that studies the implementation of capital budgets and its implications for capital misallocation across projects within a firm (e.g., Harris and Raviv 1996; Malenko 2019; Graham 2022; Bernardo et al. 2004a). Closest to our research is Harris and Raviv (1996), who also argue that unanticipated shocks to productivity can lead to inefficient investment because it is difficult to adjust actual investment relative to budgeted capital expenses in real-time. However, the key theoretical focus of this literature has been on relating this inertia to optimal mechanism design as a result of intrafirm agency and information frictions (Harris et al. 1982; Antle and Eppen 1985; Harris and Raviv 1996; Bernardo et al. 2004b; Malenko 2019). In contrast, our paper takes partial inflexibility as a given, and examine firm incentives to acquire better information about their projects ex ante in light of these inflexibilities. Moreover, our focus is on the broader macroeconomic implications of capital budgeting quality, in particular, how variations in capital budgeting quality can explain capital misallocation across firms.

Second, our paper connects to the large literature on informational rigidity and rational inattention (e.g., Sims 2003; Reis 2006; Coibion and Gorodnichenko 2015; Ilut and Valchev 2020; see, in particular, the handbook chapter by Baley and Veldkamp 2021 and the references therein). While a general thrust of the literature has focused on the implications of rational inattention for business cycle and inflation dynamics (e.g., Van Nieuwerburgh and Veldkamp 2009; Maćkowiak and Wiederholt 2015; Baley and Blanco 2019; Ilut and Saijo 2020), our paper focuses on the interaction between rational inattention and capital budgeting, and connect this channel to a broader study on dynamic capital misallocation. Importantly, our paper emphasizes the complementarity between information and productivity at both the firm and aggregate levels. At the firm level, we show that higher productivity firms always have a natural incentive to acquire better information, since they value information more.[2] At the aggregate level, we show that a pure reallocation of information towards higher productivity firms can substantially ameliorate misallocation arising from uncertainty. As a whole, these results, to the best of our knowledge, is new to the literature.

Based on costly information acquisition, our theoretical results also help inform recent empirical findings on firm expectations formation. Recent findings suggest that higher productivity firms tend to make better forecasts over both microeconomic and macroeconomic forecasts, a finding that we also replicate in our data. For instance, Bloom et al. (2021) find that high-productivity firms make better forecasts of future firm revenue and Tanaka et al. (2020) find similar results with respect to macroeconomic aggregates. Different from these papers, our model *endogenizes* information acquisition, providing an explanation for empirically observed patterns of managerial forecasting based on *rational* agents with costly learning. Our model allows us to relate firm characteristics to the cross-sectional distribution of forecast errors. We argue that this heterogeneity is important for the overall aggregate effects of managerial beliefs. Our framework also allows for a number of interesting exten-

[2]Our emphasis on heterogeneity in returns to learning across firm types mirror recent research by Broer et al. (2020), although the authors focus on household heterogeneity and consequences for wealth inequality.

sions examining the interplay of how firms choose to acquire information and choose other variables such as corporate financial or investment policy. Our approach and inference is therefore similar to Farboodi et al. (2021), who study the incentives investors have to acquire information, and present a comprehensive framework to measure the private value of financial information.

Third, our findings relate to the voluminous research studying the sources of capital misallocation following the seminal work of Restuccia and Rogerson (2008) and Hsieh and Klenow (2009). In focusing on informational rigidity and investment inflexibility, our paper most directly relates to the literature emphasizing the role of uncertainty, time-to-build, and capital adjustment frictions as a source of capital misallocation (Asker et al. 2014; David et al. 2016; David and Venkateswaran 2019). However, our analysis emphasizes the role of information misallocation as the driver of capital misallocation across firms. Our emphasis on studying the underlying intangible factors backing the distribution of tangible factors of production mirrors recent research in the literature (e.g., Bhandari and McGrattan 2020; Whited and Zhao 2021).

Finally, from an empirical perspective, we add to the literature by directly verifying and quantifying our mechanisms using a unique novel data set of firm expectations, which should provide researchers a unique and rich setting to understand corporate investment behavior. Prior studies with investment plans use managerial forecasts made by a small number of public firms, who are only a special subset of the economy, or firm panels with qualitative forecasts regarding sales or investment (e.g., Bachmann et al. 2013; Bachmann and Elstner 2015; Bachmann et al. 2017; Barrero 2021; Bloom et al. 2021). Other studies involve experiments designed to elicit the expectations formations process of private individuals or firm managers, but do not explicitly study capital investment behavior and ex post firm profitability (e.g., Coibion and Gorodnichenko 2015; Fuster et al. 2020; Coibion et al. 2020). Importantly, this setting allows us to uncover a key finding: information acquisition costs, as estimated, are sharply increasing in firm productivity. In other words, high-productivity firms face a comparative disadvantage in information acquisition, despite evidence showing that high-productivity firms make more accurate forecasts.

2 Model

We model capital budgeting and expensing parsimoniously by incorporating three components: (i) information acquisition, where firms can invest in better information to improve the accuracy of their investment plans; (ii) investment plan formation, where firms lay out their investment plans for the year; and (iii) capital expensing, where firms decide how much capital to actually purchase, subject to inertia driven by costly deviation from the initial plans. After presenting the model, Section 3 shows how information plays a similar role to a factor of production through firms' capital budgeting process, and how the endogenous allocation of information across firms affect aggregate productivity.

2.1 Economic Environment

We follow a standard setup for agents, households, and final goods producers. We consider a discrete time infinite horizon model with infinitely lived agents. The economy is populated by a representative household and representative final goods producing firm, and heterogeneous intermediate goods producing firms. Each model period is 1 year, corresponding to our sample frequency. Although the household and final goods firm play a limited role in our analysis, we present their setup as they are necessary for studying the impact of information allocation on aggregate productivity. The representative household discounts time at rate β, inelastically supplies a fixed quantity of labor $N = 1$ and has preferences over consumption of a final aggregate consumption good. The representative final goods firm produces the final aggregate good given by $Y_t = \left(\int_0^1 y_{i,t}^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}}$, where i is a specific intermediate good and η is the elasticity of substitution across goods. The usual cost minimization problem for the final goods firm yields the standard demand schedule for each good i as $p_{i,t} = y_{i,t}^{-\frac{1}{\eta}} P_t Y_t^{\frac{1}{\eta}}$, where $P_t \equiv \left(\int_0^1 p_{i,t}^{1-\eta} di\right)^{\frac{1}{1-\eta}}$ is the CES price index, giving us $P_t Y_t = \int_0^1 p_{i,t} y_{i,t} di$. For the rest of the paper, we will set the final consumption good as the numeraire and normalize P to 1.

Each intermediate goods firm i produces a differentiated good based on the production function $y_{it} = z_{it} k_{it}^{\alpha} l_{it}^{1-\alpha}$, where z_{it} is idiosyncratic stochastic productivity, k_{it} is the beginning-of-period capital stock, l_{it} is labor hired by the firm, and $\alpha \in (0, 1)$ is the capital share of the firm. We assume $\log z_{it}$ follows $\log z_{i,t} = \rho \log z_{i,t-1} + \epsilon_{it}$, where $\epsilon_{i,t} \sim N(0, \sigma_{\epsilon}^{2})$. In addition, capital depreciates at a geometric rate δ. Finally, each unit of labor costs a wage w, and we assume that l_{it} is chosen after z_{it} has been observed. As such, firms choose labor to maximize static profits which are determined by the residual demand function above. Under these assumptions, without loss of generality, we can directly rewrite the firm's gross profit function net of labor cost as

$$\pi = \mathcal{A}(w_t, Y_t)\, z_{i,t}^{\Theta_z} k_{i,t}^{\Theta_k}, \tag{1}$$

where $\Theta_z \equiv \frac{\frac{\eta-1}{\eta}}{1-(1-\alpha)\frac{\eta-1}{\eta}}$, $\Theta_k \equiv \frac{\alpha\frac{\eta-1}{\eta}}{1-(1-\alpha)\frac{\eta-1}{\eta}}$, and $\mathcal{A}(w, Y) > 0$ is a function of the endogenous aggregate wage and output.

Besides labor, managers also decide on capital inputs, which are quasi-fixed and determined in the previous period. The role of the manager is to make a sequence of investment decisions to maximize the firm's net present value, which the manager discounts at constant rate $\frac{1}{1+r}$.[3] We lay out the capital budgeting and investment decision of the manager and discuss how capital budgeting affects the value of the firm below.

2.2 Capital Budgeting and Investment

As demonstrated in Figure 1, the investment problem is broken down into 3 sub-periods to reflect capital budgeting (sub-period I and II) and expensing in practice (sub-period III). First, a manager begins a fiscal year with full information on last period productivity z_{-1} but no information about the realization of TFP innovation ϵ (and thus, no information about z). Second, the manager must make a capital budget for the year; but before laying out the budget, she has the option to purchase better information about the year's prospects ϵ (e.g.,

[3]Because managers do not have any agency costs in our model, we interchangeably use the term "manager" or "firm" when discussing corporate policy.

by doing costly market research). Finally, after setting a budget, the manager observes the current year's productivity and decides how much labor to hire and how much actual capital to spend on investment. Importantly, any deviation of investment from the capital budget is costly, providing an incentive to make better investment plans at the beginning of the fiscal year. We discuss each sub-period in more detail below.

Sub-period I: Information Acquisition. The manager of firm i begins period t with full information on last period's productivity $z_{i,t-1}$, but does not observe $\epsilon_{i,t}$. However, she is able to improve on her information by acquiring a signal with precision $1/\sigma_{i,t}^2$. For some choice of $\sigma_{i,t}$, the manager will receive a signal $s_{i,t}$ about current productivity, given by $s_{i,t} = u_{i,t} + \epsilon_{i,t}$, where $u_{i,t} \sim N\left(0, \sigma_{i,t}^2\right)$. The manager learns by Bayesian updating: after observing the signal, the manager's belief over the posterior distribution of $z_{i,t}$ is given by $\log \tilde{z}_{i,t} \sim N\left(\log \hat{z}_{i,t}, \mathbb{V}_{i,t}\right)$, where $\log \hat{z}_{i,t} \equiv \rho \log z_{i,t-1} + \frac{\sigma_\epsilon^2}{\sigma_{i,t}^2+\sigma_\epsilon^2} s_{i,t}$. $s_{i,t}$ is the expected current period productivity and $\mathbb{V}_{i,t}$ is the posterior variance of productivity (equivalently, $\mathbb{V}_{i,t}^{-1}$ is the signal precision). The ratio $\frac{\mathbb{V}_{i,t}}{\sigma_\epsilon^2}$ captures the amount of learning, reflecting the amount of chosen uncertainty relative to the prior.[4] Given our information structure, we can define a TFP shock as follows:

Definition 1. (TFP shock) A TFP shock is defined as $\log z - \log \hat{z}$, that is, the log-difference between realized and forecasted TFP.

However, information acquisition is costly and follows a cost function,

$$\mathcal{C}\left(\sigma_{i,t}, z_{i,t-1}\right) = \xi z_{i,t-1}^{\zeta} \sigma_{i,t}^{-2}, \tag{2}$$

where ζ controls the degree to which information acquisition costs are correlated with initial

[4] Note that with our specification, u_{it} becomes degenerate at 0 when $\sigma_{i,t} \to 0$; in other words, this reduces to the standard timing whereby firms perfectly observe current productivity.

productivity $z_{i,t-1}$. This specification allows us to nest a variety of potential learning costs.[5] When $\zeta = 0$, the marginal cost of improving signal precision is constant and homogeneous across firms, whereas when $\zeta > 0$ ($\zeta < 0$), higher productivity firms face higher (lower) marginal costs. For the following two sub-periods, we drop the i subscript and denote variables from past productivity with z_{-1}.

Sub-period II: Investment Plan Formation. After observing the signal and forming an updated belief $\log \tilde{z}_t$ over the distribution of ϵ_t, the manager makes an investment plan k^p_{t+1}. We assume as a model primitive that these investment plans have to be made prior to making an actual investment in order to achieve next-period capital stock k_{t+1}. Our assumption is motivated by both capital investment practice and the corporate finance literature.[6] Importantly, managers have an incentive to make accurate plans because it is costly to deviate from the plan.[7] Given some plans and actual investment, the manager faces the

[5] We allow for this generic structure since there is no clear evidence as to whether the marginal cost of learning should be increasing, constant, or decreasing in firm productivity. Importantly, as we will show, our model structure allows us to identify this parameter, which will allow us to conduct our key counterfactual analysis later in the paper. In general, increasing marginal costs ($\zeta > 0$) can happen if larger (and thus higher productivity) firms also have more complicated management structures relative to smaller (less productive) firms, which leads to greater co-ordination failure across production units and chains of command (Harris et al. (1982); Malenko (2019)). Constant marginal costs arise in standard rational inattention models which formulates a capacity constraint or fixed costs of learning that are independent of productivity (e.g., Sims (2003); Reis (2006)). Finally, decreasing marginal costs ($\zeta < 0$) can occur if high-productivity managers naturally are better at information processing due to the implementation of better management systems (e.g, Bloom et al. (2007)), or if information acquisition has a degree of increasing returns to scale (e.g.,Begenau et al. (2018)).

[6] In practice, capital budgeting is a core process in the allocation of capital within a firm, and any deleterious effects of poor capital budgeting could arise for a variety of reasons (e.g., Mao 1970; Myers 1974; Schall et al. 1978; Arnold and Hatzopoulos 2000; Ryan and Ryan 2002), which is outside the scope of this paper. In the context of the macroeconomics literature, this relates to the idea of time-to-build and time-to-plan (e.g., Kydland and Prescott (1982); Christiano and Todd (1996)), where completely inflexible investment plans are model primitives.

[7] This assumption on costly plan deviation can be micro-founded by operational inefficiencies such as those modeled by Harris and Raviv (1996) and Malenko (2019), which feature information asymmetry within a firm and costly verification (auditing) of spending within a firm. In both models, the constrained optimal capital budgeting rule is to allocate a planned amount for capital expenditure, then incur costly verification for realized spending that deviates from plans. In this framework, costs are due to information frictions in the decentralized organization, such as auditing or various meetings.

following cost function $\phi(k^p_{t+1}, k_{t+1})$ for any deviation from the plan:

$$\phi\left(k^p_{t+1}, k_{t+1}\right) = \frac{\chi}{2}\left(\frac{k_{t+1}}{k^p_{t+1}} - 1\right)^2 k^p_{t+1}, \tag{3}$$

where $\chi \geq 0$ denotes the severity of the cost function and $\chi = 0$ implies that plans are irrelevant. Although the parametrization of the cost function is analogous to a traditional convex *physical* adjustment cost function, it is a penalty for *investment plan deviations* rather than deviations from the firm's current capital stock.[8] We show later in Section 3 that our key qualitative insights regarding information allocation do not hinge directly on this functional form assumption. However, this choice of functional form permits us to more closely match the data in our quantitative exercise. Finally, given our functional form assumption, it is convenient to define an investment plan deviation as follows:

Definition 2. (Investment Plan Deviation) An investment plan deviation is defined as $u_t \equiv \log k_{t+1} - \log k^p_{t+1}$, that is, the log-difference between realized and planned next-period capital.

Sub-period III: Investment. After making an investment plan, the manager perfectly observes current period productivity $z_{i,t}$. Knowing that z_t follows an AR(1) process, the manager uses her updated information to decide how much capital to invest for use in the next period, subject to the aforementioned adjustment cost $\phi(k^p_{t+1}, k_{t+1})$ in Equation 3.

2.3 Bellman Equations

We can now define the problem recursively. Let $J(k, k^p, z)$ denote the value function of the manager after all shocks have been realized (i.e., at sub-period III), $W(k, s, z_{-1})$ the value function of the manager after a signal has been observed (i.e., sub-period II), and finally $V(k, \tilde{z}, z_{-1})$ the value function of the manager at the beginning of the period (i.e., sub-period I). To simplify notation, we denote variables in the future with a $\cdot'$, drop both i

[8] This assumption also means that there are no physical adjustment costs inducing inertia in firm investment decisions.

and t subscripts, and denote past productivity with z_{-1}.

The manager's Bellman equation at sub-period III is given by

$$J\left(k,k^{p},z\right)=\max_{k'}\pi+\left(1-\delta\right)k-k'-\frac{\chi}{2}\left(\frac{k'}{k^{p}}-1\right)^{2}k^{p}+\frac{1}{1+r}\mathbb{E}\left[V\left(k',\tilde{z}',z\right)|z\right] \quad (4)$$
$$s.t. \quad \log\tilde{z}'\sim N\left(\rho\log z,\sigma_{\epsilon}^{2}\right),$$

where next-period capital k' is chosen to maximize expected profits, taking into account that deviations from investment plans are costly. While costs show up in the dynamic program of the manager, we assume that they are "utility" costs in our baseline model. Therefore, they do not consume any resources in terms of labor, capital, or profits and hence do not directly affect firm value.

Thus, the Bellman equation in sub-period II, after the signal has been observed, is given by

$$W\left(k,s,z_{-1},\sigma\right)=\max_{k^{p}}\mathbb{E}\left[J\left(k,k^{p},z\right)|s,z_{-1},\sigma\right] \quad (5)$$
$$s.t. \quad \log z\sim N\left(\rho\log z_{-1}+\frac{\sigma_{\epsilon}^{2}}{\sigma^{2}+\sigma_{\epsilon}^{2}}s,\mathbb{V}\right),$$

where the manager chooses investment plan k^p, taking into account the potential future need to deviate from the plan. The manager makes a capital budget k^p accounting for the full distribution of possible investment choices, since she does not know what current productivity is at this point.

Finally, in anticipation of the final two sub-periods, the Bellman equation in sub-period I is given by

$$V\left(k,\tilde{z},z_{-1}\right)=\max_{\sigma}-\xi z_{-1}^{\zeta}\sigma^{-2}+\mathbb{E}\left[W\left(k,s,z_{-1},\sigma\right)\right] \quad (6)$$
$$s.t. \quad s=u+\epsilon$$
$$u\sim N\left(0,\sigma^{2}\right),$$

where the manager chooses the signal quality σ of today's productivity, whereby a more precise signal is costlier to acquire but adds value to the firm by reducing the size of ex post investment plans deviations. The manager knows the productivity from the previous period z_{-1}.

2.4 Equilibrium Definition

A stationary competitive equilibrium is defined by a measure of firms Λ; a set of policy functions $\{\sigma(z_{-1}, k; w, Y), k^p(z_{-1}, k, s; w, Y), k'(z_{-1}, k, s, z; w, Y)\}$; a set of value functions $\{V, W, J\}$; intermediate good prices $\{p_{i,t}\}_{i\in\Lambda}$; wage w; and a Markov transition function Γ induced by the policy function k' and exogenous productivity z' such that (1) the distribution of firms is invariant ($\Lambda = \Gamma(\Lambda)$), (2) the labor market clears, and (3) the intermediate goods market clears.

Having laid out the model primitives, we next turn to a discussion on how these firm-level decisions affect the aggregate information allocation across firms and discuss implications for aggregate productivity.

3 Capital Budgeting: Micro Choices and Macro Outcomes

Because capital budgeting involves acquiring and using costly information, information behaves as an implicit factor of production. This section studies the information acquisition choice that a firm makes, the distribution of information acquisition across firms, and the resulting impact on aggregate productivity. Finally, we discuss our identification strategy for mapping the model to the data and present the quantitative implications of our model in Section 5.

3.1 Within-Firm Information Acquisition

Firms have an incentive to acquire better information and make more accurate plans because deviating from investment plans is costly. To illustrate this, we focus first on the impact on firm profits assuming that investments are completely unadjustable once set in place (i.e., $\chi = \infty$). This assumption isolates the information channel, without complicating the analysis when allowing for partial adjustment, and allows us to derive a global solution. Appendix B contains all relevant formal proofs for the propositions below.

Proposition 1. *(Marginal Benefit of Learning 1) Suppose $\chi = \infty$. Then,*

1. *the ex ante (sub-period I) expected value of the firm is strictly decreasing in the posterior uncertainty $\mathbb{V}$, and*

2. *the value of increasing signal precision (i.e., decreasing $\mathbb{V}$) is increasing in initial firm productivity z_{-1}. That is, information and productivity are complements at the firm-level.*

Proposition 1 summarizes two aspects of capital budgeting. First, it points out that information is desirable, and all firms prefer to face lower uncertainty since it increases firm value. This effect can be seen in Equation 7, which expresses the overall marginal benefit of information as a constant multiplied by two terms, $z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}$ and $\frac{\partial \mathcal{F}^{\pi}(\mathbb{V}^{-1})}{\partial \mathbb{V}^{-1}}$, that correspond to a productivity effect and information effect, respectively. We show that $\frac{\partial \mathcal{F}^{\pi}(\mathbb{V}^{-1})}{\partial \mathbb{V}^{-1}} > 0$, therefore capturing the desirability of information for all firms.[9]

$$MB \equiv \frac{1}{1+r}\frac{\partial \mathbb{E}\left[V\left(\prime\right)\right]}{\partial \mathbb{V}^{-1}} = \frac{1}{1+r}\bar{h} \underbrace{z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}}_{\text{pure productivity effect}} \underbrace{\frac{\partial \mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial \mathbb{V}^{-1}}}_{\text{pure information effect}} \tag{7}$$

Second, we can see from the expression above that the term $z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}$ is increasing in z_{-1}; in other words, the marginal benefit from lowering uncertainty (i.e., incentive for information acquisition) increases in productivity. Figure 2 plots the marginal benefit of information

[9]Note that this is different from the Abel-Hartmann-Ooi effect, which emphasizes a volatility effect. This effect is captured in the $\bar{h}$ constant, which is a function of σ_{ϵ}^2. The exact derivation of $\bar{h}$ is detailed in Appendix B.

acquisition for a high and low productivity firm.[10] Beyond delivering intuition and a clear solution, the simplified model with $\chi = \infty$ has also an added advantage in emphasizing that our baselines results are not driven by our parametric assumptions on the investment plan adjustment cost channel. Rather, this result derives entirely from standard assumptions on the revenue generating function of the firm.

[Figure 2 Here]

Moreover, although higher productivity firms always have a higher benefit of learning, the amount of endogenous learning also takes into account the marginal cost function of information acquisition, stated in Proposition 2 below.

Proposition 2. *(Net Benefit of Learning) For some threshold $\bar{\zeta} > 0$, the optimal signal precision is weakly increasing in initial firm productivity when $\zeta < \bar{\zeta}$, weakly decreasing when $\zeta > \bar{\zeta}$, and does not depend on productivity when $\zeta = \bar{\zeta}$.*

The indifference condition (derived in Appendix B) that determines the firm's choice of $\mathbb{V}$ is given by

$$\underbrace{\frac{\partial \mathcal{C}}{\partial \mathbb{V}^{-1}}}_{MC} \equiv \xi z^{\zeta} = \frac{1}{1+r} \bar{h} z_{-1}^{\frac{\rho^2 \Theta_z}{1-\Theta_k}} \frac{\partial \mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial \mathbb{V}^{-1}} \equiv \frac{1}{1+r} \underbrace{\frac{\partial \mathbb{E}\left[V\left(^{\prime}\right)\right]}{\partial \mathbb{V}^{-1}}}_{MB},$$

where the right side of the equation is the marginal benefit formula from Equation 7. This indifference condition is analogous to standard equations connecting physical factor inputs to firm production. Just as factor inputs and productivity are complements in production, so are information and productivity once we account for endogenous learning. While information is not an explicit input into production, our reformulation of the firm's marginal benefit

[10] To put our model mechanism in the context of a real life example, consider the impact of capital budgeting mistakes on the profitability of a large grocery store chain as compared to a "mom-and-pop" corner store. Our model predicts that the same relative mistake (i.e., same percentage investment plan deviation) will naturally have a larger *level* impact on the bottom line of the large chain as compared to the corner store. This difference then leads the large chain to preemptively make better investment plans (for instance, by engaging external consultants) relative to the corner store.

function re-expresses all physical factor inputs as functions of information choices. This is similar to recent advances in the finance literature emphasizing the role of financial factors as an implicit input into firm production and value added (Whited and Zhao (2021)).

On the input cost side of the equation, our formulation of $\mathcal{C}$ is analogous to models where the implicit cost of factor inputs is correlated with the idiosyncratic state of the firm (e.g., where the tightness of financial constraints correlate with the cash flow or productivity of the firm as in Hennessy et al. (2007); Midrigan and Xu (2014)). In our case, we interpret ζ as a reduced-form "wedge" that affects overall information allocation in the spirit of the literature on firm-level wedge accounting (e.g., Hsieh and Klenow (2009); David and Venkateswaran (2019); Whited and Zhao (2021)). Correspondingly, we consider an economy where $\zeta = 0$ as a "first-best" reference where information allocation is not distorted by implicit wedges. For illustration, Figure 2 shows two examples, one where the marginal cost is constant (i.e., $\zeta = 0$) and one where the net benefit is constant (i.e., $\zeta = \bar{\zeta}$). When $\zeta = 0$, the optimal signal precision is increasing in productivity, reflecting the natural complementarity between information and productivity; in contrast, when $\zeta = \bar{\zeta}$, all firms choose identical signal precisions because the cost of information acquisition is rising sharply in productivity and overwhelms the complementarity between information and productivity.

Extant research using firm-level forecasting data (e.g., Tanaka et al. (2020); Bloom et al. (2021)) has shown that higher productivity firms appear to formulate more precise forecasts over fundamental variables associated with firm profitability (e.g., sales, GDP), leading to the hypothesis that higher productivity firms have a comparative advantage in forecasting. Our model delivers the same effect but provides an alternative interpretation: higher productivity firms make better forecasts because they have higher *incentives* to acquire better information. In contrast, the comparative advantage of firms in forecasting depends on the marginal cost of information acquisition. Our model does not take a stand on this, allowing for high-productivity firms to exhibit comparative advantage relative to low-productivity firms when $\zeta < 0$ (and vice versa when $\zeta > 0$). In the next section, we will use our data to quantify the degree to which high-productivity firms do exhibit a comparative advantage with regards to forecasting by inferring the value of ζ using our identification strategy discussed below.

Finally, we now return to discussing the role of the investment plan adjustment cost in our model. While our global result in Proposition 1 is derived under the special case where investment plans are immutable ($\chi = \infty$), the results remain when we return to the general model under a local approximation, as formally stated in Proposition 3 below:

Proposition 3. *(Marginal Benefit of Learning 2) Suppose* $0 < \chi < \infty$*. Then the results of Proposition 1 are preserved under a log-linear approximation for next-period capital choice* k'*.*

This gives rise to a natural corollary:

Corollary 1. *(Net Benefit in General Model) The results of Proposition 2 are preserved under a log-linear approximation for next-period capital choice* k' *for the general model with* $0 < \chi < \infty$.

While the inclusion of partially flexible investment plans might seem superfluous to our focus on information allocation, Section 4 shows that empirically, firms appear able to deviate from their plans in response to real-time productivity shocks. This implies that an assumption of immutable plans will overstate the effects of uncertainty and information reallocation. Therefore, to estimate ζ and conduct quantitative counterfactual exercises, we adopt a simple cost function to improve our model fit to the data and avoid the risk of overstating the quantitative role of information.

3.2 Across-Firm Information Allocation and Aggregate Productivity

Endogenous learning depends on firms' heterogenous productivities, which when aggregated, affect economy-wide productivity. We illustrate this by presenting an aggregation result relating arbitrary allocations of information to aggregate total factor productivity (TFP) for a simplified version of our model with $\chi = \infty$. We define $TFP \equiv \frac{\left(\int y^{\frac{\eta-1}{\eta}} d\Lambda\right)^{\frac{\eta}{\eta-1}}}{\left(\int k d\Lambda\right)^{\alpha}\left(\int l d\Lambda\right)^{1-\alpha}}$, where Λ is the distribution of firms and Proposition 4 states the result.

Proposition 4. *Information and productivity are complements in firm production at the aggregate level. For any given marginal distributions of initial period productivity z_{-1} and information choice $\mathbb{V}$, TFP is decreasing in the correlation between z_{-1} and $\mathbb{V}$.*

The result is analogous to a model where the distribution of physical factor inputs are arbitrarily imposed and the total stock of factors are fixed such that TFP gains arise solely from factor reallocation. In the context of physical factors, because productivity and factors of production are explicitly modeled as complements, it is obvious that TFP would increase if factors are reallocated towards higher productivity firms.

Our contribution is to show that a similar result holds for productivity and information—an implicit factor of production—where such complementarity is neither explicit nor obvious *prima facie*. Crucially, Proposition 4 also reflects how the information allocation aspect of capital budgeting can affect aggregate TFP. In our model, the parameter ζ determines the net benefit of information acquisition and is key to whether high productivity firms actually acquire more information in practice. Quantifying the importance of this channel will be a key exercise in the following sections.

3.3 Model Identification and External Validation

Having developed the theoretical framework to show how information acquisition through capital budgeting affects input choices, our next goal is to quantify the degree to which this information allocation mechanism affects firm value and aggregate output. We proceed in two steps. First, we present our identification strategy that connects the capital budgeting decision of the firm to key identifying moments. Next, we present two additional predictions of our model, which we will use to externally validate the primitive assumption that investment plans are not fully flexible but also not completely immutable.

3.3.1 Model Parameter Identification

Following the literature, the bulk of our model parameters either use standard values or are estimated independent of the model. Three remaining keys parameters (χ, ζ, ξ) need to be estimated. We estimate these parameters by targeting three moments and connect each moment to the corresponding parameter that it identifies.

Identification 1. *The average cross-sectional dispersion of TFP shocks, with TFP shocks defined as in Definition 1, depends on the intercept of marginal cost (ξ).*

Identification 2. *The average cross-sectional dispersion of investment plan deviations, with investment plan deviations defined as in Definition 2, depends on the adjustment cost parameter (χ).*

Identification 3. *The correlation of lagged TFP with the dispersion of TFP shocks depends on the slope of the marginal cost to information acquisition (ζ).*

We briefly discuss the intuition behind our identification strategy and defer more detailed and formal derivations to Appendix B.5. First, our marginal cost parameter ξ is identified using the dispersion of TFP shocks because ξ determines the average cost a firm needs to pay to acquire information. Per our model definition, TFP shocks are simply forecast errors. Consequently, since increases in ξ reduce the average amount of information acquired, the average size of forecast errors will increase, leading to a larger dispersion of TFP shocks.

Second, our adjustment cost parameter χ most directly affects the dispersion of investment plan deviations since it affects the degree to which plans are adjustable. However, an issue is that both χ and ξ are intrinsically tied together: a higher adjustment cost χ (independent of marginal cost ξ) increases the benefit of information and leads to more information acquisition, reducing the average dispersion of TFP shocks and plan deviations, while a lower ξ also reduces the dispersion of TFP shocks and investment plan deviations. However, because χ has an *indirect effect* on information acquisition while ξ has an *indirect effect* on investment adjustment, the non-linear map from χ and ξ to the two moments permit identification. We formally derive this result in the Appendix (see Equation 28) and show that the *ratio* of these two moments is in fact only a function of χ. To show the intuition of how the parameters are identified, we construct "isomoment" plots in Figure 3 to illustrate the intuition. Here, we construct a series of graphs that report the set of parameter pairs that give rise to the same targeted moment, holding all else constant. In the case of identifying ξ and χ, it is most useful to focus on the first two rows of figures, where we see that every pair of ξ and χ maps uniquely into the two aforementioned moment targets, whereas multiple pairs of ζ (and ξ or χ) can map into the same two moment targets.

Finally, ζ most directly relates to the correlation between uncertainty and firm productivity (Proposition 3 and Corollary 2). Most importantly, our cost function assumptions imply that the map between these variables is monotone, and that a firm's information acquisition policy will depend on its initial productivity except for the knife-edge case when $\zeta = \bar{\zeta}$. This effect is reflected in the last two panels of our isomoment plots (Panels h and i of Figure 3). Moreover, although ζ will also have an impact on the first two moments, as we already pointed out, the map between parameter pairs of ζ (and ξ or χ) to the first two moment targets is not unique. This result allows us to "isolate" ζ via the last moment target.

3.3.2 External Validation of Model

We now discuss two additional predictions that we use to validate our model. Confirming that these predictions hold in the data helps us to verify our main model channels.

Prediction 1. *In a "horse race" regression for forecasting actual firm investment, investment plans will attenuate the forecasting power of expected productivity for actual investment.*

Intuitively, the less flexible the plan, the less actual investment should adjust to TFP shocks. Prediction 1 serves to reject the null that investment plans are fully flexible, which is important since capital budgeting is only relevant if investments plans are not fully flexible. This prediction arises from the investment plan equation (Equation 23 of Appendix B), where we show that investment plans are a function of both expected productivity and the precision of these expectations. Since actual investment itself depends on investment plans, investment plans will have a stronger forecasting power than expected productivity alone, since the latter accounts for only a fraction of the variation in the former. In Appendix B.6.2, we formally show that the degree to which investment plans attenuate the forecasting power of expected productivity decreases in the flexibility of investment plans. Our next prediction formally tests whether investment plans are immutable.

Prediction 2. *The size of investment plan deviations is increasing in the size of realized TFP shocks.*

This prediction arises directly from the investment Euler equation (equation 21 of Appendix B), where we show that except for the limiting case when plans are completely immutable, actual investment will in general be positively correlated with TFP shocks. Therefore, this prediction relates to the assumption that investment plans are at least partially flexible and clarifies why we assume an adjustment cost function in our quantitative model rather than simply assuming immutable plans.

Having discussed the link between firm-level capital budgeting to firm value and aggregate productivity, we turn to quantifying the impact of this mechanism, which requires a calibration of model parameters. A challenge to calibrating our model is data availability, as the identification criteria above requires observing three key variables that are not typically jointly available in widely used datasets: (1) expected (i.e., forecasted) and realized TFP, (2) capital budgets (i.e., investment plans) and realized capital expenditures, and (3) the correlation of the variation of productivity shocks with past productivity. In the next section, we introduce a novel dataset from Japan which includes all of the necessary information.

4 Empirical Analyses

In the following subsections, we discuss our main data sources and the definition and construction of the key variables, present reduced-form analyses that validate the qualitative predictions of our model, and finally calibrate our model and discuss the results and model sensitivity.

4.1 Data and Sample Construction

Our sample comes from combining two datasets from the Japanese Ministry of Finance (MoF): the Business Outlook Survey (BOS) and the Annual and Quarterly Financial Statements Statistics of Corporations by Industry. The BOS contains firm-level forecasts/spending plans and their realizations, while the Annual (Quarterly) Financial Statistics survey provides detailed year-end (quarter-end) financial statistics such as cash holdings, debt, total employment, and cost of labor for the fiscal period. Although the two surveys are distinct, they follow the same sampling procedure and target the same firms (Appendix C). Responding to both surveys is mandatory; the response rates for large firms in our sample are between 87% and 88.8%.[11] Merging the data creates a nationally representative panel of both public and large private firms in Japan with information on their expectations and financial statements.

4.1.1 Summary Statistics

Our main sample includes all industries but drop firm-year observations with zero sales, missing labor costs, and missing net plants, property, and equipment information. The final sample contains around 6,000 firms a year, ranges from fiscal year 2005 to 2016, and accounts for around 60% of total employment in Japan.[12]

[Table 1 Here]

Table 1 reports the general characteristics of our merged sample, including the mean, standard deviation, skewness, and the 25th, 50th, and 75th percentile of key variables. Panel A reports balance sheet variables from the Annual Financial Statistics survey and Panel B reports constructed variables. The average total assets of a firm in our sample is 88 billion yen (approx. US$7,500,000) with a median of 26.7 billion yen (approx. US$2,300,000),

[11] Charoenwong et al. (2020) considers the strategic reporting motives of firms and shows that the forecasted profit and sales data exhibit similar, if not less biased and more accurate, distributional properties than publicly reported forecasts among publicly listed companies.

[12] In Japan, a fiscal year ranges from April to March. For example, fiscal year 2005 ranges from April 2005 to March 2006.

and the average capital stock of 50 billion yen (approx. US$4,300,000) with a median of 12.3 billion yen (approx. US$1,060,000). On average, firms plan to invest 7.7% of capital and actually invest around around 7.6% of capital. Although the interquartile range of the investment plan deviations are roughly the same order of magnitude as for the actual realized investment, the standard deviation is 5.6% compared to the realized investment rate of 9.6% and the distribution of investment plan deviations as a percentage of capital is tighter and less skewed than those of the realized or planned investment rate.[13]

4.1.2 Definition and Construction of Key Variables

Firms are surveyed around six weeks into the first quarter about their forecasted sales, profits, and investment spending plans for the full fiscal year. For variables derived from sales and profits, we refer to the differences between actual and realized variables as "*shocks*." Our interpretation, following from our model assumptions and consistent with Bachmann and Elstner (2015), is that differences between actual and predicted sales or profits are largely driven by circumstances outside of a firm's control and thus can be interpreted as innovations to a firm's information set. In contrast, for investment plans, we refer to the differences between realized and planned values as "*deviations from plan*", since investment spending is under a firm's control and thus differences between realized and planned values are *choices* made by a firm. Such deviations reflect both a firm's reactions to real-time innovations to its information set and the degree of flexibility to adjust realized capital expenses relative to its planned expenses.[14]

[13] While not a focus of the paper, we report more detailed investment characteristics in Table C.4 (Appendix C.2), drawn from the Annual Financial Statistics survey. Around 16% of firms in our sample exhibit negative gross investment, while around 17% report a gross investment rate in absolute value lower than 1%. In general, the statistics involving investment are comparable to public firms in the United States (for instance, see Clementi and Palazzo (2019) for statistics computed using data from Compustat North America).

[14] To verify that the BOS corporate plans and forecasts are economically relevant, we conduct regression analyses similar to that of Gennaioli et al. (2015), and report the results in Appendix Section A.1. Observed capital spending plans appear informative for realized spending, generating results quantitatively similar to or even larger than those in Gennaioli et al. (2015).

TFP, Expected TFP, and TFP Shocks The firm "fundamental" we consider is revenue-based total factor productivity (TFP). For computing realized TFP, we follow Asker et al. (2014) and assume firms operate Cobb-Douglas production functions with capital (labor) intensity α $(1-\alpha)$ and face isoelastic demand curves with elasticity η. Given values for α and η, we can back out TFP z using the identity

$$z = \frac{(py)^{\frac{\eta-1}{\eta}}}{k^{\alpha} l^{1-\alpha}}, \tag{8}$$

where py is the firm's total value-added for the fiscal year, k is the firm's physical capital stock measured at the end of the last fiscal year, and l is the number of full-time equivalent labor hired as of the end of the current fiscal year. We defer our full description to Appendix A.2.

We also compute an *approximate* measure of the firm's expected TFP using data on expected sales and balance sheet variables and use it to infer the TFP shock the firm faces. This is unlike prior research, which typically uses the fitted value of an AR(1) regression as a proxy for a firm's expected TFP and the residuals as the TFP shock a firm faces. Specifically, we define expected TFP z^e as

$$z^e \equiv \frac{(py^e)^{\frac{\eta-1}{\eta}}}{k^{\alpha} l^{1-\alpha}}, \tag{9}$$

where py^e is the firm's *expected* total value-added for the fiscal year (which we compute from expected sales), and k and l are the same as above. We define a TFP "shock" in the data $\tilde{\epsilon}$ as

$$\tilde{\epsilon} \equiv \log z - \log z^e. \tag{10}$$

This follows our model definition of a TFP shock as being the log-difference of realized and forecasted TFP ($\log z - \log \hat{z}$). However, because the definition above uses only an approximation of expected TFP (specifically, due to Jensen's inequality, $\log \hat{z} \neq \log z^e$), it consequently mismeasures the TFP shock. Nonetheless, in Appendix A.2.4, we show that given the timing convention and parametric assumptions of our model in Section 2, these measurement issues do not affect our qualitative results and are unlikely to be quantitatively large. The main intuition is that $\log z^e$ essentially suffers from classical measurement error, introducing an attenuation bias when expected TFP is used as an explanatory variable in our analyses, which does not affect the sign of a regression coefficient. Relatedly, for the case of $\tilde{\epsilon}$, the TFP shock is (i) amplified by a fixed factor that is a function of only η and α, meaning the measured shock will be too large relative to the true shock, and (ii) suffers from measurement error coming from variations in the firm's degree of uncertainty (i.e., $\mathbb{V}$). The former can be directly corrected for in our analyses, and the latter variation is quantitatively small relative to log realized TFP.

Finally, we need to compute firm value-added to operationalize our measurement strategy. Unfortunately, the Financial Statistics survey does not give us a direct measure of intermediate inputs, so we cannot compute value-added directly. Therefore, we impute the firm's material share based on industry input-output tables provided by the Ministry of Economy, Trade and Industry (analogous to those produced by the Bureau of Economic Analysis in the United States) which follows standard approaches in the literature (e.g., David and Venkateswaran 2019). Our exact measurement strategy is detailed in Appendix A.2.

To ensure that our strategy is robust, we also consider an alternative measure of value-added using sales minus the costs of goods sold. As a robustness check, we verify that all of our results would be qualitatively similar using this measure as a measure of value-added. However, an issue with using costs of goods sold is that while it captures the total cost of intermediate inputs, it also includes part (but not all) of labor costs that would systematically bias our estimates of value added and labor intensity downwards. In contrast, our imputed material shares approach introduces measurement error without systematic bias. Therefore, we choose the imputed material share approach as our baseline estimate.

Investment, Investment Plans and Deviations from Investment Plans Firms are surveyed regarding their planned and actual investment-related spending, including information over three broad categories of spending: plants, property and equipment (PP&E), land, and software. We use PP&E as our definition of "capital", and consider investment *rates* rather than investment *levels* by normalizing investment values by the end-of-previous-period capital stock. Planned investment (denoted by i^p) are reported planned investments from the beginning of the fiscal year, while actual investment (denoted by i) are the realized spending at the end of the fiscal year. For investment plan deviations, we use definition 2 by imputing next-period capital stock using the perpetual inventory method with a fixed geometric rate of depreciation.

A caveat for our empirical analysis is that the investment values in the survey data only reflect capital expenses. This means that firms that plan to and / or actually disinvest will report $i^p = 0$ and $i = 0$. This data truncation biases our measure of investment plan deviations toward 0, which means that we will underestimate investment flexibility. We discuss this bias in more detail in Appendix A.2.6 and argue that based on our simulations, we do not believe that this bias meaningfully affects our empirical results. In addition, we will also incorporate this measurement error in the model's calibration to the relevant targeted moments.

4.1.3 Summary Statistics of Key Variables

As discussed in the previous section, our calibration requires three pieces of information: (1) dispersion of TFP shocks, (2) dispersion of investment plan deviations, and (3) the correlation of TFP with dispersion of TFP shocks. While we do not directly observe these three statistics in the data, we can compute their empirical counterparts per our definitions in the preceding subsections. For additional information, we report the statistics for TFP and investment at the industry level in Tables A.2 and A.3, respectively. Finally, in the next section, we document the correlation between lagged productivity and the dispersion of TFP shocks.

We highlight two important caveats. First, as we already noted, the above-mentioned TFP dispersion is mismeasured. Consequently, as we can see in Column 6 of Table A.2, the dispersion of TFP is often larger than that of the corresponding volatility of TFP. However, we show in the Appendix that, we can directly invert back the true TFP shock dispersion if we assume homogenous uncertainty (Equation 18). Therefore, as an additional check, Column 7 shows readjusted values according to our formula. We see that all values are now typically around one-half that of the volatility of TFP. Second, in the BOS, we only observe i and $i \geq 0$. We will account for these empirical issues when calibrating our model.

4.2 External Empirical Validation

We now present reduced-form empirical analyses to validate our model's extended predictions. We report validation tests in Table 2. For notation, subscript i indexes a firm, $j(i)$ indexes the industry firm i is in, and t is a fiscal year.

First, we run a regression specification of the form

$$\frac{i_{i,t}}{k_{i,t}} = \alpha_{j(i),t} + \beta \frac{i^p_{i,t}}{k_{i,t}} + \gamma \log z^e_{i,t} + \varepsilon_{i,t}, \tag{11}$$

where $\alpha_{j(i),t}$ represents industry and year fixed effects. The coefficients of interest are γ, which captures the relation between expected TFP and investment, and β, which captures the relation between investment plans and investment. To conduct the "horse race", we first run the regression with only expected TFP alone to show the benchmark result, and then run the same regression again including investment plans. Column 1 shows the relationship between investment and expected TFP. Column 2 adds investment plans as a regressor. As predicted, while expected TFP is strongly and positively correlated with investment, the presence of investment plans sharply attenuates the importance of these expected TFP by over two-thirds and improve R^2 almost 7-fold. This finding suggests that investment plans might indeed feature a high degree of inflexibility.

[Table 2 Here]

Prediction 2 suggests the size of plan deviations is increasing in the size of TFP shocks. To test this, we run a regression specification of the form:

$$\log\left(1+|u_{i,t}|\right)=\alpha_{j(i),t}+\beta\log\left(1+|\tilde{\epsilon}_{i,t}|\right)+u_{i,t}, \tag{12}$$

where $\alpha_{j(i),t}$ represents industry and year fixed effects. The coefficient of interest is β, which relates the size of investment deviations to the size of TFP shocks. We report our result in Column 3, which shows that a 1% increase in the size of TFP shocks leads to a 0.13% increase in the size of investment plan deviations.

Finally, in Column 4, we present a third statistic of interest: the correlation between lagged productivity and the dispersion of TFP shocks. We proxy for the dispersion of TFP shocks using the absolute size of firm forecast errors for TFP, as given by $\log\left(1+|\tilde{\epsilon}|\right)$. We then run a regression specification of the form

$$\log\left(1+|\tilde{\epsilon}_{i,t}|\right)=\alpha_{j(i),t}+\beta\log z_{i,t-1}+\eta_{i,t}, \tag{13}$$

where for notation, subscript i indexes a firm, $j(i)$ indexes the industry firm i is in, t is a fiscal year, and $\alpha_{j(i),t}$ represents industry and year fixed effects. The coefficient of interest is β, which relates the size of TFP forecast errors to the firm's last period productivity. The results show that for our baseline measure of TFP, a 1% increase in the level of TFP leads to a 0.009% decrease in the size of TFP forecast errors. In other words, realized TFP in the previous fiscal year is negatively correlated with the size of TFP shocks in the current fiscal year. Table A.5 in Appendix A.2 also reports this correlation at the industry level.

4.3 Model Calibration

We now briefly discuss our model calibration strategy, having described our data source. First, we fix a subset of parameters based on values from existing literature: $r = 0.02$, consistent with an annual real interest rate of 2%; $\delta = 0.06$ and $\eta = 4$ following standard assumptions; and set $N = 1$ and $P = 1$ as a normalization. Then, we set another subset of parameters (σ_ϵ, ρ, α) based on direct estimation from the data. These parameters are estimated at the industry level, but we use only the estimates for the manufacturing sector in our benchmark calibration. For our baseline calibration, we use the estimates derived with an assumption of $\eta = 4$, following the literature (Bloom et al. 2018). As a robustness check, we also consider $\eta = 3$ and $\eta = 6$, spanning the range of commonly used values for the elasticity of substitution. All parameter values are reported in Panel A of Table 3. A detailed description of the estimation strategy, as well as full estimates for all industries, are deferred to Appendix A.2.

Finally, we calibrate our model in general equilibrium for the remaining three key parameters: the cost of information acquisition ξ, the cost of deviation of investment plans χ, and ζ, which captures the relationship between the marginal cost of information acquisition and productivity. ξ, χ, and ζ are jointly calibrated to three moments: (i) dispersion of TFP shocks relative to the volatility of TFP, (ii) the dispersion of investment deviations, and (iii) the regression coefficient on lagged productivity. This strategy follows our identification argument as discussed in Section 3.3. Moreover, we simultaneously solve for the endogenous wages and aggregate output, such that these values are consistent with our market-clearing conditions at the calibrated parameter values.[15] All calibrated parameter values are reported in Panel B of Table 3.

We now address an important concern with calibrating our model to the empirical measures of TFP and investment. As explained in Section 4.1.3, TFP shocks and investment

[15]To compute model moments in the steady state, we simulate a panel of 10,000 firms for 50 years and drop the first 30 years as a burn-in.

plan deviations are mismeasured due to the exact variable definitions in the survey. To address the issue of constructing expected TFP from expected sales, we recreate the exact same mismeasurement in our calibration; similarly, to address the truncation bias in investment expenditures, for any simulated firms that made (or planned on making) a negative investment, we replace i and i^p with zeros in our calibration.

4.4 Calibration Results and Sensitivity

We report our calibrated parameter values in Panel B of Table 3, while Panel A of Table 4 reports our calibration results for our benchmark and robustness exercises, demonstrating a nearly perfect model fit for all three specifications. Moreover, we find qualitatively similar parameter values across all three specifications.

We further validate the model's calibrated parameters by considering the same regressions as in Section 4.2 and comparing the model-predicted regression results with the data. The results are reported in Panel B of Table 4. For Prediction 1, we report the relative change in the coefficient on TFP before and after controlling for investment plans, while for Prediction 2, we report the regression coefficient on TFP shocks. Relative to the data, our calibrated model predicts a weaker attenuation from including investment plans as an additional control variable, and larger responses of investment deviations to TFP shocks, indicating that our model overstates the degree of investment plan flexibility. However, we believe the fit is reasonable given the simplicity of our model. Importantly, this over-prediction of investment flexibility implies that any of the negative effects of capital budgeting as practiced, as we discuss later, can be seen as a lower bound of estimates, since the quantitative impact of capital budgeting is decreasing in investment flexibility. Finally, given that our results are qualitatively similar across all specifications, in the interest of space, we only discuss them in the context of our baseline calibration.

While investments appear to be somewhat flexible, we find that the total implied cost incurred by information acquisition is almost fifteen times that of investment plan adjustments (when interpreted as resource costs, they amount to around 0.15% and 0.012% of aggregate output respectively). This implies that there is a strong preference by firms to be correct ex ante rather than to respond to forecast errors ex post by deviating from plans. As such, we focus on the learning channel in this paper, and leave for future research a further study of the underlying microfoundations of investment flexibility. This finding also shows why we needed to model some flexibility, since ignoring it would overstate the importance of learning.

Figure A.1 reports the full distribution of firm residual uncertainty in terms of the posterior uncertainty $\mathbb{V}$ as a share of prior uncertainty σ_ϵ^2 (i.e., $\frac{\mathbb{V}}{\sigma_\epsilon^2}$), and reports the broad statistics associated with it. We see that an average firm has a residual uncertainty of around 27.6% that of its prior, and that there is relatively small dispersion in firm uncertainty, with an interquartile range of around 3.1%. Our overall findings suggest that there is substantial resolution of uncertainty, similar to recent findings in the literature that rely on indirect inference techniques.[16] Notably, if we had directly applied our formula (Equation 18) for inverting out the true dispersion of TFP shocks, we would have obtained a similar ratio of around 27.8%. This is consistent with our argument that the mismeasurement driven by unobserved heterogeneity in $\mathbb{V}$ is relatively small.

In the baseline calibration, ζ is estimated to be large and positive, implying that the marginal cost of learning increases sharply in firm productivity. Our finding implies that high-productivity firms face a comparative disadvantage in information acquisition relative to low-productivity firms. That said, the negative correlation between productivity and firm uncertainty that we find in the data indicates that the gross benefits of information outweigh the costs, leading high-productivity firms to acquire relatively more information. Our results provide a useful interpretation for similar recent empirical findings that high-productivity firms make more precise forecasts of profitability and related fundamentals (e.g., Tanaka et al. (2020); Bloom et al. (2021)). While a natural hypothesis is that high-productivity firms have a comparative advantage in forecasting, our results suggest the opposite: the correlations

[16] For instance, David et al. (2016) develop a model with similar timing restrictions to ours and find $\frac{\mathbb{V}}{\sigma_\epsilon^2} \approx 0.41$ in the United States.

observed in the data happen *in spite of* the comparative disadvantage high-productivity firms have.[17]

To further investigate the importance of private incentives for information acquisition with respect to aggregate information allocation, we next calibrate a counterfactual model imposing $\zeta = 0$, and target only the first two moments. In other words, we uncover the counterfactual distribution of information across firms, holding the average investment uncertainty and flexibility fixed. Doing so allows us to quantify the degree to which higher productivity firms would have acquired more information if marginal costs were equal across firms. We find that high-productivity firms have substantial private incentives to acquire better information; when $\zeta = 0$, the correlation between productivity and uncertainty becomes almost eight times higher (Panel B of Table 4). Moreover, while the average uncertainty is identical across both models, the counterfactual predicts a substantially larger dispersion—with an interquartile range of around 27.2%—indicating much stronger sorting of firms into high- and low-information.

[Table 4 Here]

Our results so far suggest that there is a substantial role of information "misallocation". As discussed in Section 3, we view our parametrization of $\mathcal{C}$ as akin to a form of dynamic wedge accounting, similar to the approach by David and Venkateswaran (2019). At the individual firm level, our counterfactual suggests that high-productivity firms are substantially constrained in the amount of learning they can do, which might negatively impact firm value. At the aggregate level, our inferred distribution of wedges also substantially reverses the strong negative covariance between productivity and uncertainty (i.e., $cov\left(z_{-1}, \mathbb{V}\right)$) that would have naturally arose due to the complementarity between productivity and information. Consequently, this can lead to losses in aggregate productivity. In the next section,

[17] In previous versions of this paper, we have shown that similar derivations can be made for the benefit side of the equation if we allow for aggregate shocks to profitability or permanent firm characteristics with respect to productivity. While these extensions map more directly to the findings of the empirical literature, they are abstracted from this paper since this is not our main focus.

our goal will be to present the quantitative implications of our findings with respect to firm value and aggregate outcomes.

5 Quantitative Model Implications

The results in this section emphasize the gains to information reallocation. To that end, we compare our baseline model to a counterfactual featuring the same average uncertainty and degree of investment flexibility, but with a different joint distribution of productivity and uncertainty that arises when $\zeta = 0$. Our choice of this counterfactual relates to Section 3, emphasizing our study of information as an implicit factor of production. The premise is that the correlation between the cost of information acquisition and productivity arises solely due to a distortive wedge as parameterized by ζ, and we are quantifying the implications of these wedges for firm value and aggregate outcomes. In doing so, we follow the literature on factor misallocation in assuming that the cost of factor inputs are not correlated with productivity (e.g. Hsieh and Klenow (2009); David and Venkateswaran (2019); Whited and Zhao (2021)).

We interpret the concept of an economy with the *same average degree of investment flexibility and uncertainty* as a recalibrated counterfactual model that replicates the average uncertainty in our baseline model, and the degree to which firms adjust their investment spending relative to plans. The recalibration is done in partial equilibrium using prices from our baseline model to avoid the confounding effects of price changes on our inference.[18] We then consider the impact in general-equilibrium.

5.1 Private Gains to Information Reallocation

Our first analysis focuses on the impact of information allocation on firm value. As we are emphasizing the effect of information allocation across firms of different productivities, we

[18] An alternative is to impose χ at the baseline calibrated value, then recalibrate ξ to only target the average dispersion of TFP shocks. Therefore, "same flexibility" in this case would be interpreted as one where the cost of adjustment stays the same; however, this latter calibration will not be able to simultaneously match the average dispersion of investment plan deviations. Nonetheless, we have experimented with this specification and found that both alternatives predict almost identical results.

define firm value from an ex ante perspective similar to equation 6, where for a firm with a given prior productivity z_{-1}, firm value is computed as

$$\Pi\left(z_{-1}\right) \equiv \mathbb{E}_{z_{-1}} \sum_{t=1}^{\infty} \left(\frac{1}{1+r}\right)^{t} \left(\pi\left(z_t, k_t\right) - \left(r+\delta\right) k_t\right).$$

In other words, we compute the ex-dividend value of a firm using our Bellman equations (equations 4 to 6), subtracting the dollar cost of capital but not information acquisition costs or investment plan adjustment costs, since we do not consider them as direct costs of operations in our baseline. For simplicity, and to emphasize the capital budgeting process, we net out the dividend flow at period 0 (i.e., $\pi_0 - (r+\delta) k_0$) because it does not depend on the capital budgeting decision of the firm at time 0. To operationalize the computation, for a given value of z_{-1}, we simulate a panel of 5,000 firms and track the panel for 20 years. We then compute the discounted net present value of each realized stream of dividend flow and average them across all firms. We then repeat the exercise over a grid of z_{-1} using the exact same draw of TFP and noise shocks, using a grid that covers 2.5 standard deviations of TFP values on both sides of the mean. The improvement in firm value for each state z_{-1} in the baseline relative to the counterfactual is calculated as:

$$\Delta\Pi\left(z_{-1}\right) = 100 \times \left(\frac{\Pi^{counterfactual}\left(z_{-1}\right)}{\Pi^{baseline}\left(z_{-1}\right)} - 1\right).$$

[Figure 4 Here]

Figure 4 shows the improvement in firm value as a function of z_{-1}, broadly increases in z_{-1} following the intuition of our model that high-productivity firms have the most to gain from acquiring better information. We also see that improvements are across the board, ranging from around 2.8% to 4.1% of firm value. Noticeably, even low-productivity firms benefit from such a reallocation of information. For low-productivity firms, a positive correlation between productivity and information acquisition costs ($\zeta > 0$) implies that they enjoyed an implicit subsidy in information acquisition through lower costs; as such, information is in fact being reallocated *away* from them in our counterfactual exercise.

The net gains of reallocation for low-productivity firms arise because firms are forward-looking and productivity is mean-reverting. Even though low-productivity firms are subsidized when $\zeta > 0$, firms understand that they will eventually end up in a high-productivity state in the future and face an implicit tax on information acquisition. Consequently, this inter-temporal consideration means that even low-productivity firms prefer having information reallocated away from them now, in anticipation of having information reallocated towards them in the future when the benefits are substantially larger.

5.2 Aggregate Gains to Information Reallocation

Having discussed the implications of information reallocation at the firm-level, we now turn to quantifying the aggregate gains from information reallocation. To conduct our analysis, we first quantify the total TFP loss arising from uncertainty alone, since this helps us to put in context the maximum possible gains from information reallocation. We consider a model where uncertainty is homogeneous across firms, but the degree of investment flexibility is otherwise the same as our baseline model, and compute TFP relative to a model where firms face no uncertainty about current productivity. Column 1 in Table 5 shows the amount of TFP loss from uncertainty alone is about 0.66%.

Next, we quantify the degree to which information reallocation mitigates these losses. For our baseline (Column 2), we find that TFP loss is about 0.53%, an improvement over the counterfactual where information is homogeneously allocated across firms. This finding follows directly from our Proposition 4. In our baseline model, since the calibrated $\zeta < \bar{\zeta}$, the covariance between productivity and uncertainty is negative, leading to mitigation of TFP losses. Turning to the $\zeta = 0$ counterfactual, we find that TFP loss is about 0.06%. In other words, if firms did not face these implicit wedges to learning, aggregate TFP loss due to uncertainty would have been virtually eliminated via an improvement allocation of information.

Overall, our results suggest that there are substantial gains from reallocating information towards high-productivity firms. However, our baseline calibration implies that few of these gains are realized due to the aforementioned wedges on the cost of information acquisition. However, while our counterfactual exercise demonstrates that substantial gains can be achieved if these wedges are removed, our analyses so far precludes a discussion of the costs of such reallocation as well as potential general-equilibrium effects. We now turn to a brief exploration of the cost and net benefits of information reallocation.

[Table 5 Here]

5.3 The Cost of Reallocation

To provide a simple quantification, we assume that a social planner is able to impose targeted subsidies $\tau(z_{-1})$ at the firm-level such that $\xi z_{-1}^{\zeta}\tau(z_{-1}) = \xi^{*}$, where ξ^{*} is the parameter value in the counterfactual model. In other words, the social planner places additional wedges such that firms perceive that the marginal cost of learning is no longer correlated with productivity, but in a way such that average uncertainty across both economies are preserved. The total cost of such a reallocation is computed as $\int_{z_{-1}} \xi z_{-1}^{\zeta}\left(\sigma^{-2}(z_{-1}; \tau(z_{-1})) - \sigma^{-2}(z_{-1})\right)$, where $\sigma^{-2}(\cdot)$ is the information acquisition policy function under each scenario.

In partial equilibrium, keeping prices fixed, we continue to find net gains to all firms' values if we assume that these costs are distributed out as a lump-sum to firms (Panel a of Figure 5). This result arises because the overall costs are relatively small, reaching a maximum of around 0.75% of pre-reform firm value. Likewise, we find substantial net benefits at an aggregate level, with long-run aggregate output increasing by about 2.6%, whereas the total cost amounts to around 0.15% of pre-reform aggregate output. However, a weakness of this partial-equilibrium approach is that it does not respect the resource constraints in our model, namely that labor supply is inelastic and demand is endogenous. We therefore further explore the impact of the reallocation exercise in general-equilibrium.

In general equilibrium, conducting a similar exercise as in the PE counterfactuals but allowing wages and prices to adjust endogenously, we find that gains are no longer unconditional across the distribution. Instead, Panel b of Figure 5 shows that lower productivity firms appear to be hurt by the reallocation exercise. The difference between partial and general equilibrium counterfactual analyses arises from our assumption that labor supply is inelastic. Specifically, due to the improvement in information allocation, labor demand by higher productivity firms increases (Figure 6) while labor demand by lower productivity firms decreases. However, aggregate labor demand is disproportionally driven by higher productivity firms and total demand increases. This translates purely to a rise in wages in general-equilibrium due to inelastic labor supply, which causes firms to reduce the amount of workers hired. Consequently, firm revenue and profits are relatively lower in general-equilibrium. Notably, while demand for firm output also rises in equilibrium, which in turn increases firm value, the rise in labor cost is sharper than that of demand, decreasing firm profits. In our calibrated model, this in turn implies that the gains to lower productivity firms, which are only realized in the longer term, are outweighed by the permanent rise in labor cost associated with this policy change and they are hurt by the information reallocation.

At the aggregate level, we see a similar effect playing out. While the total cost of the reform still amounts to around 0.15% of pre-reform aggregate output, similar to our partial equilibrium analyses, long-run aggregate output increases by a smaller level of around 0.66%. While the gains are more muted in general equilibrium, these improvements are still substantial in comparison to the cost of the reallocation exercise.

Overall, we view the partial- and general- equilibrium approaches as both valid counterfactual analyses that highlight the importance of labor market effects. In particular, any policymakers considering whether such an information reallocation policy, for example by subsidizing the purchase of information technology to improve internal controls and internal reporting/information gathering, should consider how labor supply may react to the technology's impact on labor demand. On the one hand, if labor supply is able to increase and fully absorb the rise in labor demand such that wages remain unchanged, the net effects will be closer to those in the partial equilibrium analyses. On the other hand, if labor supply cannot accommodate the higher labor demand due to the information reallocation, then gains are muted and low productivity firms may be worse off.

6 Conclusion

In this paper, we develop a theory of capital budgeting, featuring a trade-off between information acquisition for improved capital budgeting against the cost of information acquisition. We theoretically and quantitatively examine the interplay of these two forces, and find that while firms of higher productivity benefit strongly from better information, the cost of information acquisition is also sharply rising in productivity, leading most firms to not internalize these benefits. We show that a theoretical "reallocation" of information from low- to high-productivity firms can generate substantial private and aggregate benefits, although the degree of gains depend on the assumptions of the labor market.

References

Antle, R. and Eppen, G. D. (1985). Capital Rationing and Organizational Slack in Capital Budgeting. *Management Science*, 31(2):163–174.

Arnold, G. C. and Hatzopoulos, P. D. (2000). The theory-practice gap in capital budgeting: Evidence from the United Kingdom. *Journal of Business Finance and Accounting*, 27(5-6):603–626.

Asker, J., Collard-Wexler, A., and De Loecker, J. (2014). Dynamic Inputs and Resource (Mis)Allocation. *Journal of Political Economy*, 122(5):1013–1063.

Bachmann, R., Caballero, R. J., and Engel, E. M. (2013). Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model. *American Economic Journal: Macroeconomics*, 5(4):29–67.

Bachmann, R. and Elstner, S. (2015). Firm Optimism and Pessimism. *European Economic Review*, 79:297–325.

Bachmann, R., Elstner, S., and Hristov, A. (2017). Surprise, surprise - Measuring firm-level investment innovations. *Journal of Economic Dynamics and Control*, 83:107–148.

Baley, I. and Blanco, A. (2019). Firm uncertainty cycles and the propagation of nominal shocks. *American Economic Journal: Macroeconomics*, 11(1):276–337.

Baley, I. and Veldkamp, L. (2021). Bayesian Learning. In *Handbook of Economic Expectations*.

Barrero, J. M. (2021). The Micro and Macro of Managerial Beliefs.

Begenau, J., Farboodi, M., and Veldkamp, L. (2018). Big data in finance and the growth of large firms. *Journal of Monetary Economics*, 97:71–87.

Bernardo, A. E., Cai, H., and Luo, J. (2004a). Capital budgeting in multidivision firms: Information, agency, and incentives. *The Review of Financial Studies*, 17(3):739–767.

Bernardo, A. E., Cai, H., and Luo, J. (2004b). Capital Budgeting in Multidivision Firms: Information, Agency, and Incentives. *Review of Financial Studies*, 17(3):739–767.

Bhandari, A. and McGrattan, E. R. (2020). Sweat Equity in U.S. Private Business. *Quarterly Journal of Economics*.

Bloom, N., Bond, S., and Reenen, J. V. (2007). Uncertainty and Investment Dynamics. *The Review of Economics Studies*, 74(2):391–415.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-eksten, I., and Terry, S. J. (2018). Really Uncertain Business Cycles. *Econometrica*, 86(3):1031–1065.

Bloom, N., Kawakubo, T., Meng, C., Mizen, P., Riley, R., Senga, T., and John Van Reenen (2021). Do Well Managed Firms Make Better Forecasts?

Broer, T., Kohlhas, A., Mitman, K., and Kathrin Schlafmann (2020). Heterogenous Information Choice in General Equilibrium.

Charoenwong, B., Kimura, Y., and Kwan, A. (2020). How and Why Do Managers Use Public Forecasts to Guide the Market ? *Working Paper*.

Christiano, L. J. and Todd, R. M. (1996). Time to plan and aggregate fluctuations. *Federal Reserve Bank of Minneapolis Quarterly Review*.

Clementi, G. L. and Palazzo, B. (2019). Investment and the Cross-Section of Equity Returns. *Journal of Finance*, 74(1):281–321.

Coibion, O. and Gorodnichenko, Y. (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *American Economic Review*, 105(8):2644–2678.

Coibion, O., Gorodnichenko, Y., and Ropele, T. (2020). Inflation Expectations and Firm Decisions: New Caual Evidence. *Quarterly Journal of Economics*, 135(1):165–219.

David, J. M., Hopenhayn, H. A., and Venkateswaran, V. (2016). Information, Misallocation, and Aggregate Productivity. *Quarterly Journal of Economics*, 131(2):943–1005.

David, J. M. and Venkateswaran, V. (2019). The Sources of Capital Misallocation. *American Economic Review*, 109(7):2531–67.

Dyreng, S., Ferracuti, E., and Morris, A. (2021). Renegotiation costs and contract design. In *28th Annual Conference on Financial Economics and Accounting, Paris December 2017 Finance Meeting EUROFIDAI-AFFI.*

Farboodi, M., Veldkamp, L., Singal, D., and Venkateswaran, V. (2021). Valuing Financial Data. *SSRN Electronic Journal*, pages 1–43.

Fuster, A., Perez-Truglia, R., Wiederholt, M., and Zafar, B. (2020). Expectations With Endogenous Information Acquisition: An Experimental Investigation. *The Review of Economics and Statistics (forthcoming).*

Gennaioli, N., Ma, Y., and Shleifer, A. (2015). Expectations and Investment. In *NBER Macroeconomics Annual*, number 30, pages 379–442.

Graham, J. (2022). Presidential Address: Corporate Finance and Reality. *Journal of Finance.*

Harris, M., Kriebel, C. H., and Raviv, A. (1982). Asymmetric Information, Incentives and Intrafirm Resource Allocation. *Management Science*, 28(6):604–620.

Harris, M. and Raviv, A. (1996). The capital budgeting process: Incentives and information. *The Journal of Finance*, 51(4):1139–1174.

Harris, M. and Raviv, A. (1998). Capital budgeting and delegation. *Journal of Financial Economics*, 50(3):259–289.

Hennessy, C. A., Levy, A., and Whited, T. M. (2007). Testing Q theory with financing frictions. *Journal of Financial Economics*, 83(3):691–717.

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

Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(5795):6.

Ilut, C. and Saijo, H. (2020). Learning, Confidence, and Business Cycles. *Journal of Monetary Economics.*

Ilut, C. and Valchev, R. (2020). Economic agents as imperfect problem solvers. *Working Paper.*

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, 50(6).

Maćkowiak, B. and Wiederholt, M. (2015). Real Business Cycle Dynamics under Rational Inattention. *The American Economic Review.*

Malenko, A. (2019). Optimal dynamic capital budgeting. *Review of Economic Studies*, 86(4):1747–1778.

Mao, J. C. T. (1970). Survey of Capital Budgeting: Theory and Practice. *Journal of Finance: Papers and Proceedings*, 25(2):349–360.

Midrigan, V. and Xu, D. Y. (2014). Finance and Misallocation: Evidence from Plant-Level Data. *American Economic Review*, 104(2):422–458.

Myers, S. C. (1974). Interactions of Corporate Financing and Investment Decisions - Implications for Capital Budgeting. *The Journal of Finance*, 29(1):1–25.

Plambeck, E. L. and Taylor, T. A. (2007). Implications of breach remedy and renegotiation design for innovation and capacity. *Management Science*, 53(12):1859–1871.

Reis, R. (2006). Inattentive producers. *Review of Economic Studies*, 73(3):793–821.

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

Ryan, P. a. and Ryan, G. P. (2002). Capital Budgeting Practices of the Fortune 1000: How Have Things Changed? *Journal of Business & Management*, 8(4):355.

Schall, L. D., Sundem, G. L., and Geijsbeek, W. R. (1978). Survey and Analysis of Capital Budgeting Methods. *Journal of Finance*, 33(1):281–287.

Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690.

Tanaka, M., Bloom, N., David, J. M., and Koga, M. (2020). Firm Performance and Macro Forecast Accuracy. *Journal of Monetary Economics*, 18(9):1–38.

Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *Journal of Finance*, 64(3):1187–1215.

Whited, T. M. and Zhao, J. (2021). The Misallocation of Finance. *Journal of Finance*, (2009).

7 Tables and Figures

Table 1: Summary Statistics

The table below shows the summary statistics of our firm-year panel. The total number of firms in our sample is 5,989, of which 2,273 are publicly listed and the rest are private companies. To reduce the influence of outliers on these summary statistics, we winsorize variables at the 1% level. Capital stock is the Net Plants, Property & Equipment. Profits are reported ordinary profits according to Japanese Generally-Accepted Accounting Principles (GAAP). The Wage Bill is the sum of total salary cost for employees and company officers as well as the bonus for employees and company officers. Investment plans are represented as a percentage of end-of-previous-fiscal-year capital stock. Employment is the number of employees is represented as the number of full-time equivalent workers and may include fractions. TFP shocks are defined in the main text with additional details in Appendix Section A.2. All numbers are rounded to three significant digits or three decimal points, whichever results in fewer decimal points.

Panel A: Firm Fundamentals						
					Percentile	
Variable	Mean	SD	Skew	25^{th}	50^{th}	75^{th}
Total Assets (mn)	88,100	205,000	4.81	11,060	26,700	67,300
Capital Stock (mn)	50,000	131,000	5.14	4,590	12,300	33,100
Employment (count, FTE)	1,180	2,100	3.63	162	466	1,180
COGS (mn)	57,100	125,000	4.62	4,950	17,400	49,800
Wage Bill (mn)	1,670	3,120	4.11	229	649	1,650
Sales (mn)	72,700	152,000	4.47	7,810	23,800	66,300
Expected Sales (mn)	75,300	157,000	4.43	8,100	24,700	68,800
Profits (mn)	3,760	9,050	4.56	207	920	3,070
Expected Profits (mn)	3,700	8,650	4.57	247	930	3,000
Total Capital Investment (mn)	3,060	8,500	5.23	121	545	1,990
Total Capital Investment Plan (mn)	3,250	9,160	5.20	100	554	2,030
Panel B: Constructed Variables						
Investment (% of Capital)	7.63	9.55	2.67	1.63	4.58	9.76
Capital Investment Plan (% of Capital)	7.73	9.78	2.67	1.40	4.69	10.1
Investment Plan Deviations (% of Capital)	-0.140	5.55	0.63	-1.64	-0.0610	0.947
log TFP Shock	-0.127	0.480	-1.06	-0.275	-0.043	0.082

Table 2: Testing Model Predictions

In the table below, Columns 1 and 2 report Prediction 1 from Section 3. Column 3 tests Prediction 2 from Section 3. Column 4 reports the β from Equation (13) for the pooled sample. Expectations, actual values, and shocks are winsorized at the 1% level. All regressions include industry and year fixed effects. Standard errors are clustered by firm and shown in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Dependent Variable:	$i_{i,t}/k_{i,t}$		$\log(1+\lvert u_{i,t}\rvert)$	$\log(1+\lvert\tilde{\epsilon}_{i,t}\rvert)$
	(1)	(2)	(3)	(4)
$\log z^e_{i,t}$	1.759***	0.489***		
	(0.170)	(0.126)		
$\frac{i^p_{i,t}}{k_{i,t}}$		0.638***		
		(0.048)		
$\log(1+\lvert\tilde{\epsilon}_{i,t}\rvert)$			0.013***	
			(0.003)	
$\log z_{i,t-1}$				-0.009***
				(0.002)
Observations	30,441	30,441	30,441	30,441
R^2	0.055	0.410	0.031	0.240

Table 3: Parameter estimates

The table below reports our parameter estimates. Panel A reports the parameter values estimated directly off the data, while Panel B reports calibrated parameters. All values are rounded to 3 decimal placeswhere appropriate.

Parameters	$\eta = 4$	$\eta = 3$	$\eta = 6$
	(1)	(2)	(3)
	Panel A: Estimated parameters		
α	0.353	0.273	0.418
ρ	0.947	0.954	0.941
σ	0.263	0.260	0.268
	Panel B: Calibrated parameters		
χ	0.1110	0.0803	0.1431
ξ	2.4867	1.5656	4.3761
ζ	2.1543×10^5	1.4381×10^5	0.3898×10^5

Table 4: Model Fit

The tables below reports our model fit for our baseline model and counterfactual model with homogenous learning cost. We report results for both our benchmark calibration assuming $\eta = 4$, as well as our alternative assumptions with $\eta = 3$ and $\eta = 6$. In panel A, we reported model fit per our targeted moments. Note that we also clear the market for labor and output down to a tolerance of 10^{-8}. In panel B of each table, we report model fit per untargeted moments, namely, our predictions from the earlier section. All results are rounded to 3 decimal places or 3 significant figures.

Panel A: Targeted Moments				Panel B: Untargeted Moments			
Moments	Data	Baseline	Homog. Cost	Moments	Data	Baseline	Homog. Cost
				$\eta = 4$			
$\frac{\sigma_{\hat{z}-z}}{\sigma_\epsilon}$	1.024	1.024	1.024	Prediction 1	0.279	0.426	0.353
$\sigma\left(\frac{k'}{k^p}\right)$	0.073	0.073	0.073	Prediction 2	0.013	0.231	0.240
$\rho\left(z_{-1}, \lvert\tilde{\epsilon}\rvert\right)$	-0.0087	-0.0087	—	$\rho\left(z_{-1}, \lvert\tilde{\epsilon}\rvert\right)$	-0.0087	—	-0.065
$(N, P) = (1, 1)$	—	$(1, 1)$	$(1, 1)$				
				$\eta = 3$			
$\frac{\sigma_{\hat{z}-z}}{\sigma_\epsilon}$	1.006	1.006	1.006	Prediction 1	0.249	0.324	0.254
$\sigma\left(\frac{k'}{k^p}\right)$	0.073	0.073	0.073	Prediction 2	0.012	0.243	0.334
$\rho\left(z_{-1}, \lvert\tilde{\epsilon}\rvert\right)$	-0.010	-0.010	—	$\rho\left(z_{-1}, \lvert\tilde{\epsilon}\rvert\right)$	-0.010	—	-0.089
$(N, P) = (1, 1)$	—	$(1, 1)$	$(1, 1)$				
				$\eta = 6$			
$\frac{\sigma_{\hat{z}-z}}{\sigma_\epsilon}$	1.038	1.038	1.038	Prediction 1	0.303	0.218	0.154
$\sigma\left(\frac{k'}{k^p}\right)$	0.073	0.073	0.073	Prediction 2	0.014	0.220	0.232
$\rho\left(z_{-1}, \lvert\tilde{\epsilon}\rvert\right)$	-0.0026	-0.0025	—	$\rho\left(z_{-1}, \lvert\tilde{\epsilon}\rvert\right)$	-0.0026	—	-0.118
$(N, P) = (1, 1)$	—	$(1, 1)$	$(1, 1)$				

Table 5: Capital Budgeting and Capital Misallocation

The table below reports aggregate TFP relative to a model where firms do not face any uncertainty regarding their contemporaneous productivity. The measures are reported as percentage *decreases* relative to the reference. All results are rounded to 3 decimal places or 3 significant figures.

	Fixed $\mathbb{V}$, $\chi < \infty$	Baseline	Counterfactual
	(1)	(2)	(3)
$\eta = 4$	0.657	0.535	0.059

Figure 1: Model Timing

This figure illustrates the three sub-periods of our model.

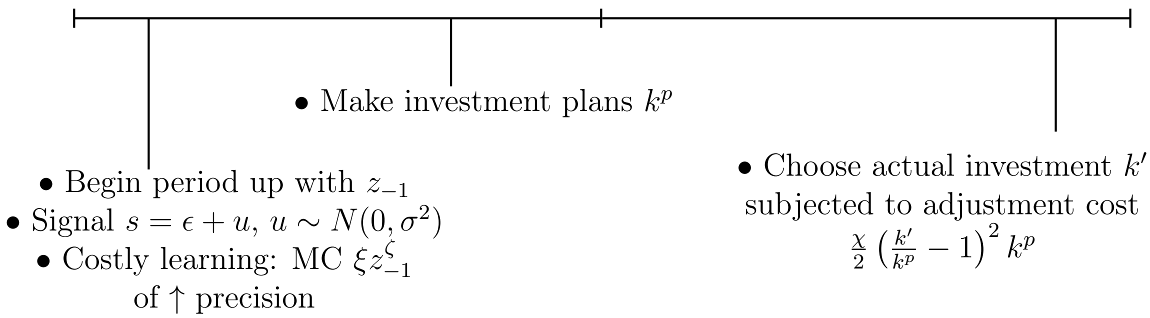


Figure 2: Marginal Benefit and Marginal Cost of Improving Signal Precision

The figures below plot illustrative graphs of how the marginal benefit and cost of learning vary with firm productivity. Graphs are plotted against $\mathbb{V}^{-1}$ (i.e., increasing signal precision). Panel a plots two representive solutions when marginal costs are constant (i.e., $\zeta = 0$). Panel b plots two representive solutions when the net benefit is constant in productivity (i.e., $\zeta = \bar{\zeta}$). Vertical dash lines indicate where the marginal cost meets marginal benefit for each value of productivity.

(a) Constant marginal cost ($\zeta = 0$). (b) Constant net benefit ($\zeta = \bar{\zeta}$).

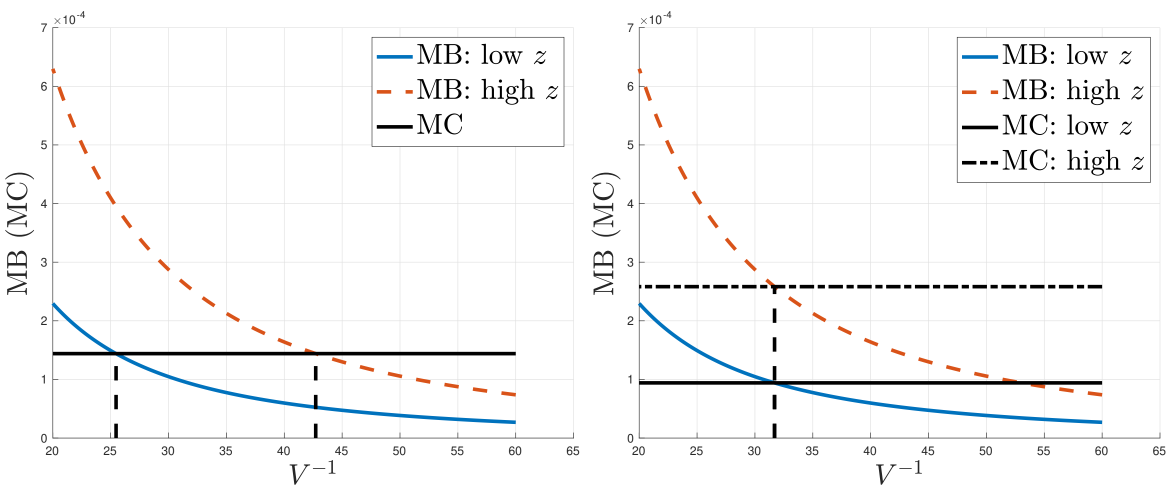
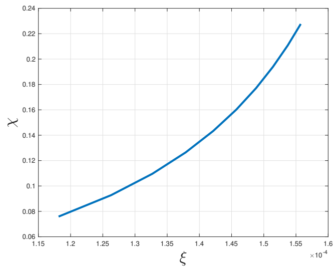


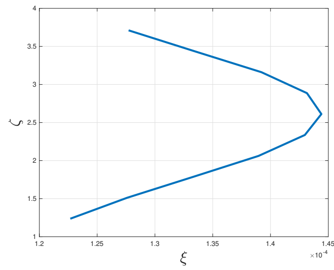
Figure 3: Isomoment Plots

The figures below plot isomoment plots for each possible combination of parameter pairs. The first row reports isomoment plots for the average dispersion of TFP shocks, the second row reports isomoment plots for the average dispersion of investment plan deviations, and the last row plots the correlation of lagged TFP with dispersion of TPF shocks. All results are computed using our targeted moments as per Panel A of Table 4.

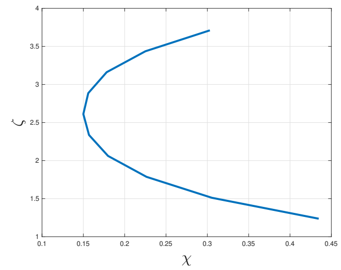
(a) Moment 1: ξ vs χ



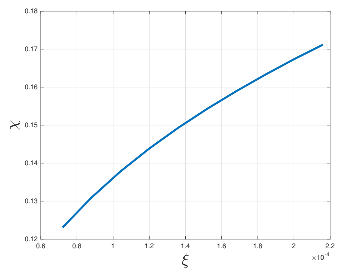
(b) Moment 1: ξ vs ζ



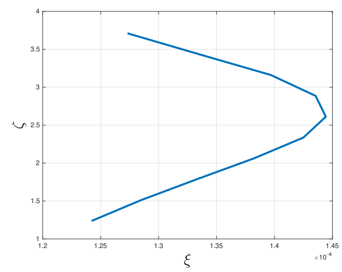
(c) Moment 1: χ vs ζ



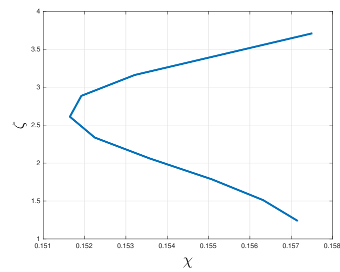
(d) Moment 2: ξ vs χ



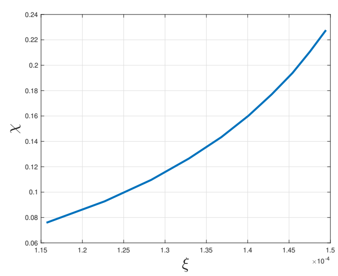
(e) Moment 2: ξ vs ζ



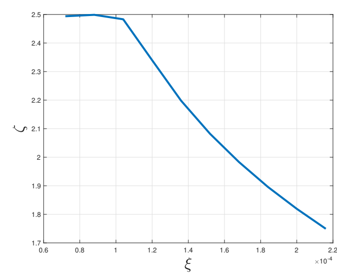
(f) Moment 2: χ vs ζ



(g) Moment 3: ξ vs χ



(h) Moment 3: ξ vs ζ



(i) Moment 3: χ vs ζ

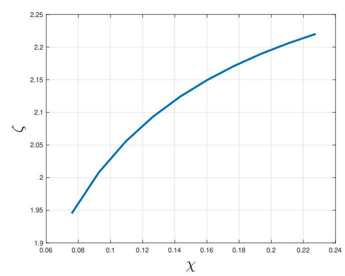


Figure 4: Gains to Firm Value Post-Reallcation

The figures below plots the gains in net present firm value after removal of the information acquisition wedges.

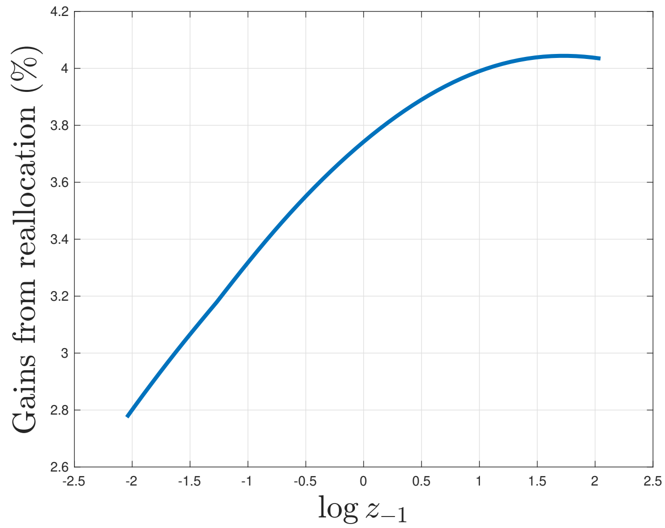
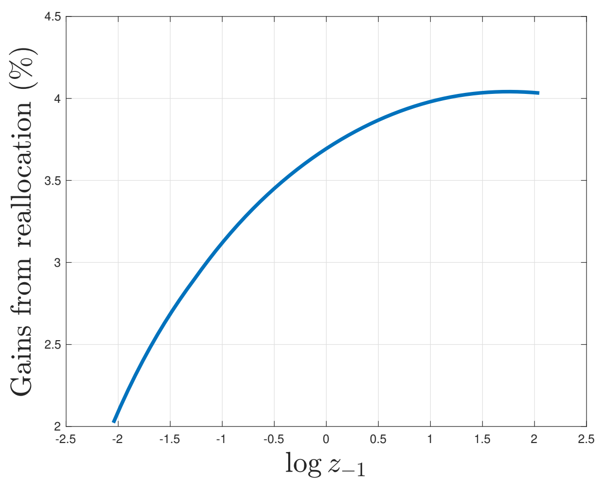


Figure 5: Gains to Firm Value After Accounting for Costs

The figures below plot gains to net present firm value relative to the baseline model, computed in partial-equilibrium (Panel a) and general-equilibrium (Panel b). The net present value includes the life time cost of pay for the policy.

(a) PE



(b) GE

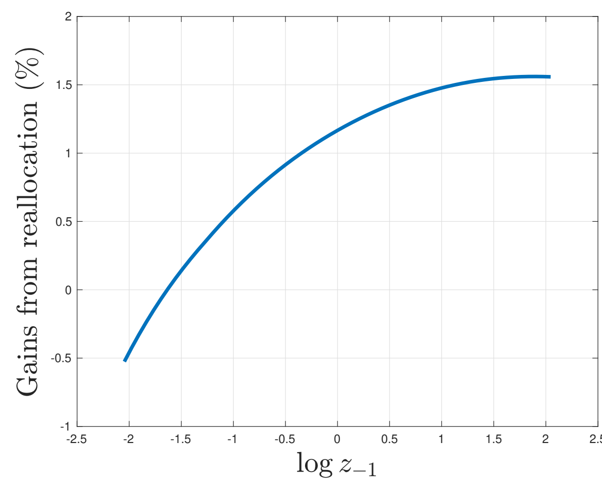
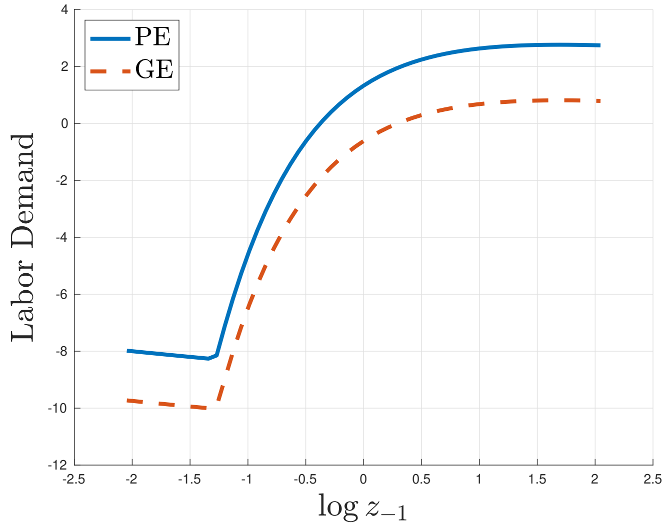


Figure 6: Changes In Labor Demand

The figures below plot the change in labor demand relative to the baseline model. A negative number implies that labor demand decreases in the counterfactual. Blue solid line: PE. Red dashed line: GE.



Online Appendix

A Empirical Methodology and Data Details

A.1 How Good are Forecasts and Plans?

We verify that the forecasts and plans reported by a firm do in fact predict its realized counterparts. In other words, these forecasts and plans are relevant. To do so, we use regressions of the form:

$$y_{i,t} = \alpha_{j(i),t} + \alpha_i + \beta \mathbb{E}\left[y_{i,t}\right] + \varepsilon_{i,t} \tag{14}$$

where the i subscript indexes a firm, $j(i)$ indexes the MoF industry of the firm, and t indexes a fiscal year. The variables $\mathbb{E}\left[y_{i,t}\right]$ is the forecast or plan of the outcome variable $y_{i,t}$ made from the first quarter of the same fiscal year, α_i denotes firm fixed effects, and $\alpha_{j(i),t}$ denotes industry-by-year fixed effects. Standard errors are clustered by firm. The outcome variables considered are realized capital investment, land investments, software expenses, profits, and sales. As discussed, for investment, land investment, and software expenses, the expectation variable is interpreted as a spending plan; while for profits and sales, the expectation variable is interpreted as a forecast.

The empirical specification in Equation 14 controls for industry and macroeconomic shocks and compares firms with higher expected spending plans with those in the same industry and year with lower investment plans. The coefficient of interest is β – capturing the importance of plans on actual realizations across different fiscal years. $\beta = 1$ corresponds to perfectly sticky plans which do not permit ex post adjustments while $\beta = 0$ corresponds to completely uninformative plans for actual spending. Finally, in addition to reporting the point estimate, we also consider the statistical importance represented by the relative improvement of the R^2 in the regressions compared against a regression specification without the expectation variable.

Our empirical results suggest the corporate plans and forecasts contained in the BOS are good predictors of actual realizations, and hence economically relevant. Columns 1 through 3 of Panel A in Table A.1 study spending plans for physical capital investment, land investment, and software spending, while Columns 4 and 5 evaluate profits and sales forecasts. For realized physical capital investment, plans have an estimated coefficient of 0.666 and increase the R^2 by 28% relative to a model with only firm and industry-by-quarter fixed effects. For land investment, plans have an estimated coefficient of 0.904 and increase the R^2 by 32%. For software expenses, plans have an estimated coefficient of 0.494 and increase the R^2 by only 7%. These results for both the point estimates and relative statistical fit improvements are consistent with an intuitive ranking of how costly it is to adjust your investment plans: changing a firm's land purchasing plan incurs the highest cost, followed by physical capital investment, and then by software spending. Notably, spendings for costlier deviation from plans will carry a coefficient closer to one and account for a larger statistical variation of actual realized investment.

Table A.1: Spending Plans and Forecasts

The table below reports the results of estimating equation 14. Expectations and actual values are winsorized at the 1% level. All regressions include firm and industry-by-year fixed effects. Standard errors are clustered by firm and shown in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01.

Dependent Variable:	y_t				
y_t =	PP&E	Land Purchases	Software Spending	Profits	Sales
	(1)	(2)	(3)	(4)	(5)
$\mathbb{E}[y_t]$	0.666***	0.904***	0.494***	0.759***	0.915***
	(0.015)	(0.0401)	(0.028)	(0.030)	(0.012)
Observations	26,718	26,588	26,412	26,718	26,718
R^2	0.739	0.442	0.839	0.885	0.995
R^2 without Expected Value	0.578	0.335	0.787	0.802	0.940
Relative % Increase	28%	32%	7%	10%	36%

A.2 Measurement Strategy and Issues

In this section, we detail our measurement strategy for the variables described in the main text, and address some potential issues that arise with our measurement strategy.

A.2.1 Imputing value-added

The basic assumption in our model is that we have a physical production function given by $y = zk^{\alpha}l^{1-\alpha}$. Under assumptions of constant elasticity of substitution (CES) across goods and monopolistic competition, we get the "sales" production function $py = z^{\theta}k^{\alpha\theta}l^{(1-\alpha)\theta}$, where $\theta \equiv \frac{\eta-1}{\eta}$. An issue arises for our measurement strategy because although this is conceptually a "value-added" production function, we do not observe value-added in the data—only total sales. As such, even if we knew the exact parameters of the production function, we are not able to directly back out TFP using our data.

Therefore, to address this issue, we impute value-added by constructing it using total sales scaled by one minus the fraction of industry-level material costs to total sales based on aggregate statistics. This is common practice in the literature when direct measures of firm intermediate costs are not available (e.g., David and Venkateswaran (2019)). Our industry-level material share estimates are obtained from the Japanese Ministry of Economy, Trade, and Industry (METI), which we report in Column 1 of Table A.2.

Table A.2: TFP and TFP Shocks by Industry

The table below reports our estimates using our imputed materials share approach, for value-added, capital cost share, TFP parameters, and the dispersion of TFP shocks relative to the volatility of TFP, with the assumption that $\eta = 4$. All estimates are computed at the industry level, except for the overall results which are pooled regressions. Value-added to total sales ratios (i.e., one minus the imputed materials share) come from Ministry of Economics, Trade, and Industry. The adjusted ratio uses the formula in Equation 18, assuming that the variance of $\mathbb{V}$ is approximately 0. All results are rounded to 3 decimal places or 2 significant figures, where appropriate.

Industry	Value-added, As Fraction of Total Sales (1)	Capital Cost Share ($\theta\alpha$) (2)	AR(1) Coef (ρ) (3)	SD AR(1) Resid. Error (σ_ϵ) (4)	SD TFP Shock ($\sigma_{\tilde{\epsilon}}$) (5)	$\frac{\sigma_{\tilde{\epsilon}}}{\sigma_\epsilon}$ (6)	Adjusted ratio (7)
Agriculture	0.497	0.409	0.903*** (0.028)	0.176	0.213	1.208	0.621
Mining and Construction	0.425	0.491	0.961*** (0.007)	0.191	0.218	1.141	0.586
Manufacturing	0.304	0.264	0.942*** (0.008)	0.218	0.346	1.583	0.814
Metals and Machinery	0.324	0.238	0.920*** (0.006)	0.249	0.384	1.541	0.792
Retail and Wholesale Trade	0.690	0.635	0.989*** (0.003)	0.181	0.132	0.728	0.374
Eateries and Real Estate	0.709	0.595	0.953*** (0.008)	0.276	0.162	0.586	0.301
Transportation	0.544	0.383	0.972*** (0.004)	0.193	0.149	0.770	0.396
Utilities and Services	0.489	0.415	0.963*** (0.007)	0.287	0.200	0.697	0.368
Medical, School, and Misc. Services	0.609	0.305	0.901*** (0.014)	0.422	0.283	0.671	0.345
Overall	0.483	0.398	0.947*** (0.003)	0.263	0.270	1.024	0.526

A.2.2 Estimating α

Similar to Asker et al. (2014), we assume that capital is quasi-fixed (like in our model) but labor is free to adjust every period after productivity z has been observed. Let profits net of labor cost be $py - wl$. Then the optimality condition for labor is

$$MRPL \equiv (1-\alpha)\,\theta z^{\theta} k^{\alpha\theta} l^{(1-\alpha)\theta} = w$$
$$\implies \frac{wl}{py} = (1-\alpha)\,\theta.$$

In other words, we can identify $(1-\alpha)\,\theta$ by simply computing the labor share in value-added. We follow Asker et al. (2014) by computing the industry median from the individual firm-year level labor cost shares:

$$\widehat{(1-\alpha)}\,\theta = median\left\{\frac{wl}{py}\right\}.$$

Finally, α can be backed out using a given assumed value of η, keeping in mind the identity that $\theta = \frac{\eta-1}{\eta}$. As our benchmark, we assume $\eta = 4$, following the literature (e.g., Asker et al. (2014); Bloom et al. (2018)). However, we also consider values of $\eta = 3$ and $\eta = 6$ as a robustness check. We report our estimates for our benchmark assumption (in the form of capital cost shares) in Column 2 of Table A.2.

A.2.3 Estimating TFP, Expected TFP and TFP Shocks

With θ and α in hand, we compute realized TFP (in levels) as

$$z = \left(\frac{py}{k^{\alpha\theta} l^{(1-\alpha)\theta}}\right)^{\frac{1}{\theta}},$$

where py is value-added for the fiscal year, used in computing α. In computing the statistical moments associated with z, we residualize $\log z$ using year (and industry, where relevant) fixed effects, and fit AR(1) regressions to $\log z$. The autocorrelation and volatility of TFP, used in the model calibration, are reported in Columns 3 and 4 of Table A.2.

To compute expected TFP z^e (in levels), we assume that

$$z^e = \left(\frac{py^e}{k^{\alpha\theta} l^{(1-\alpha)\theta}} \right)^{\frac{1}{\theta}},$$

where py^e is forecasted value-added for the fiscal year. Finally, we define TFP "shocks" as

$$\tilde{\epsilon} \equiv \log z - \log z^e,$$

that is, the difference between forecasted TFP and actual TFP. The dispersion of TFP shocks is reported in Column 5 of Table A.2, while in Column 6, we report the dispersion of TFP shocks relative to the volatility of TFP.

It is clear that our measured z^e is only an approximation of the "true" expected TFP. To be precise, our measurement strategy gives a value of expected TFP that can be expressed as

$$z^e = \left(\frac{\mathbb{E}\left[z^{\theta + \frac{\theta^2(1-\alpha)}{1-(1-\alpha)\theta}} \right]}{z^{\frac{\theta^2(1-\alpha)}{1-(1-\alpha)\theta}}} \right)^{\frac{1}{\theta}},$$

which is not $\mathbb{E}[z]$. This measurement is, in large part, why the dispersion of TFP shocks is generally larger than the volatility of TFP. However, in the next subsection, we proceed to explain why this mismeasurement is not an area of concern for us.

A.2.4 Mismeasured Expected TFP: Why It Happens and Why It Does Not Matter For Our Results

We structure our discussion here in three stages. First, we show that under certain circumstances, measured expected TFP will always exhibit a negative bias, but the bias is small and irrelevant. Second, we show why, due to our data limitations, the bias in z^e can be large, and cannot be signed in general. In particular, we formally derive an expression to quantify this (see equation 15 in this appendix). Finally, we show that the bias in $\tilde{\epsilon}$ (i.e., TFP shocks, our object of interest) can always be signed. Specifically, the bias is only in one direction: our mismeasured TFP shocks is always smaller (or more negative) than the true TFP shock. Importantly, because our focus is on studying $\tilde{\epsilon}$, not z^e, this means that our qualitative results are not affected by the mismeasurement in z^e.

Best case scenario: Bias in z^e is small and can be signed We begin by emphasizing here that it is impossible to directly estimate an unbiased measure of expected TFP using just balance sheet data alone, even if we observe all possible expectations of these balance sheet variables. For instance, in our setup, suppose we observed expected labor or $\alpha = 1$. In both cases, the expression above is reduced to

$$z^{e^*} = \left(\mathbb{E}\left[z^\theta\right]\right)^{\frac{1}{\theta}},$$

where we denote z^{e^*} as the expected TFP one would back out under the assumptions above. Due to Jensen's inequality, $z^{e^*} < \mathbb{E}[z]$ since $\theta < 1$. That is to say, the mismeasured expected TFP is always smaller than the true expected TFP. For example, suppose we assume that the firm's expectations follow our model, then

$$\begin{aligned} z^{e^*} &= \left[z_{-1}^{\rho\theta} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2}\theta s\right) \exp\left(\frac{1}{2}\theta^2 \mathbb{V}\right)\right]^{\frac{1}{\theta}} \\ &= z_{-1}^{\rho} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s\right) \exp\left(\frac{1}{2}\theta \mathbb{V}\right), \end{aligned}$$

where s is the private signal as observed by the firms in our model, and σ (and corresponding $\mathbb{V}$) is the endogenous choice of uncertainty. The unbiased expected value of z is

$$\mathbb{E}[z] = z_{-1}^{\rho} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s\right) \exp\left(\frac{1}{2}\mathbb{V}\right).$$

This gives us a biased estimated of expected TFP, specifically,

$$z^{e^*} - \mathbb{E}\left[z\right] = z_{-1}^{\rho} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s\right) \left[\exp\left(\frac{1}{2}\theta\mathbb{V}\right) - \exp\left(\frac{1}{2}\mathbb{V}\right)\right] < 0,$$

where the bias shows up because of the terms in the square brackets. However, because the volatility of TFP (i.e, σ_ϵ) is typically small, and $\mathbb{V} < \sigma_\epsilon^2$, the bias will be relatively small. In other words, $z^{e^*} \approx \mathbb{E}\left[z\right]$.

The generic case: Bias in z^e is large and cannot be signed In our case, we do not observe expected labor, and α is clearly less than unity. However, if we follow our model's assumptions (as in Section 2), we can make further headway into understanding the source of the bias, by expressing expected TFP as

$$z^e = \left(\frac{z_{-1}^{\rho\theta + \rho\hat{\theta}} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2}\left(\theta + \hat{\theta}\right) s\right) \exp\left(\frac{1}{2}\left(\theta + \hat{\theta}\right)^2 \mathbb{V}\right)}{z^{\hat{\theta}}}\right)^{\frac{1}{\theta}},$$

where $\hat{\theta} \equiv \frac{\theta^2(1-\alpha)}{1-(1-\alpha)\theta}$. The above expression can be further reduced to

$$\begin{aligned}
z^e &= z_{-1}^{\rho} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s\right) \left(\frac{\exp\left(\frac{1}{2}\left(\theta + \hat{\theta}\right)^2 \mathbb{V}\right)}{\exp\left(\hat{\theta}\epsilon\right)}\right)^{\frac{1}{\theta}} \\
&= z^{e^*} \left(\frac{\exp\left(\left(\theta\hat{\theta} + \frac{1}{2}\hat{\theta}^2\right)\mathbb{V}\right)}{\exp\left(\hat{\theta}\epsilon\right)}\right) \\
\Leftrightarrow \log z^e &= \log z^{e^*} + \left(\theta\hat{\theta} + \frac{1}{2}\hat{\theta}^2\right)\mathbb{V} - \hat{\theta}\epsilon \qquad (15)
\end{aligned}$$

where we utilize the fact that $z = z_{-1}^{\rho} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s\right) \exp\left(\epsilon\right)$ with $\epsilon \sim N\left(0, \mathbb{V}\right)$ being the mean 0 innovations arising under the posterior distribution (i.e., true TFP shocks), per our model assumptions; and we further substituted in the definition for z^{e^*}. Notice that even after taking z^{e^*} as our reference, there is no clear direction for the bias in expected TFP, which depends on the exact innovation the firm receives. Notably, the implication here is that $\tilde{\epsilon}$ introduces attenuation bias into our regression framework. That said, as we reported in Table 2, we find that investment is positively and significantly correlated with $\log z^e$ despite the attenuation bias, suggesting that the bias might not be that severe.

Large unsigned bias in z^e does not matter for our results We now turn to relating mismeasurement in z^e to mismeasurement in the forecast errors. Specifically, we can derive a bias for forecast errors as

$$\begin{aligned}
\frac{z^e}{z} &= \left(\frac{z_{-1}^{\rho\theta+\rho\hat{\theta}} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}\left(\theta+\hat{\theta}\right)s\right)\exp\left(\frac{1}{2}\left(\theta+\hat{\theta}\right)^2\mathbb{V}\right)}{z^{\hat{\theta}+\theta}}\right)^{\frac{1}{\theta}} \\
&= \left(\frac{z_{-1}^{\rho\theta+\rho\hat{\theta}} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}\left(\theta+\hat{\theta}\right)s\right)\exp\left(\frac{1}{2}\left(\theta+\hat{\theta}\right)^2\mathbb{V}\right)}{z_{-1}^{\rho\left(\hat{\theta}+\theta\right)} \exp\left(\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}\left(\theta+\hat{\theta}\right)s\right)\exp\left(\left(\theta+\hat{\theta}\right)\epsilon\right)}\right)^{\frac{1}{\theta}} \\
&= \left(\frac{\exp\left(\frac{1}{2}\left(\theta+\hat{\theta}\right)^2\mathbb{V}\right)}{\exp\left(\left(\theta+\hat{\theta}\right)\epsilon\right)}\right)^{\frac{1}{\theta}} \\
\Leftrightarrow \tilde{\epsilon} \equiv \log z - \log z^e &= \frac{1}{\theta}\left(\left(\theta+\hat{\theta}\right)\epsilon - \frac{1}{2}\left(\theta+\hat{\theta}\right)^2\mathbb{V}\right) \\
&= \frac{\theta+\hat{\theta}}{\theta}\left(\epsilon - \frac{1}{2}\left(\theta+\hat{\theta}\right)\mathbb{V}\right). \qquad (16)
\end{aligned}$$

Note that the corresponding unbiased measure of TFP shocks is simply ϵ.

It is clear now why neither our main empirical results in Section 4 nor that in our calibration in Section 5 are affected by the bias in expected values — the bias in $\tilde{\epsilon}$ is constant (it is shifted by $\frac{1}{2}\left(\theta+\hat{\theta}\right)\mathbb{V}$). We further relate this mismeasurement to our empirical results and calibration strategy as follows.

Relationship to Section 4 As an example, consider two firms with $\mathbb{V}$, one that receives an innovation of $\epsilon = 0$ and another with an innovation of $\epsilon = \frac{1}{2}\left(\theta + \hat{\theta}\right)^2 \mathbb{V}$. Our shocks measure is biased: For the first firm, we record it as having a negative shock, whereas for the second firm, it has a shock of zero. However, since our regressions use cross-sectional variations in identifying the effect of shocks to investment deviations (or investment itself), relative to the first firm, the second firm still receives a "larger" shock, so our regressions are still consistent. To be precise, suppose we fit a regression of the form

$$y = \alpha + \beta\tilde{\epsilon} + u,$$

where u is the usual error term, and y is either $\frac{\Delta i}{k}$ or $\frac{i}{k}$. Our interest is in using ϵ as a regressor, which we do not observe, and is thus replaced with $\tilde{\epsilon}$ as in our empirical strategy. Then this gives us,

$$\begin{aligned}
\hat{\beta} &= \frac{cov\left(\tilde{\epsilon}, y\right)}{var\left(\tilde{\epsilon}\right)} \\
&= \frac{cov\left(\frac{1}{\theta}\left(\left(\theta + \hat{\theta}\right)\epsilon - \frac{1}{2}\left(\theta + \hat{\theta}\right)^2 \mathbb{V}\right), y\right)}{var\left(\frac{1}{\theta}\left(\left(\theta + \hat{\theta}\right)\epsilon - \frac{1}{2}\left(\theta + \hat{\theta}\right)^2 \mathbb{V}\right)\right)} \\
&= \frac{\theta}{\theta + \hat{\theta}} \frac{var\left(\epsilon\right)}{var\left(\epsilon\right) + var\left(\frac{1}{2}\left(\theta + \hat{\theta}\right)\mathbb{V}\right) - 2cov\left(\epsilon, \frac{1}{2}\left(\theta + \hat{\theta}\right)^2 \mathbb{V}\right)} \left[\beta - \frac{cov\left(\frac{1}{2}\left(\theta + \hat{\theta}\right)\mathbb{V}, y\right)}{var\left(\epsilon\right)}\right].
\end{aligned}$$

Note that $\epsilon \sim N\left(0, \mathbb{V}\right)$; therefore, $\mathbb{E}\left[\epsilon \mid \mathbb{V}\right] = \mathbb{E}\left[\epsilon\right]$, implying $cov\left(\epsilon, \mathbb{V}\right) = 0$. Therefore, the expression above reduces to

$$\hat{\beta} = \frac{\theta}{\theta + \hat{\theta}} \frac{var\left(\epsilon\right)}{var\left(\epsilon\right) + var\left(\frac{1}{2}\left(\theta + \hat{\theta}\right)\mathbb{V}\right)} \left[\beta - \frac{cov\left(\frac{1}{2}\left(\theta + \hat{\theta}\right)\mathbb{V}, y\right)}{var\left(\epsilon\right)}\right]. \tag{17}$$

If we assume that $\mathbb{V}$ is homogeneous across firms, as is typically done in the literature (for example, Bloom et al. (2018); Tanaka et al. (2020)), equation 17 reduces to

$$\hat{\beta} = \frac{\theta}{\theta + \hat{\theta}} \beta.$$

In other words, our estimated $\hat{\beta}$ will be smaller than the true unbiased elasticity of investment with respect to TFP shocks β; however, the qualitative correlation will always remain the same (i.e., investment is positively correlated with realized shocks).

However, as we show in our model, $\mathbb{V}$ is heterogeneous across firms. That said, if y is investment deviations $(\frac{\Delta i}{k})$, $cov\left(\mathbb{V}, \frac{\Delta i}{k}\right) \approx 0$. This is because, through the lens of our model, firms with higher uncertainty make larger ex post mistakes in both directions. In other words, while the absolute size of investment deviations are increasing in $\mathbb{V}$, it is not correlated with $\mathbb{V}$. Therefore, the mismeasurement is not an issue for us when studying investment deviations.

Relationship to Section 5 In the case of our calibration, a similar logic follows. For simplicity, again first assume that all firms have the same $\mathbb{V}$. Our calibration strategy then depends solely on the dispersion of TFP shocks being proportional to the posterior variance. Specifically, if we could observe "correct" TFP shocks, then our identification strategy would simply be to map $var\left(\tilde{\epsilon}\right)$ to our model parameters (i.e., we directly observe the posterior variance). For our shocks measure, we observe, assuming that firms have the same $\mathbb{V}$, $var\left(\tilde{\epsilon}\right) = \left(\frac{\theta+\hat{\theta}}{\theta}\right)^2 var\left(\epsilon\right)$ — but this is simply a scaled measure of the true posterior variance. As such, our indirect inference strategy will remain consistent in estimating the true amount of posterior variance.

However, as we note, $\mathbb{V}$ is heterogenous across firms, and so we cannot directly invert out

$var\left(\epsilon\right)$. Specifically, as we already derived,

$$var\left(\tilde{\epsilon}\right)=\left(\frac{\theta+\hat{\theta}}{\theta}\right)^{2}\left[var\left(\epsilon\right)+var\left(\frac{1}{2}\left(\theta+\hat{\theta}\right)\mathbb{V}\right)\right]. \tag{18}$$

As such, in our model calibration, we choose $var\left(\tilde{\epsilon}\right)$ as a target for calibration, rather than simply invert $var\left(\epsilon\right)$ directly from the data.

A.2.5 Summary statistics of investment and investment plans at the industry level

Table A.3: Investment, Investment Plans, and Investment Plan Deviations by Industry

The table below reports our statistics for investment, investment plans, and deviation from investment plans as defined in the main text. All estimates are computed at the industry-by-industry level, except for the overall results which are pooled regressions. All results are rounded to 3 decimal places or 2 significant figures, where appropriate.

	i^p/k		$\log\frac{k'}{k^p}$	
Industry	Mean	std. dev.	Mean	std. dev.
	(1)	(2)	(3)	(4)
Agriculture Forestry	10.976	11.639	1.025	0.104
Textfiles, Food, Paper Products	5.745	6.763	1.009	0.103
Manufacturing	9.842	10.086	1.002	0.113
Manufacturing - Electronic, Computer Equipment	9.687	10.488	1.012	0.302
Transportation	4.615	6.177	1.014	0.101
Misc Retail and Real Estate	4.641	7.443	1.014	0.189
Non-retore retailers and finance	8.772	11.955	1.039	0.386
Utilities and Services	8.843	13.65	1.063	0.523
Medical, School, and Miscellaneous Services	5.210	10.889	1.027	0.198
Overall	7.612	9.931	1.02	0.273

A.2.6 Issues with the investment variable

We briefly discuss the issue with our investment variable here. To be precise, there are three cases when this happens, summarized in Table A.4 below.

	Case I	Case II	Case III
True i and i^p	$i>0$ and $i^p<0$	$i<0$ and $i^p<0$	$i<0$ and $i^p>0$
Observed i and i^p	$i>0$ and $i^p=0$	$i=0$ and $i^p=0$	$i=0$ and $i^p>0$

Table A.4: Three cases where ΔI is mismeasured.

We account for this issue using three strategies. First, the crudest of the three, we simply drop any observation for which i or i^p is reported as zero.[19] We find that our main results are robust to this data treatment. Second, we use the fact that we observe actual capital expenses in the Annual Financial Statistics survey to impute actual (dis)investment done by the BOS firms. This would in theory address the bias generated in Case III. However, the capital expenses in the Financial Statements do not line up perfectly with the BOS due to discrepancies in accounting treatment in the two surveys. Moreover, this approach does not address Case I or II. As such, we consider this imputation method only as a robustness check. We do find that our results are robust to this alternative source of capital expenses data. Third, when calibrating our model, we use an indirect inference approach similar to how we address the bias in estimating expected TFP. This method addresses all three cases.

A.2.7 Summary statistics for correlation of lagged TFP with TFP shocks at the industry level

Table A.5: TFP Shocks and Initial Firm Productivity

The table below reports the regression coefficient β from the regression specification as described in Equation 13 at the industry-by-industry level, as well as the pooled sample. Expectations, actual values, and shocks are winsorized at the 1% level. Standard errors are clustered by firm and shown in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Industry	Coefficient
Agriculture	-0.031 (0.019)
Mining and Construction	0.007 (0.005)
Manufacturing	-0.021 (0.007)***
Metals and Machinery	-0.040 (0.005)***
Retail and Wholesale Trade	-0.005 (0.002)**
Eateries and Real Estate	0.010 (0.004)***
Transportation	0.007 (0.003)**
Utilities and Services	-0.008 (0.003)**
Medical, School, and Miscellaneous Services	-0.019 (0.005)***
Overall	-0.009 (0.002)***

[19] This strategy borrows from Bachmann et al. (2017) with details from footnote 12 of their paper.

A.3 Additional Model-Implied Figures and Tables

We report here the additional figures referenced in the main text, namely the distribution of firm uncertainty and associated descriptive statistics. Figure A.1 below plots the distribution of firm uncertainty, and the table below reports broad descriptive statistics associated with this distribution.

Figure A.1: The distribution of signal precision.

The figure below plots the distribution of $\frac{\mathbb{V}}{\sigma^2}$ for the baseline model. The solid red line is a reference for the signal precision under the counterfactual assumption that the distribution of information is degenerate.

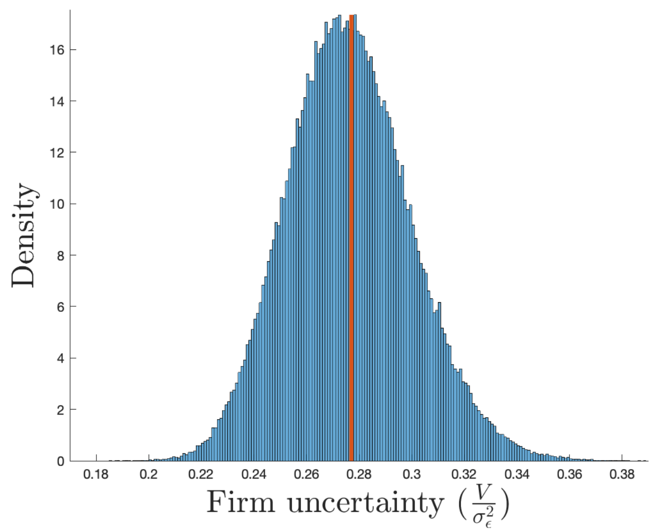


Table A.6: The distribution of signal precision.

The table below reports selected descriptive statistics associated with the distributions in Figure A.1.

				Percentile		
Model	Mean	SD	Skew	25^{th}	50^{th}	75^{th}
$\eta = 4$	0.642	0.089	0.413	0.580	0.636	0.698

B Model Proofs

B.1 Proof of Proposition 1

For the case where $\chi = 0$, we can compress the sub-periods into just the first two sub-periods, and impose $k' = k^p$. With this simplification, the Euler equation for investment plans after a signal has been observed is simply

$$\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}\left(w\right)\mathbb{E}\left[z^{\Theta_z\rho}|z_{-1},s\right]k^{p\Theta_k-1} = r+\delta.$$

With some rearrangement, we get that optimal investment plans are given by

$$\begin{aligned} k^p &= \left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}\left(w\right)}{r+\delta}\mathbb{E}\left[z^{\Theta_z\rho}|z_{-1},s\right]\right)^{\frac{1}{1-\Theta_k}} \\ &= \left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}\left(w\right)}{r+\delta}z_{-1}^{\rho^2\Theta_z}\right)^{\frac{1}{1-\Theta_k}}\exp\left(\frac{\rho\Theta_z}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(\frac{1}{2}\frac{\left(\rho\Theta_z\right)^2}{1-\Theta_k}\mathbb{V}\right) \quad (19) \end{aligned}$$

By substituting in equation 19 into the profit function, realized (next-period) profits, net of capital cost, can be expressed as

$$\begin{aligned} \pi' = \mathcal{A}\left(w\right)z'^{\Theta_z}&\left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}\left(w\right)}{r+\delta}z_{-1}^{\rho^2\Theta_z}\right)^{\frac{\Theta_k}{1-\Theta_k}}\exp\left(\frac{\rho\Theta_z\Theta_k}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(\frac{1}{2}\frac{\Theta_k}{1-\Theta_k}\left(\rho\Theta_z\right)^2\mathbb{V}\right) - \ldots \\ \ldots\left(r+\delta\right)&\left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}\left(w\right)}{r+\delta}z_{-1}^{\rho^2\Theta_z}\right)^{\frac{1}{1-\Theta_k}}\exp\left(\frac{\rho\Theta_z}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(\frac{1}{2}\frac{\left(\rho\Theta_z\right)^2}{1-\Theta_k}\mathbb{V}\right) \end{aligned}$$

Therefore, ex ante expected profits, gross of the information acquisition cost, can be expressed as,

$$
\begin{aligned}
\pi^e &\equiv \mathbb{E}\left[\mathbb{E}\left[\pi'|s\right]|\mathbb{V}, z_{-1}\right] \\
&= \mathbb{E}\Bigg[\mathcal{A}(w)\, z'^{\Theta_z}\left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}(w)}{r+\delta} z_{-1}^{\rho^2\Theta_z}\right)^{\frac{\Theta_k}{1-\Theta_k}} \exp\left(\frac{\rho\Theta_z\Theta_k}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(\frac{1}{2}\frac{\Theta_k}{1-\Theta_k}\left(\rho\Theta_z\right)^2\mathbb{V}\right) - \ldots \\
&\qquad \ldots (r+\delta)\left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}(w)}{r+\delta} z_{-1}^{\rho^2\Theta_z}\right)^{\frac{1}{1-\Theta_k}} \exp\left(\frac{\rho\Theta_z}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(\frac{1}{2}\frac{\left(\rho\Theta_z\right)^2}{1-\Theta_k}\mathbb{V}\right)|\mathbb{V}, z_{-1}\Bigg] \\
&= \mathbb{E}[\left(\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\mathcal{A}(w)\right)^{\frac{1}{1-\Theta_k}}(r+\delta)^{-\frac{\Theta_k}{1-\Theta_k}} z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}\exp\left(\frac{\rho\Theta_z}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\times \\
&\qquad \exp\left(\frac{1}{2}\frac{\left(\rho\Theta_z\right)^2}{1-\Theta_k}\mathbb{V}\right)\Theta_k^{\frac{1}{1-\Theta_k}}\left(\frac{1-\Theta_k}{\Theta_k}\right)|\mathbb{V}, z_{-1}] \\
&= \underbrace{\Theta_k^{\frac{1}{1-\Theta_k}}\left(\frac{1-\Theta_k}{\Theta_k}\right)\left(\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\mathcal{A}(w)\right)^{\frac{1}{1-\Theta_k}}(r+\delta)^{-\frac{\Theta_k}{1-\Theta_k}}\exp\left(\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2\sigma_\epsilon^2\right)}_{\equiv\bar{h}} z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}\times \\
&\qquad \underbrace{\exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2\Theta_k\mathbb{V}\right)}_{\equiv\mathcal{F}^\pi\left(\mathbb{V}^{-1}\right)} \\
&= \bar{h} z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}\mathcal{F}^\pi\left(\mathbb{V}^{-1}\right)
\end{aligned}
$$

The marginal benefit of information is then

$$
\begin{aligned}
\frac{1}{1+r}\frac{\partial\mathbb{E}\left[V\left(\prime\right)|\mathbb{V}, z_{-1}\right]}{\partial\mathbb{V}^{-1}} &= \frac{1}{1+r}\frac{\partial\pi^e}{\partial\mathbb{V}^{-1}} \\
&= \frac{1}{1+r}\bar{h} z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}\frac{\partial\mathcal{F}^\pi\left(\mathbb{V}^{-1}\right)}{\partial\mathbb{V}^{-1}}
\end{aligned}
$$

which is the expression in the main text for equation 7. Notice that the term $\bar{h} > 0$, while the partial derivative term evaluates to

$$\begin{aligned}\frac{\partial \mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial \mathbb{V}^{-1}} &= -\mathbb{V}^{2}\frac{\partial \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_{z}}{1-\Theta_{k}}\right)^{2}\Theta_{k}\mathbb{V}\right)}{\partial \mathbb{V}} \\ &= -\mathbb{V}^{2}\times\left(-\frac{1}{2}\left(\frac{\rho\Theta_{z}}{1-\Theta_{k}}\right)^{2}\Theta_{k}\right)\mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right) \\ &= \frac{1}{2}\left(\frac{\rho\Theta_{z}}{1-\Theta_{k}}\right)^{2}\Theta_{k}\mathbb{V}^{2}\mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right) > 0.\end{aligned}$$

In other words, the marginal benefit of purchasing more information is positive, which proves point 1 of Proposition 1. The proof of point 2 follows through trivially, since the exponent term on z_{-1} is strictly larger than 0.

B.2 Proof of Proposition 2

We next turn to proving Proposition 2.

Proof. From Proposition 1, we already showed that the marginal benefit of a lower posterior variance is increasing in z_{-1}. The marginal cost is given by

$$\frac{\partial}{\partial \mathbb{V}^{-1}}\xi z_{-1}^{\zeta}\sigma^{-2} = \xi z_{-1}^{\zeta}$$

where the cost of lowering posterior uncertainty is increasing, and its dependence on z_{-1} is determined by ζ. Equating the marginal cost and benefit, and with some trivial rearrangement of terms, gives us

$$\xi\left(1+r\right)z_{-1}^{\zeta-\frac{\rho^{2}\Theta_{z}}{1-\Theta_{k}}} = \bar{h}\frac{\partial \mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial \mathbb{V}^{-1}}, \tag{20}$$

where $\mathbb{V}^{-1}$ solves the implicit equation above. Importantly, notice that the right-hand side term is not a function of z_{-1}. Note that the expression $\frac{\partial \mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial \mathbb{V}^{-1}}$ has the following proporties: (1) $\lim_{\mathbb{V}^{-1}\to 0}\frac{\partial \mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial \mathbb{V}^{-1}} = 0$; (2) is increasing when $\mathbb{V}^{-1} < \tilde{\mathbb{V}}^{-1}$ for some threshold $\tilde{\mathbb{V}}^{-1}$; (3) bounded at the threshold $\tilde{\mathbb{V}}^{-1}$; (4) is decreasing for the same threshold when $\mathbb{V}^{-1} > \tilde{\mathbb{V}}^{-1}$;

and (5) $\lim_{\mathbb{V}^{-1}\to\infty}\frac{\partial\mathcal{F}^{\pi}\left(\mathbb{V}^{-1}\right)}{\partial\mathbb{V}^{-1}}=0$. In other words, the marginal benefit is increasing for values of $\mathbb{V}^{-1}$ (low information), but decreasing as the amount of information increases; correspondingly, the value function is convex at low levels of $\mathbb{V}^{-1}$ and concave otherwise. Consequently, if an interior solution to equation 20 exists, then two solutions $\left(\mathbb{V}_l^{-1},\mathbb{V}_h^{-1}\right)$ exist, with $\mathbb{V}_l^{-1}<\tilde{\mathbb{V}}^{-1}<\mathbb{V}_h^{-1}$. However, the value of $\mathbb{V}$ that maximizes the value function is $\mathbb{V}_h^{-1}$, so if an interior solution exists, the right-hand side term is decreasing in $\mathbb{V}^{-1}$. Otherwise, the solution will simply be $\tilde{\mathbb{V}}^{-1}$.

On the left-hand side, we see that the term is strictly decreasing in z_{-1} if $\zeta<\frac{\rho^2\Theta_z}{1-\Theta_k}$, constant in z_{-1} if $\zeta=\frac{\rho^2\Theta_z}{1-\Theta_k}$, and increasing in z_{-1} if $\zeta>\frac{\rho^2\Theta_z}{1-\Theta_k}$. In other words, the threshold $\bar{\zeta}$ described in the main text is simply $\frac{\rho^2\Theta_z}{1-\Theta_k}$. Therefore, an application of the inverse function theorem tell us that, for an interior solution, the choice of $\mathbb{V}^{-1}$ is increasing in z_{-1} if $\zeta<\bar{\zeta}$, constant in z_{-1} if $\zeta=\bar{\zeta}$, and decreasing in z_{-1} if $\zeta>\bar{\zeta}$. Otherwise, for cases where $\mathbb{V}=\tilde{\mathbb{V}}$, productivity is uncorrelated with learning. $\square$

B.3 Proof of Proposition 3

We now derive the proof for our general model. To prove this proposition, we will first prove the following two lemmas, which will establish that the value of the firm, gross of the signal acquisition cost but net of adjustment costs, is strictly increasing in the posterior uncertainty.

Lemma 1. *The expected ex ante adjustment cost is increasing in* $\mathbb{V}$*.*

Lemma 2. *The expected ex ante value of the firm, gross of adjustment costs, is decreasing in* $\mathbb{V}$*.*

We first derive some common terms that will be useful in proving the two lemmas. We begin by deriving the solution to k'. To do so, first recall that the Euler equation for investment reduces to

$$\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}\left(w\right)z^{\Theta_z\rho}k'^{\Theta_k-1}-\left(1+r\right)\chi\frac{k'}{k^p}=r+\delta-\left(1+r\right)\chi. \tag{21}$$

The log-linear approximate solution for k', around the non-stochastic steady state, is therefore given by

$$\Delta k' = \phi_z \Delta z + \phi_k \Delta k^p$$
$$\implies k' = \bar{k}^{1-\phi_k} z^{\phi_z} \left(k^p\right)^{\phi_k}, \tag{22}$$

where for some generic variable x and steady state value $\bar{x}$, $\Delta x \equiv x - \bar{x}$; $\phi_z = \frac{\Theta_z \rho (r+\delta)}{(1+r)\chi + (r+\delta)(1-\Theta_k)}$ and $\phi_k = \frac{(1+r)\chi}{(1+r)\chi + (r+\delta)(1-\Theta_k)}$; and $\bar{k} = \left[\frac{\Theta_k \mathcal{A}(w)}{r+\delta}\right]^{\frac{1}{1-\Theta_k}}$
(with $\exp\left(\frac{1}{2}\Theta_z^2 \sigma_\epsilon^2\right) \Theta_k \mathcal{A}(w) z^{\Theta_z \rho} k'^{\Theta_k - 1} = r + \delta$ in the non-stochastic steady-state). We see that $\phi_z > 0$ and $\phi_k \in (0,1)$, where in particular, $\lim_{\chi \to 0} \phi_z = \frac{\rho}{1-\Theta_k}$ and $\lim_{\chi \to 0} \phi_k = 0$ returns us to the usual frictionless model, and $\lim_{\chi \to \infty} \phi_z = 0$ and $\lim_{\chi \to \infty} \phi_k = 1$ moves us to a model where only plans matter.

Next, we can substitute this solution for k' into the Bellman equation in the main text, and obtain

$$W(k, s, z_{-1}, \sigma) = \max_{k^p} \mathbb{E}\left[\pi + (1-\delta)k - k' - \frac{\chi}{2}\left(\frac{k'}{k^p} - 1\right)^2 k^p + \beta \mathbb{E}\left[V\left(k', \tilde{z}', z\right) |z\right] |s, z_{-1}, \sigma\right]$$
$$s.t.$$
$$\log z \sim N\left(\rho \log z_{-1} + \frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s, \mathbb{V}\right),$$

where we write k' as a function of k^p and z. This gives us the expected value of the firm after a signal has been observed, but before a plan has been made. Taking first order conditions, and following some algebra, we can obtain the optimal k^p as

$$k^p = \sqrt{\mathbb{E}\left[k'^2 |s\right]}$$
$$\implies k^p = \sqrt{\mathbb{E}\left[\left(\bar{k}^{1-\phi_k} z^{\phi_z} \left(k^p\right)^{\phi_k}\right)^2\right]}$$
$$\implies k^p = \bar{k} z_{-1}^{\frac{\rho \phi_z}{1-\phi_k}} \left(\sqrt{\mathbb{E}\left[\exp\left(2\phi_z \epsilon\right) |s\right]}\right)^{\frac{1}{1-\phi_k}},$$

where ϵ is the underlying innovations of z. Noting that the posterior distribution of ϵ (i.e., after the signal s has been observed) is given by $\epsilon \sim N\left(\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s, \mathbb{V}\right)$ we can express k^p as

$$k^p = \bar{k} z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}} \left(\sqrt{\exp\left(2\phi_z \frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2} s + 2\phi_z^2 \mathbb{V}\right)} \right)^{\frac{1}{1-\phi_k}}$$
$$= \bar{k} z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}} \exp\left(\frac{\phi_z}{1-\phi_k} \frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2} s \right) \exp\left(\frac{\phi_z^2}{1-\phi_k} \mathbb{V} \right). \tag{23}$$

Before moving on, it is useful to note that planned investment is increasing in $\mathbb{V}$ for small enough s (in particular, it is always increasing in $\mathbb{V}$ when $s \to 0$). This reflects a precautionary term coming from insurance against any upside risk, which becomes increasingly dominant as s becomes smaller (in the limit, there is only upside risk and no downside risk). This term exists, in part, because our specific formulation of the adjustment cost is bounded below while unbounded above.

We can now formally derive a proof for Lemma 1.

Proof. We begin by substituting the solution for k' and k^p back into the original cost function, which gives

$$\phi(k', k^p) = \frac{\chi}{2} \left(\frac{k'}{k^p} - 1 \right)^2 k^p$$
$$= \frac{\chi}{2} \left(\exp(\phi_z \epsilon) \exp\left(-\frac{\phi_z \sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2} s \right) \exp\left(-\phi_z^2 \mathbb{V}\right) - 1 \right)^2 \ldots$$
$$\ldots \left(\bar{k} z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}} \exp\left(\frac{1}{1-\phi_k} \frac{\phi_z \sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2} s \right) \exp\left(\frac{\phi_z^2}{1-\phi_k} \mathbb{V} \right) \right).$$

The ex ante cost function prior to the realization of signals is therefore

$$
\begin{aligned}
\mathbb{E}\left[\phi\left(k',k^p\right)|\sigma\right] &= \mathbb{E}\left[\frac{\chi}{2}\left(\frac{k'}{k^p}-1\right)^2 k^p|\sigma\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\frac{\chi}{2}\left(\frac{k'}{k^p}-1\right)^2 k^p|s\right]|\sigma\right] \\
&= \mathbb{E}\left[\frac{\chi}{2}\mathbb{E}\left[\left(\exp\left(-\phi_z\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(-\phi_z^2\mathbb{V}\right)\exp\left(\phi_z\epsilon\right)\right)^2 ...\right.\right. \\
&\qquad ... -2\exp\left(-\phi_z\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(-\phi_z^2\mathbb{V}\right)\exp\left(\phi_z\epsilon\right)+1|s...\Big] \\
&\qquad ... \left(\bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}\exp\left(\frac{1}{1-\phi_k}\frac{\phi_z\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s\right)\exp\left(\frac{\phi_z^2}{1-\phi_k}\mathbb{V}\right)\right)|\sigma\Big] \\
&= \chi\bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}\mathcal{F}^A\left(\mathbb{V}\right),
\end{aligned}
\tag{24}
$$

where we define $\mathcal{F}^A\left(\mathbb{V}\right) \equiv \exp\left[\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2\right]\left(1-\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\right)\exp\left(\frac{\phi_z^2}{2}\left(\frac{1-2\phi_k}{(1-\phi_k)^2}\right)\mathbb{V}\right)$ for notational convenience, and noting that $\chi\bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}} > 0$.

We see that the $\mathcal{F}^A$ term is the only term in the expression that depends on $\mathbb{V}$. To study the impact of $\mathbb{V}$ on the expected adjustment cost, it therefore suffices to study the marginal effect of changing $\mathbb{V}$ on $\mathcal{F}^A$. To do so, we can take the derivative of $\mathcal{F}^A$ with respect to $\mathbb{V}$, obtaining,

$$
\begin{aligned}
\frac{\partial\mathcal{F}^A\left(\mathbb{V}\right)}{\partial\mathbb{V}} &= \exp\left[\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2\right]\left(\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\frac{1}{2}\phi_z^2\exp\left(\frac{\phi_z^2}{2}\left(\frac{1-2\phi_k}{(1-\phi_k)^2}\right)\mathbb{V}\right)+...\right. \\
&\qquad ...\left.\left(1-\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\right)\exp\left(\frac{\phi_z^2}{2}\left(\frac{1-2\phi_k}{(1-\phi_k)^2}\right)\mathbb{V}\right)\frac{\phi_z^2}{2}\left(\frac{1-2\phi_k}{(1-\phi_k)^2}\right)\right) \\
&= \exp\left[\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2\right]\exp\left(\frac{\phi_z^2}{2}\left(\frac{1-2\phi_k}{(1-\phi_k)^2}\right)\mathbb{V}\right)\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2 ... \\
&\qquad ...\left(2\phi_k-1\right)\left(\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\left(\frac{\phi_k^2}{2\phi_k-1}\right)-1\right).
\end{aligned}
$$

We now proceed to show our proof for three cases.

Case 1: $\phi_k \leq \frac{1}{2}$.

In this case, $2\phi_k - 1 < 0$ and likewise $\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\left(\frac{\phi_k^2}{2\phi_k-1}\right) - 1 < 0$. Therefore, it is clear from the above expression that $\frac{\partial\mathcal{F}^A(\mathbb{V})}{\partial\mathbb{V}} > 0$ for all $\mathbb{V}$. In other words, the expected adjustment cost is always strictly increasing in $\mathbb{V}$ if $\phi_k \leq \frac{1}{2}$.

Case 2: $\frac{1}{2} < \phi_k < 1$

In this case, $2\phi_k - 1 > 0$ and likewise $\frac{\phi_k^2}{2\phi_k-1} > 1$. To show that $\frac{\partial\mathcal{F}^A(\mathbb{V})}{\partial\mathbb{V}} > 0$, we need to show that $\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\left(\frac{\phi_k^2}{2\phi_k-1}\right) - 1 > 0$. This implies that $\mathbb{V} < \frac{2}{\phi_z^2}\log\left(\frac{\phi_k^2}{2\phi_k-1}\right)$, or equivalently, $\mathbb{V} < 2\left(\frac{1}{\rho(\eta-1)}\right)^2\left(\frac{1}{1-\phi_k}\right)^2\log\left(\frac{\phi_k^2}{2\phi_k-1}\right)$. Now recall that because $\mathbb{V} < \sigma_\epsilon^2$, as long as $\sigma_\epsilon^2 < 2\left(\frac{1}{\rho(\eta-1)}\right)^2\left(\frac{1}{1-\phi_k}\right)^2\log\left(\frac{\phi_k^2}{2\phi_k-1}\right)$ is satisfied, we can establish case 2. Keeping in mind that $\left(\frac{1}{1-\phi_k}\right)^2\log\left(\frac{\phi_k^2}{2\phi_k-1}\right) > 1$ when $\frac{1}{2} < \phi_k < 1$, this implies that for most reasonable calibrations, $\sigma_\epsilon^2 \ll 2\left(\frac{1}{\rho(\eta-1)}\right)^2\left(\frac{1}{1-\phi_k}\right)^2\log\left(\frac{\phi_k^2}{2\phi_k-1}\right)$. Therefore, $\frac{\partial\mathcal{F}^A(\mathbb{V})}{\partial\mathbb{V}} > 0$ for all feasible choices of $\mathbb{V}$. In other words, the expected adjustment cost is again always strictly increasing in $\mathbb{V}$ if $\frac{1}{2} < \phi_k < 1$.

Case 3: $\phi_k = 1$

Note that this is the limiting case for which $\chi \to \infty$, where we also have $\phi_z = 0$. In this limiting case, we see from the expression that $\lim_{\chi\to\infty}\exp\left(-\frac{1}{2}\phi_z^2\mathbb{V}\right)\left(\frac{\phi_k^2}{2\phi_k-1}\right) - 1 = 0$, so $\lim_{\chi\to\infty}\frac{\partial\mathcal{F}^A(\mathbb{V})}{\partial\mathbb{V}} = 0$. Intuitively, since adjustment costs are infinitely large, the marginal effect of improving information is trivially zero. In other words, we have shown that the expected cost of violating the adjustment friction is increasing in the posterior uncertainty $\mathbb{V}$. This concludes the proof for Lemma 1. □

We can also now formally derive a proof for Lemma 2.

Proof. To begin, recall that since the choice of $\mathbb{V}$ only affects next-period profits, this implies that $\mathbb{V}$ only affects the value of the firm through this channel. Recalling that expected profits net of investment cost (but gross of the adjustment cost), π^e, is given by the expression

$$\pi^e \equiv \mathbb{E}\left[-k' + \frac{1}{1+r}\left(\mathbb{E}\left[\mathcal{A}\left(w,Y\right)z'^{\Theta_z}k'^{\Theta_k}|z\right] + (1-\delta)\,k'\right)|\sigma\right],$$

it suffices to show that $\frac{\partial\pi^e}{\partial\mathbb{V}} < 0$ to show that higher posterior uncertainty has a negative impact on the firm's value gross of adjustment cost. To do so, it is convenient to rewrite the expected next-period profit as

$$\pi^e = \frac{1}{1+r}\mathbb{E}\left[\mathbb{E}\left[\mathcal{A}\left(w,Y\right)z'^{\Theta_z}k'^{\Theta_k}|z\right] - \left(r+\delta\right)k'|\sigma\right].$$

We will now proceed to derive this as a function of $\mathbb{V}$ (and other initial conditions and parameters). We can show that

$$\mathbb{E}\left[k'|\sigma\right] = \bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}\exp\left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2\right)\exp\left(-\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2\mathbb{V}\right)$$

and

$$\mathbb{E}\left[\mathbb{E}\left[\mathcal{A}\left(w\right)z'^{\Theta_z}k'^{\Theta_k}|z\right]|\sigma\right] = \mathcal{A}\left(w,Y\right)\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\bar{k}^{\Theta_k}z_{-1}^{\rho\frac{\phi_z}{1-\phi_k}}\exp\left(\frac{1}{2}\vartheta\sigma_\epsilon^2\right)...$$

$$...\exp\left(-\frac{1}{2}\Theta_k^2\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2\frac{2-\Theta_k}{\Theta_k}\mathbb{V}\right),$$

with $\vartheta \equiv \left(\rho\Theta_z\right)^2 + 2\rho\Theta_z\Theta_k\frac{\phi_z}{1-\phi_k} + \left(\Theta_k\frac{\phi_z}{1-\phi_k}\right)^2$.

Before proceeding, it is worth discussing briefly the economic intuition here. Notice that the $\mathbb{E}\left[k'|\sigma\right]$ term is decreasing in $\mathbb{V}$, that is, expected investment is decreasing in the posterior uncertainty. This contrasts with the $\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2$ term, which says that expected investment is increasing in the dispersion of productivity. This term is the usual volatility effect that predicts investment increasing with uncertainty (due to its real option value), whereas subjective uncertainty $\mathbb{V}$ decreases investment. Notice however that the term here is pre-multiplied by ϕ_k, which is the weight on the investment plan in the manager's investment policy function. A more standard model of Bayesian learning would imply that $\phi_k = 1$. In our model, we nest this standard framework by allowing partial flexibility of plans. As such, the ability to weakly deviate from planned investment (i.e., $\phi_k < 1$) dampens the effect of subjective uncertainty.

Moreover, notice that $\mathbb{E}\left[\mathbb{E}\left[\mathcal{A}\left(w\right)z'^{\Theta_z}k'^{\Theta_k}|z\right]|\sigma\right]$ (expected revenue) is also decreasing in

$\mathbb{V}$. Since expected investment is decreasing in $\mathbb{V}$, this result is not surprising. Now with these two terms in hand, we can rewrite the expected profits as

$$\pi^e = \frac{1}{1+r}\left[\mathcal{A}(w,Y)\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\bar{k}^{\Theta_k} z_{-1}^{\rho\frac{\phi_z}{1-\phi_k}} \exp\left(\frac{1}{2}\vartheta\sigma_\epsilon^2\right)\dots\right.$$
$$\dots\exp\left(-\frac{1}{2}\Theta_k^2\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2\frac{2-\Theta_k}{\Theta_k}\mathbb{V}\right)-$$
$$\dots(r+\delta)\,\bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}\exp\left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2\right)$$
$$\left.\dots\exp\left(-\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2\mathbb{V}\right)\right]$$
$$= \frac{1}{1+r}\bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}\exp\left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\sigma_\epsilon^2\right)\mathcal{F}^\pi(\mathbb{V}), \tag{25}$$

where $\mathcal{F}^\pi(\mathbb{V})$ is a function of the posterior uncertainty and other model parameters. Critically, this function does not include z_{-1}, which pre-multiplies this function. Therefore, the effect of $\mathbb{V}$ is always scaled by z_{-1}. To show that increasing posterior uncertainty decreases expected profits, we simply need to show that $\frac{\partial\mathcal{F}^\pi(\mathbb{V})}{\partial\mathbb{V}} < 0$. We derive

$$\frac{\partial\mathcal{F}^\pi(\mathbb{V})}{\partial\mathbb{V}} = \left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2\right)\exp\left(-\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2\mathbb{V}\right)\dots$$
$$\left[\mathcal{A}(w,Y)\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\bar{k}^{\Theta_k-1}\exp\left(\frac{1}{2}\left(\vartheta-\left(\frac{\phi_z}{1-\phi_k}\right)^2\right)\sigma_\epsilon^2\right)\dots\right.$$
$$\left.\dots(-\Theta_k(2-\Theta_k))\exp\left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2(1-\Theta_k(2-\Theta_k))\mathbb{V}\right)+(r+\delta)\right],$$

which means that for the previous condition to hold, we need the following condition

$$\mathcal{A}(w,Y)\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\bar{k}^{\Theta_k-1}\exp\left(\frac{1}{2}\left(\vartheta-\left(\frac{\phi_z}{1-\phi_k}\right)^2\right)\sigma_\epsilon^2\right)\dots > r+\delta$$
$$\dots(\Theta_k(2-\Theta_k))\exp\left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2(1-\Theta_k(2-\Theta_k))\mathbb{V}\right).$$

But recall that $\bar{k} = \left[\frac{\Theta_k \mathcal{A}(w)}{r+\delta}\right]^{\frac{1}{1-\Theta_k}}$, which means that $\bar{k}^{\Theta_k - 1} = \frac{r+\delta}{\Theta_k \mathcal{A}(w,Y)}$. Substituting this back into the equation above, we get

$$\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\exp\left(\frac{1}{2}(\rho\Theta_z)^2\frac{\Theta_k}{1-\Theta_k}\sigma_\epsilon^2\right)(2-\Theta_k)\exp\left(\frac{1}{2}\left(\frac{\phi_z}{1-\phi_k}\right)^2\phi_k^2(1-\Theta_k(2-\Theta_k))\mathbb{V}\right) > 1.$$

Further substituting in $\vartheta - \left(\frac{\phi_z}{1-\phi_k}\right)^2 = (\rho\Theta_z)^2\frac{\Theta_k}{1-\Theta_k}$, we reduce the expression to,

$$\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2 + \frac{1}{2}(\rho\Theta_z)^2\frac{\Theta_k}{1-\Theta_k}\sigma_\epsilon^2 + \frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2\phi_k^2\left(1-\Theta_k\left(1-\frac{1}{2}\Theta_k\right)\right)\mathbb{V}\right) > \frac{1}{2-\Theta_k}.$$

Notice that since $\Theta_k \in (0,1)$, $\frac{1}{2-\Theta_k} \in \left(\frac{1}{2},1\right)$, the above relation is trivially true for all parameter values, so $\frac{\partial\mathcal{F}^\pi(\mathbb{V})}{\partial\mathbb{V}} < 0$. Therefore, expected profits, gross of adjustment cost, is decreasing in the posterior variance. This concludes the proof for Lemma 2. □

With Lemma 1 and 2 in hand, we can now derive a proof for our proposition.

Proof. First, Lemma 2 tells us that the ex ante value of the firm, gross of signal acquisition costs but net of adjustment cost, is strictly decreasing in $\mathbb{V}$. This gives us, for the general case, the equivalent of point 1 in Proposition 1.

To show point 2 of Proposition 1 in our general case, we can simply combine both the expected adjustment cost and expected profits and take the first derivative with respect to $\mathbb{V}$. To recall, the first derivative of the expected adjustment cost is

$$\frac{\partial\mathbb{E}\left[\phi\left(k',k^p\right)|\sigma\right]}{\partial\mathbb{V}} = \chi\bar{k}z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}\frac{\partial\mathcal{F}^A(\mathbb{V})}{\partial\mathbb{V}},$$

and recall $\frac{\partial\mathcal{F}^A(\mathbb{V})}{\partial\mathbb{V}} > 0$. The marginal effect of $\mathbb{V}$ on firm value, net of adjustment cost, can then be expressed as

$$-\frac{\partial \mathbb{E}\left[\phi\left(k', k^{p}\right)|\sigma\right]}{\partial \mathbb{V}}+\frac{\partial \pi^{e}}{\partial \mathbb{V}}=-\chi \bar{k} z_{-1}^{\frac{\rho \phi_{z}}{1-\phi_{k}}} \frac{\partial \mathcal{F}^{A}(\mathbb{V})}{\partial \mathbb{V}}+\frac{1}{1+r} \bar{k} z_{-1}^{\frac{\rho \phi_{z}}{1-\phi_{k}}} \exp \left(\frac{1}{2}\left(\frac{\phi_{z}}{1-\phi_{k}}\right)^{2} \sigma_{\epsilon}^{2}\right) \frac{\partial \mathcal{F}^{\pi}(\mathbb{V})}{\partial \mathbb{V}}$$

$$=\bar{k} z_{-1}^{\frac{\rho \phi_{z}}{1-\phi_{k}}}\left(-\chi \frac{\partial \mathcal{F}^{A}(\mathbb{V})}{\partial \mathbb{V}}+\frac{1}{1+r} \exp \left(\frac{1}{2}\left(\frac{\phi_{z}}{1-\phi_{k}}\right)^{2} \sigma_{\epsilon}^{2}\right) \frac{\partial \mathcal{F}^{\pi}(\mathbb{V})}{\partial \mathbb{V}}\right),$$

Note that this gives us, for the general case, the equivalent expression in Equation 7 in the main text. Here, we also see that the term $-\chi \frac{\partial \mathcal{F}^{A}(\mathbb{V})}{\partial \mathbb{V}}+\frac{1}{1+r} \exp \left(\frac{1}{2}\left(\frac{\phi_{z}}{1-\phi_{k}}\right)^{2} \sigma_{\epsilon}^{2}\right) \frac{\partial \mathcal{F}^{\pi}(\mathbb{V})}{\partial \mathbb{V}}<0$ and does not depend on z_{-1}. In other words, initial productivity z_{-1} has a pure scaling effect; that is, firms with higher initial productivity face a steeper cost of having a more dispersed signal. Conversely, the benefits to get a better signal is increasing in initial productivity. Formally, this statement is seen in the cross-derivative, which is given by

$$\frac{\partial}{\partial z_{-1}}\left(-\frac{\partial \mathbb{E}\left[\phi\left(k', k^{p}\right)|\sigma\right]}{\partial \mathbb{V}}+\frac{\partial \pi^{e}}{\partial \mathbb{V}}\right)=\frac{\rho \phi_{z}}{1-\phi_{k}} \bar{k} z_{-1}^{\frac{\rho \phi_{z}}{1-\phi_{k}}-1} \ldots$$

$$\ldots\left(\chi \frac{\partial \mathcal{F}^{A}(\mathbb{V})}{\partial \mathbb{V}}+\frac{1}{1+r} \exp \left(\frac{1}{2}\left(\frac{\phi_{z}}{1-\phi_{k}}\right)^{2} \sigma_{\epsilon}^{2}\right) \frac{\partial \mathcal{F}^{\pi}(\mathbb{V})}{\partial \mathbb{V}}\right)$$

$$<0$$

Therefore, the benefits of having a lower posterior uncertainty is increasing in initial firm productivity z_{-1}. This thus concludes the proof, for the general case, of the equivalent of point 2 in Proposition 1. □

B.3.1 Derivation of Corollary 1

The derivation of this corollary is similar to the derivation of Proposition 1. Because the marginal cost function is identical, we can see that the indifference condition for the general model is given by

$$\xi z_{-1}^{\zeta-\frac{\rho \phi_{z}}{1-\phi_{k}}}=\bar{k}\left(-\chi \frac{\partial \mathcal{F}^{A}(\mathbb{V})}{\partial \mathbb{V}^{-1}}+\frac{1}{1+r} \exp \left(\frac{1}{2}\left(\frac{\phi_{z}}{1-\phi_{k}}\right)^{2} \sigma_{\epsilon}^{2}\right) \frac{\partial \mathcal{F}^{\pi}(\mathbb{V})}{\partial \mathbb{V}^{-1}}\right), \tag{26}$$

where $\mathbb{V}^{-1}$ solves the implicit equation above. This formula is more complicated than the special case for $\chi = \infty$, but reflects the same insights. In particular, on the left-hand side, we see that the term is strictly decreasing in z_{-1} if $\zeta < \frac{\rho\phi_z}{1-\phi_k}$, constant in z_{-1} if $\zeta = \frac{\rho\phi_z}{1-\phi_k}$, and increasing in z_{-1} if $\zeta > \frac{\rho\phi_z}{1-\phi_k}$. In other words, the threshold $\bar{\zeta}$ in this general case is simply $\frac{\rho\phi_z}{1-\phi_k}$, which in fact evaluates to $\frac{\rho^2\Theta_z}{1-\Theta_k}$. Finally, an application of the inverse function theorem tell us that the choice of $\mathbb{V}^{-1}$ is increasing in z_{-1} if $\zeta < \bar{\zeta}$, constant in z_{-1} if $\zeta = \bar{\zeta}$, and decreasing in z_{-1} if $\zeta > \bar{\zeta}$.

B.4 Proof of Proposition 4

We now derive the proof showing that information and productivity are complements in production. To do this, we first derive an expression for TFP, per the definition in the main text as follows,

Claim 1. TFP can be expressed as

$$TFP = \left\{\mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}\right]\mathbb{E}\left[\exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2\Theta_k\mathbb{V}\right)\right] + cov\left(z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}, \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2\Theta_k\mathbb{V}\right)\right)\right\}^{\frac{1-\nu}{\nu}}$$

Proof. To derive this expression, we first note that the production function can be re-expressed without labor as

$$y = z^{\Theta_z\nu^{-1}}k^{\Theta_k\nu^{-1}},$$

Next, we simply utilize Equation 23 by imposing $k' = k^p$, and substituting the solution back into the production function,

$$y = z^{\Theta_z\nu^{-1}}\left[\left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}(w)}{r+\delta}z_{-2}^{\rho^2\Theta_z}\right)^{\frac{1}{1-\Theta_k}}\exp\left(\frac{\rho\Theta_z}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s_{-1}\right)\exp\left(\frac{1}{2}\frac{(\rho\Theta_z)^2}{1-\Theta_k}\mathbb{V}_{-1}\right)\right]^{\Theta_k\nu^{-1}}$$

$$\Leftrightarrow y^\nu = z^{\Theta_z}\left[\left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k\mathcal{A}(w)}{r+\delta}z_{-2}^{\rho^2\Theta_z}\right)^{\frac{1}{1-\Theta_k}}\exp\left(\frac{\rho\Theta_z}{1-\Theta_k}\frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2}s_{-1}\right)\exp\left(\frac{1}{2}\frac{(\rho\Theta_z)^2}{1-\Theta_k}\mathbb{V}_{-1}\right)\right]^{\Theta_k},$$

Now, to compute $\int y^{\nu} d\Lambda$, we simply integrate over (i) all TFP shocks given a signal, (ii) all possible signals for each z_{-2} and $\mathbb{V}_{-1}$, and (iii) all possible z_{-2} and $\mathbb{V}_{-1}$. This gives us,

$$\begin{aligned}\int y^{\nu} &= \mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1}\right)\right] \times \\ &\quad \left[\exp\left(\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \sigma_\epsilon^2\right) \exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right) \left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k \mathcal{A}(w)}{r+\delta}\right)^{\frac{\Theta_k}{1-\Theta_k}}\right] \\ &= \mathcal{X}_1^{\nu} \mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1}\right)\right],\end{aligned}$$

where $\mathcal{X}_1$ is a constant that is a function of model parameters and prices, and the expectation operator is computed over the joint distribution of z_{-2} and $\mathbb{V}_{-1}$. This implies that aggregate output Y can be expressed as

$$Y = \mathcal{X}_1 \mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}\right)\right]^{\frac{1}{\nu}},$$

where we have iterated the variables one time step forward. Using a similar logic, we can derive aggregate capital and labor as follows,

$$K = \left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k \mathcal{A}(w)}{r+\delta}\right)^{\frac{1}{1-\Theta_k}} \exp\left(\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \sigma_\epsilon^2\right) \mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1}\right)\right]$$

$$\begin{aligned}L = &\left(\frac{(1-\alpha)\nu}{w}\right)^{\frac{1-\alpha}{1-(1-\alpha)\nu}} \exp\left(\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \sigma_\epsilon^2\right) \exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right) \left(\frac{\exp\left(\frac{1}{2}\Theta_z^2\sigma_\epsilon^2\right)\Theta_k \mathcal{A}(w)}{r+\delta}\right)^{\frac{\Theta_k}{1-\Theta_k}} \times \\ &\mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1}\right)\right]\end{aligned}$$

This gives us

$$
\begin{aligned}
K^{\alpha}L^{1-\alpha} &= \mathcal{X}_2 \mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1} \right) \right]^{\alpha} \mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1} \right) \right]^{1-\alpha} \\
&= \mathcal{X}_2 \mathbb{E}\left[z_{-2}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V}_{-1} \right) \right] \\
&= \mathcal{X}_2 \mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V} \right) \right],
\end{aligned}
$$

where in the last line, we also iterate the variables one time step forward. Coming both expressions, we obtain,

$$
\begin{aligned}
TFP &= \frac{Y}{K^{\alpha}L^{1-\alpha}} \\
&\propto \frac{\mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V} \right) \right]^{\frac{1}{\nu}}}{\mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V} \right) \right]} \\
&= \mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V} \right) \right]^{\frac{1-\nu}{\nu}} \\
&= \left\{ \mathbb{E}\left[z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}} \right] \mathbb{E}\left[\exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V} \right) \right] + cov\left(z_{-1}^{\frac{\rho^2\Theta_z}{1-\Theta_k}}, \exp\left(-\frac{1}{2}\left(\frac{\rho\Theta_z}{1-\Theta_k}\right)^2 \Theta_k \mathbb{V} \right) \right) \right\}^{\frac{1-\nu}{\nu}},
\end{aligned}
$$

where the last line is the formula given in the claim. With the above expression, it is now straightforward to prove Proposition 4. To see this, it is trivial to see that TFP is increasing in the covariance term for any given marginal distributions of z_{-1} and $\mathbb{V}$; likewise, TFP is increasing in the correlation of the arguments to the covariance term. Note that the covariance term is positive if z_{-1} and $\mathbb{V}$ are negatively correlated, or equivalently, if z_{-1} and $\mathbb{V}^{-1}$ are *positively correlated.* In other words, as stated in the proposition, productivity and information are complements in production, and TFP is increasing in their correlation. □

B.5 Derivations for Identification

We derive our identification argument for Section 3.3.1. For concreteness, we will refer to the dispersion of a variable as the standard deviation of that variable. TFP shocks and investment plan deviations follow our model definition, namely $\log z - \mathbb{E}(\log z | z_{-1}, s)$ and $\log \frac{k'}{k^p}$ respectively.

B.5.1 Derivation for Identification 1 (ξ)

Here, it is useful to examine the indifference condition as discussed in Equation 26. Notice that in this case, the only parameters that affect firm learning are ξ and ζ, since $\chi = \infty$. Now, consider the case where $\zeta = \bar{\zeta}$, such that all firms face the same marginal cost of information acquisition. As such, the solution to the firm's information acquisition problem only depends on the fixed cost ξ. Since the left-hand side term is increasing in ξ, $\mathbb{V}^{-1}$ (which is homogenous across all firms), is correspondingly weakly decreasing in ζ. Consequently, the dispersion of TFP shocks is increasing in ζ. For the generic case where $\zeta \neq \bar{\zeta}$, the dispersion of TFP shocks is simply given by

$$var(\log z - \mathbb{E}(\log z | z_{-1}, s)) = \mathbb{E}[\mathbb{V}(z_{-1})].$$

Since we have already demonstrated that $\mathbb{V}(z_{-1})$ is increasing in ξ, the dispersion of TFP shocks is also increasing in ξ.

Effect on identification 2 An obvious issue with this partial identification argument is that changes in χ, which affects investment flexibility, will also affect the incentives of firms to acquire information ex-ante. We will turn to jointly discussing this identification issue when we discuss the identification of χ.

Effect on identification 3 As we can see from the case where $\zeta = \bar{\zeta}$, ξ has no direct effect on the correlation between information acquisition choices and productivity.

B.5.2 Derivation for Identification 2 (χ)

By combining Equations 22 and 23, we can express investment plan deviations as

$$\log \frac{k'}{k^p} = \phi_z \tilde{\epsilon} - \phi_z^2 \mathbb{V}, \tag{27}$$

where $\tilde{\epsilon} \equiv \epsilon - \frac{\sigma_\epsilon^2}{\sigma^2+\sigma_\epsilon^2} s \sim N(0, \mathbb{V})$. The variance of this object for a given z_{-1} is then

$$var\left(\log \frac{k'}{k^p} | z_{-1}\right) = \phi_z^2 \left(\mathbb{V}(z_{-1}) + \phi_z^2 var\left(\mathbb{V}(z_{-1})\right)\right),$$

where the covariance term drops out since the TFP shock is uncorrelated with productivity. Finally, using the law of total variance, we obtain

$$\begin{aligned} var\left(\log \frac{k'}{k^p}\right) &= \mathbb{E}\left[var\left(\log \frac{k'}{k^p} | z_{-1}\right)\right] + var\left(\mathbb{E}\left[\log \frac{k'}{k^p} | z_{-1}\right]\right) \\ &= \mathbb{E}\left[\phi_z^2 \left(\mathbb{V} + \phi_z^2 var(\mathbb{V})\right)\right] + \phi_z^2 \mathbb{E}[\mathbb{V}] \\ &= 2\phi_z^2 \mathbb{E}[\mathbb{V}] + \phi_z^4 var(\mathbb{V}) \\ &\approx 2\phi_z^2 \mathbb{E}[\mathbb{V}]. \end{aligned}$$

where the last approximation is fairly accurate as the dispersion of TFP shocks are typically orders of magnitude larger than the dispersion of the dispersion itself. As we can see, holding the optimal learning decision fixed, ϕ_z is decreasing in χ and consequently $var\left(\log \frac{k'}{k^p}\right)$ is decreasing in χ. This result is intuitive: if investment flexibility is decreasing, naturally the dispersion of investment plan deviations must decrease.

Effect on identification 1 An obvious issue with the partial identification argument above is that changes in χ in turn changes the optimal learning decision. In particular, as demonstrated in the indifference condition, increases in χ will prompt firms to preemptively acquire more information, since the marginal benefit of information is higher. However, what this implies is that the average uncertainty $\mathbb{E}[\mathbb{V}]$ is decreasing in χ. Correspondingly,

$var\left(\log\frac{k'}{k^p}\right)$ is monotonically decreasing in χ. The key difference between ξ and χ, therefore, is simple. ξ affects $var\left(\log\frac{k'}{k^p}\right)$ only through its effect on $\mathbb{E}\left[\mathbb{V}\right]$; in contrast, χ has a "direct" effect by changing ϕ_z, and an "indirect effect" from changing $\mathbb{E}\left[\mathbb{V}\right]$. Importantly, as we discussed in the main text, the ratio of the two moments of interest can be expressed as

$$\frac{var\left(\log z-\mathbb{E}\left(\log z|z_{-1},s\right)\right)}{var\left(\log\frac{k'}{k^p}\right)}=2\phi_z^2, \tag{28}$$

From before, we showed that ϕ_z is only a function of χ. In other words, we have "two equations in two unknowns" that are exactly identified.

Effect on identification 3 Like ξ, as we can see from the case where $\zeta=\bar{\zeta}$, χ has no direct effect on the correlation between information acquisition choices and productivity.

B.5.3 Derivation for Identification 3 (ζ)

In this case, as we already derived, the firm's choice of information $\mathbb{V}$ is the solution to the indifference condition in Equation 26, where as we already saw, $\mathbb{V}$ is a weakly monotone function of z_{-1}. More formally, the correlation between the two variables can be captured by the following:

$$\begin{aligned}\frac{\partial\mathbb{V}}{\partial z_{-1}}&=\frac{\partial\mathbb{V}}{\partial z_{-1}^{\frac{\rho\phi_z}{1-\phi_k}}}\frac{\partial z_{-1}^{\zeta-\frac{\rho\phi_z}{1-\phi_k}}}{\partial z_{-1}}\\&=\frac{\partial\mathbb{V}}{\partial z_{-1}^{\zeta-\frac{\rho\phi_z}{1-\phi_k}}}\left(\zeta-\frac{\rho\phi_z}{1-\phi_k}\right)z^{\zeta-\frac{\rho\phi_z}{1-\phi_k}-1}\end{aligned}$$

Using our indifference equation, we see that the term $\frac{\partial\mathbb{V}}{\partial z_{-1}^{\zeta-\frac{\rho\phi_z}{1-\phi_k}}}$ does not depend on ζ. Consequently, holding all else constant, $\frac{\partial\mathbb{V}}{\partial z_{-1}}$ is strictly increasing in ζ per the second term. In turn, the correlation of $\mathbb{V}$ (i.e., dispersion of TFP shocks) is increasing in ζ.

Effect on identification 1 and 2 Changes to ζ will affect the average cost of information acquisition, holding all else constant, and therefore has an indirect effect on the average dispersion of TFP shocks (and thus investment plan deviations). However, changes to ζ do not map in a monotonic way to the dispersion of TFP shocks. This is because for any increase in ζ, lower productivity firms (specifically, when $z < 1$) face a lower cost of information acqusition, whereas the converse is true for higher productivity firms. This means that changes to average information acquisition come from weighting increased information acquisition from low-productivity firm and weighting reduced information acquisition from high-productivity firms. This non-monotonic effect therefore (loosely speaking) allows us to use the first two moments to pin down ξ and χ. This can be seen in the "isomoment plots" in Figure 3 (Panels b, c, e, and f), where we see that multiple values of ζ can be combined with ξ (or χ) to produce the same two moments. Notice that the connection between TFP shocks and investment plan deviations are also reflected in how the isomoment plots for the dispersion of TFP shocks is qualitatively the same as that of the dispersion of investment plan deviations.

B.6 Derivations for External Validation Exercises

We derive our derivations for the predictions discussed in Section 3.3.2.

B.6.1 Derivation for Prediction 1

Following the earlier proofs, it is straightforward to rewrite the investment Euler equation, in terms of expected TFP, as

$$\log k' = \log \bar{k} + \left(\phi_z + \phi_k \frac{\rho \Theta_z}{1 - \Theta_k}\right) \mathbb{E}\left[\log z | z_{-1}\right] - \phi_z \frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s + \frac{\phi_k \phi_z^2}{1 - \phi_k} \mathbb{V} + \phi_z \epsilon,$$

where ϵ is an i.i.d. TFP shock. Therefore, when we simply run a regression of investment on expected productivity, we are estimating the combination of parameters $\phi_z + \phi_k \frac{\rho \Theta_z}{1-\Theta_k}$.[20]

In contrast, when we include investment plans as an additional regressor, our estimating

[20]This discussion assumes that we observe the true expected productivity. As discussed earlier, we do not directly observe $\mathbb{E}\left[\log z | z_{-1}\right]$, so in general, such a regression will suffer from attenuation bias.

equation becomes,

$$\log k' = \log \bar{k} + \phi_z \mathbb{E}\left[\log z | z_{-1}\right] + \phi_k \log k^p - \phi_z \frac{\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} s + \phi_z \epsilon.$$

In other words, the importance of expected TFP is attenuated since we are now just estimating ϕ_z. Noticeably, the degree of attenuation is increasing in ϕ_k. In particular, if investment plans were fully flexible (i.e., $\phi_k = 0$), we would not see any attenuation in the importance of expected TFP.

B.6.2 Derivation for Prediction 2

Prediction 2 follows trivially from Equation 27. Squaring both sides of the equation, we can see that $\left(\log \frac{k'}{k^p}\right)^2 \propto \phi_z^2 \tilde{\epsilon}^2$ where $\tilde{\epsilon}$ is the TFP shock. In other words, the size of investment plan deviations is increasing in the size of realized TFP shocks so long as $\phi_z \neq 0$. However, this does not reject the null that investment plans are fully flexible ($\chi = 0$), for which ϕ_z is a finite-valued parameter.

Not Intended For Publication

C More Details of the Business Outlook and Financial Statements Survey

C.1 Details of the Business Outlook Survey

The Ministry of Finance (MoF) and Cabinet Office of Japan conduct the Business Outlook Survey (BOS) in order to help forecast the economy and inform fiscal planning. The survey is implemented primarily to evaluate the investment condition of firms and is conducted separately from other Japanese survey data that have been used by existing researchers such as the Tankan survey conducted by the Bank of Japan that "aims to provide an accurate picture of business trends of enterprises in Japan, thereby contributing to the appropriate implementation of monetary policy". The MoF uses some information from the BOS in producing its Monthly Economic Report, which is made publicly available in the form of aggregated statistics. None of the disaggregated individual firm-level data ever become public.

The BOS has been conducted from the first quarter of Fiscal Year (FY) 2004 as a General Statistical Survey under the Statistics Act. The first Statistics Act (Act No. 18 of 1947) was later revised in Act No. 53, which was passed on May 23, 2007 and implemented in April 2009. The survey targets are non-financial corporations with paid-in-capital of at least 10 million yen (approx. 100,000 USD) and utilities and financial institutions with paid-in-capital of at least 100 million yen (approx. 1 million USD). Responses are collected both via mail and online, and the sampling is based on the corporations covered by quarterly surveys of Financial Statements Statistics of Corporations by Industry. Industries in the aggregated survey result are based on Japanese SIC (J-SIC) codes (the 2-digit figure for manufacturing firms and larger "alphabet" group classifications for non-manufacturing industries) and grouped into 45 MoF industries. This means that some 2-digit J-SIC industries are grouped into a larger classification for reporting. In all our analyses, we use these MoF industry definitions as our definition of industries.

The survey is administered as repeated cross sections with stratified sampling across seven categories by registered capital and industry to be representative of the Japanese economy. The seven categories are (1) 0.01 - 0.02 billion, (2) 0.02 - 0.05 billion, (3) 0.05 - 0.1 billion, (4) 0.1 - 0.5 billion, (5) 0.5 - 1 billion, (6) 1 - 2 billion, (7) over 2 billion. Before FY2010, the border between (4) and (5) was 0.6 billion yen. In the aggregated results, firms are grouped by size into three categories: "small-medium" corporations (1, 2, 3), "medium-sized" corporations (4, 5) and "large" corporations (6, 7). In addition, the Twelfth revision of the J-SIC (November 2007) was enforced on April 1, 2008, and the stratification by industry changed from the first quarter (April - June) of FY 2009. Originally, the sampling was stratified into 43 categories by industry for non-financial corporations.

Table C.1 shows the sampling probabilities for each broad size strata. Due to the high sampling probability among middle to large firms, we are able to construct a panel. The overall response rate is nearly 80% on average and nears 90% for large firms. Table C.2 shows the average annual response rates by year from FY 2005 to FY 2016. Firms are sampled on an annual basis, aligned with standard fiscal period ends for Japanese firms. Firms that are sampled are assigned a unique company identifier that is only for use in the MOF. The disaggregated information and responses at the firm-level do not leave the MoF building and are stored on air-gapped computers.

The BOS asks both qualitative and quantitative questions quarterly. Surveys start in the late of first month in each quarter. Firms are required to answer their forecasts and plans in the middle of the second month in each quarter as a reference date. Qualitative questions ask about the business condition, domestic economic conditions, employment, and others. These questions provide options like "up", "same", "down", and "unknown". For example, "what is your business condition of the current quarter (the next quarter, the one after next) compared with the previous one?" Respondents can answer one of the four choices "up, same, down and unknown."

Table C.1: Survey Sampling

The table below shows the sampling procedure for non-financial and financial corporations. Financial corporations include both banks, insurance companies, and other financial institutions. * means that 60% of overall small corporations selected by quarterly surveys of Financial Statements Statistics for Corporations by Industry must be sampled and the target number in the sample is around 6,000 firms. In addition, when the number of total sampled corporations with capital less than 500 million yen is less than 30 for each stratum, more firms will be additionally sampled to increase the total number in that stratum to 30.

Panel A: Non-Financial Corporations			
Corporation Type	Size (Yen)	Approx. Size (USD)	Sample Probability
Large	$\geq$ 2 billion	$\geq$ 18 million	100%
	1 - 2 billion	9-18 million	50%
Medium-Sized	0.5 - 1 billion	4.5 - 9 million	50%
	0.1 - 0.5 billion	1 - 4.5 million	Remaining
Small-Medium	0.01 - 0.1 billion	0.1 - 1 million	to hit 6,000 firms*
Panel B: Financial Corporations			
Corporation Type	Size (Yen)	Approx. Size (USD)	Sample Probability
Large	$\geq$1 billion	9 million	100%
Medium-Sized	0.5 - 1 billion	4.5 - 9 million	50%
	0.1 - 0.5 billion	1 - 4.5 million	Remaining
Small-Medium	0.01 - 0.1 billion	0.1 - 1 million	0%

Quantitative questions are about realized and expected values of sales, profits, and investment spending. A key feature of the survey is that full fiscal year values are always reported at all horizons. Questions about intra-year forecasts of quantitative items are different according to quarter. Regarding sales and current profit, the surveys in the first quarter (April - June) and second quarter (July - September) asks the first and second half-year forecasts for the current fiscal year. The survey in the third quarter (October - December) has realizations in the first half-year and forecasts for the second half-year for the current fiscal year. The survey in the fourth quarter (January - March) has realizations in the first half-year and forecasts for the second half-year for the current fiscal year as well as the first and second half-year forecasts in the next fiscal year. All surveys ask the semi-annual realizations in the previous fiscal year.

Table C.2: Response Rates

The table below shows the average annual response rate of the BOS, which was administered from FY 2005 (April 2005 to March 2006) to FY 2016 (April 2016 to March 2017).

Fiscal Year	Response Rate (%) (All Firms)	Response Rate (%) (Large Firms)	Fiscal Year	Response Rate (%) (All Firms)	Response Rate (%) (Large Firms)
2005	78.9	88.8	2011	78.6	87
2006	78.7	88.1	2012	79.2	88.1
2007	78.6	86.8	2013	80.2	88.4
2008	79.2	87.3	2014	81.0	88.1
2009	79.4	87.2	2015	80.9	88.6
2010	79.0	87.1	2016	80.9	88.1

C.2 Additional Detailed Summary Statistics

We report additional detailed summary statistics here in the interest of providing more detailed information on the Business Outlook Survey and the Financial Statements survey. These statistics are not used at any point in the paper.

Table C.3: Additional Summary Statistics

The table below reports additional summary statistics of our firm-year panel. Here, we report statistics for the remaining set of investment types reported in the BOS: Land and Software. To reduce the influence of outliers on these summary statistics, we winsorize variables at the 1% level. All numbers are rounded to three significant digits or three decimal points, whichever results in fewer decimal points.

Panel A: Firm Fundamentals						
				Percentile		
Variable	Mean	SD	Skew	25^{th}	50^{th}	75^{th}
Land Investment (mn)	135	583	5.88	0.000	0.000	0.000
Land Investment Plan (mn)	59.7	308	6.34	0.000	0.000	0.000
Software Spending (mn)	190	562	4.89	0.000	15	100
Software Spending Plan (mn)	191	580	4.99	0.000	9	100
Panel B: Constructed Variables						
Land Purchasing Plan (% of Assets)	0.127	0.631	6.05	0.000	0.000	0.000
Land Purchasing Plan Deviations (% of Assets)	0.187	0.964	5.38	0.000	0.000	0.000
Software Spending Plan (% of Assets)	0.867	2.650	5.64	0.000	0.0690	0.557
Software Spending Plan Deviations (% of Assets)	0.003	1.160	0.725	-0.080	0.000	0.083

Table C.4: Additional Summary Statistics on Gross Investment Rates From Financial Statements

The table below reports the summary statistics for investment derived from the Financial Statements (*not BOS)* of our firm-year panel. The total number of firms in our sample is 5,989, of which 2,273 are publicly listed and the rest are private companies. To calculate gross investment rates from financial statements (which are reported annually), we assume a capital depreciation rate of 6%. All numbers are rounded to three decimal points when in percentages and four decimal points in probabilities.

Panel A: Across Years

	(%)	(%)	Capital Sales			Inaction Region		From BOS	
Fiscal Year	$\mathbb{E}[\frac{i}{k}]$	$\mathbb{E}\left[\frac{i}{k} \mid \frac{i}{k} < 0\right]$	$P\left(\frac{i}{k} < 0\right)$	$P\left(\frac{i}{k} < -1\%\right)$	$P\left(\frac{i}{k} < -5\%\right)$	$P\left(\frac{\lvert i\rvert}{k} < 1\%\right)$	$P\left(\frac{\lvert i\rvert}{k} < 5\%\right)$	$P\left(\frac{i}{k} < 1\%\right)$	$P\left(\frac{i}{k} < 5\%\right)$
2005	4.655	-4.261	0.181	0.106	0.040	0.168	0.4801	0.207	0.525
2006	4.491	-3.428	0.196	0.122	0.032	0.166	0.496	0.193	0.504
2007	5.697	-2.771	0.144	0.076	0.021	0.148	0.449	0.192	0.503
2008	5.901	-5.400	0.130	0.072	0.024	0.131	0.430	0.178	0.504
2009	3.970	-3.493	0.160	0.083	0.030	0.191	0.527	0.250	0.608
2010	2.947	-2.794	0.235	0.131	0.029	0.225	0.600	0.224	0.587
2011	3.631	-2.277	0.176	0.084	0.014	0.214	0.565	0.220	0.575
2012	3.885	-3.612	0.187	0.101	0.027	0.193	0.536	0.197	0.537
2013	5.359	-4.079	0.136	0.077	0.024	0.152	0.462	0.182	0.518
2014	5.196	-3.693	0.129	0.069	0.023	0.154	0.466	0.193	0.518
2015	5.414	-4.333	0.123	0.074	0.026	0.133	0.450	0.182	0.503
2016	4.406	-3.508	0.162	0.095	0.029	0.169	0.502	0.183	0.509
Total	4.591	-3.579	0.163	0.091	0.027	0.170	0.496	0.200	0.531

Panel B: Public versus Private Companies

	(%)	(%)	Capital Sales			Inaction Region		From BOS	
	$\mathbb{E}[\frac{i}{k}]$	$\mathbb{E}\left[\frac{i}{k} \mid \frac{i}{k} < 0\right]$	$P\left(\frac{i}{k} < 0\right)$	$P\left(\frac{i}{k} < -1\%\right)$	$P\left(\frac{i}{k} < -5\%\right)$	$P\left(\frac{\lvert i\rvert}{k} < 1\%\right)$	$P\left(\frac{\lvert i\rvert}{k} < 5\%\right)$	$P\left(\frac{i}{k} < 1\%\right)$	$P\left(\frac{i}{k} < 5\%\right)$
Publicly-Listed	4.916	-3.265	0.133	0.079	0.022	0.132	0.484	0.153	0.541
Private Companies	4.232	-3.767	0.189	0.101	0.031	0.202	0.506	0.240	0.523