Communication Friction, Managerial Expectation, and Firm Dynamics *

Haoran Wang †

February 12, 2026

Abstract

Large firms are subject to internal communication frictions. This paper develops a theory to incorporate imperfect within-firm communication among input choices into firm dynamics. Firms are modeled as teams of input managers who make decisions based on segmented information. Using this framework, I provide a novel theory-based test to identify communication friction from managerial expectation data. Applying this methodology to managerial forecast data of US public firms yields strong evidence consistent with the existence of imperfect communication between capital and labor choices. Quantitatively, communication friction distorts capital-labor ratio and helps match the low capital-labor correlation in the US data. Communication friction, along with other information frictions, accounts for a substantial proportion of observed dispersion in capital and labor, and generates a sizable 10% of aggregate TFP loss.

*This paper was previously titled "Implications of Narrow Framing for Firm Dynamics". I am most indebted to my advisors, Luminita Stevens and John Haltiwanger, for their continuous guidance and support on this project. I thank George-Marios Angeletos, Boragan Aruoba, Bence Bardoczy, Andrew Caplin, Pierre De-Leo, Thomas Drechsel, Emel Filiz-Ozbay, Pawel Krolikowski, Chen Lian, Pablo Ottonello, Benjamin Pugsley, Raffaella Sadun, Karthik Sastry, John Shea and Tao Wang for their extremely helpful comments. I acknowledge the financial support from Summer Research Fellowship, offered by the Graduate School of University of Maryland, College Park.

†Department of Economics, University of Maryland, College Park. Email: hw2688@umd.edu

1 Introduction

As the economy becomes more concentrated over time (Autor, Dorn, Katz, Patterson, and Van Reenen (2020); De Loecker, Eeckhout, and Unger (2020)), frictions faced by large firms become more important in aggregate (Gabaix (2011)). Yet imperfect within-firm communication is one of the large-firm frictions that has been abstracted from macroeconomics studies. Despite strong evidence from micro organizational economics showing that communication frictions abound within large organizations (e.g. Mian (2006); Giroud (2013); Impink, Prat, and Sadun (2024)), macro firm dynamics models typically assume that (potentially noisy) information is perfectly shared among the decision-makers within a firm. As a result, little is known about whether and how communication friction should be modeled in macro, and what aggregate consequences it could generate.

This paper studies the macroeconomic implications of imperfect within-firm communication, leveraging the increasingly available managerial expectation data. Theoretically, I develop a model that incorporates internal communication friction into firms' dynamic choices of capital and labor inputs. To capture communication friction, I model the firm as a team of a capital manager and a labor manager, who are allowed to make their own decisions based on different information sets. A key implication of the model is that the forecast error of a manager's decision (e.g. investment) can be predicted by the other manager's past decision (e.g. lag hiring) under communication frictions.

Motivated by this finding, I propose a novel methodology to identify communication friction between capital and labor decisions. The methodology relies on the sign pattern of predictability of investment forecast error by past hiring (other manager's decision) and by past investment (the capital manager's own decision). I show that communication friction is necessary to match the pattern that the two predicatbilities are in opposite direction. In the data of US public firms, I confirm the sign pattern of predictabilities is strongly consistent with the existence of communication friction. In the quantitative exercise, I use this predictability pattern to discipline the parameters of information and communication frictions. When the model is calibrated to match the data, my estimates suggest that information and communication friction explains a larger proportion of capital and labor misallocation than other firm-level frictions, and can generate a sizable 10% aggregate TFP loss.

Theory I develop a model that embeds both imperfect information and imperfect communication into an otherwise standard problem of firms choosing capital and labor dynamically.

Motivated by Marschak and Radner (1972), I model the firm as a team of two managers, one responsible for capital and the other for labor, both of whom choose their own input to jointly maximize the firm's lifetime profit. To capture communication friction, I allow the managers to observe different private signals about the firm's productivity, so that their information sets are non-nested with each other. The capital and labor choices of the firm, together with the belief dynamics of managers, are characterized by the Markov Perfect Bayesian Equilibrium (MPBE) of this dynamic team production game with incomplete information.

The theory allows us to analyze how communication friction distorts the managerial expectation of the firm. I show that communication friction makes investment forecast error, defined as the difference between the actual and one-period-ahead forecast of capital expenditure, dependent upon the gap between actual hiring and the capital manager's nowcast of hiring in the previous period. The intuition for this result is that, under communication friction, the capital manager is not able to observe the private signals of the labor manager, and hence his belief of the labor manager's decision in that period would in general be different from the labor manager's action. This makes it possible for the investment forecast error to be predictable by the labor manager's past hiring decisions, which is the key to the empirical identification of communication friction.

Identifying Communication Friction from Expectation Although the predictability of investment forecast error by past hiring is sensitive to the firm's internal information structure, other behavioral frictions, such as extrapolative beliefs (Angeletos, Huo, and Sastry (2021)), cognitive discounting (Gabaix (2020)) and diagnostic expectation (Bordalo, Gennaioli, Ma, and Shleifer (2020)), are likely to distort the managerial expectation and result in similar patterns. To address these confounding factors, I extend the baseline framework to accommodate these alternative behavioral frictions and consider the predictability of investment error by both past hiring (the other manager's decision) and past investment (the capital manager's own decision). I show that without communication friction, the predictability of investment forecast error by past hiring tends to be in the same direction as the predictability by past investment. Communication friction, on the other hand, makes it possible for them to have opposite signs. Hence, if we observe that investment forecast under-reacts to past hiring and over-reacts to past investment, then communication friction is necessary to match the pattern.

The intuition for this result is that under imperfect communication, the predictability by past hiring and past investment convey different messages to the capital manager. Predictability of investment forecast error by past investment reflects how the capital manager utilizes his own past decision, a piece of information that he knows when he makes the forecast. Absent communication friction, the labor manager's past decision of hiring is based on same information as investment and has similar information content. The predictability by past investment and by past hiring reflect the common frictions or behavioral biases that both managers face and therefore tend to be on the same direction. This symmetry is broken under communication friction: the existence of labor manager's private information makes the capital manager interpret his own past decision and the labor manager's past decision differently, which can lead to opposite directions of predictability.

I apply this methodology to the merged IBES/Compustat dataset, which contains both managerial expectation and balance-sheet outcomes for US public firms. I document that the investment forecast systematically under-reacts to past hiring (i.e. high past hiring predicts large investment forecast error) but over-reacts to past investment (i.e. high past investment predicts low investment forecast error), which is consistent with the existence of communication friction. I show that this pattern cannot be matched by the models without communication friction.

Quantification I then extend the framework to quantify the aggregate effect of communication friction and show that it is quantitatively important and sheds new lights on the role of information frictions on firm dynamics. In the extended model, I consider the following types of frictions: time-to-build, adjustment frictions on both capital and labor, information/behavioral frictions (with imperfect communication + extrapolative beliefs), correlated and uncorrelated distortions, and firm characteristics. I target a variety of moments to identify and discipline the sizes of each forces. In particular, the information friction parameters are disciplined by the predictability of investment forecast errors by past hiring and by past investment, the key fact from the expectation test.

The calibrated model allows me to investigate the role of communication friction among the potential sources of three key features in US public firm data: the low investment-hiring correlation, the low responsiveness of capital and labor adjustments to productivity changes, and the large observed dispersion in marginal revenue products of capital and labor.

Our estimated model suggests that communication friction is a promising source of the low investment-hiring correlation in the US data. Under my parameterization, communication friction alone can generate low, even negative investment-hiring correlation in the US data, while none of the other frictions alone can break the perfect correlation between capital and labor adjustments. This is because the existence of private information breaks the perfect correlation between expectations of different managers. From the perspective of Hsieh and Klenow (2009), these manager-level private signals should be regarded as a factor-specific distortion that alters the capital-labor ratio and is a different type of distortion from firm-level, common information, which affects all the inputs symmetrically. Ignoring communication friction will likely overstate the importance of these factor-specific distortions in firm dynamics.

Our estimated model also suggests that information friction attenuates investment response to productivity more than labor response to productivity. The disproportionate attenuation pattern is also due to communication friction, as the signal noise of the capital manager is estimated to be larger than that of the labor manager. Lastly, our estimates show large aggregate effects of information friction. I find that information friction alone explains about half of capital misallocation (defined as the MRPK dispersion) and 37 percent of labor misallocation (defined as the MRPN dispersion) in the data, which translate to a sizable 10% of aggregate TFP loss. Other within-firm frictions, including adjustment costs, correlated distortions and uncorrelated distortions, explain a much more modest proportion of misallocation and effective loss by themselves.

Related Literature This paper is related to four strands of literature. First, the theoretical part of the paper is related to the literature that introduces incomplete information into macroeconomic models (Woodford (2001), Angeletos and Lian (2016), Angeletos and Lian (2018), Angeletos and Huo (2021)). My paper makes two contributions to this literature. First, the existing literature has been focusing on introducing information asymmetry across households or firms,[1] while my paper applies this modeling technique to study within-firm frictions. Second, this literature is largely theoretical and does not provide empirical evidence for imperfect common knowledge. My paper provides a novel methodology to empirically identify communication friction from the expectation data and shows evidence in support of communication friction.

Second, the empirical methodology I develop to identify communication friction belongs to the empirical literature that investigates the expectation formation process through the lens of predictability of forecast errors (Coibion and Gorodnichenko (2012, 2015)). Most works in this literature focuses on firm's macroeconomic expectations. Some recent works,

[1]An exception is Lian (2021), who studies coordination frictions among multiple decisions of households.

notably Ma, Ropele, Sraer, and Thesmar (2020), Barrero (2022) and Asriyan and Kohlhas (2024), study the distortions in the sales expectation of firms, while my paper focuses on investment forecasts. My paper jointly tests the predictability of investment forecast error by past investment and by past hiring and documents simultaneous under- and over-reaction patterns in investment forecast. This is closely related to Kohlhas and Walther (2021), who documents simultaneous over- and under-reaction of expectation in a different context.

Third, my work contributes to the organizational economics literature (Malenko (2022)) by investigating how the communication friction affect the managerial expectations. Previous works in this literature document evidence for imperfect communication in two ways. One strand of literature (e.g. Dessein et al. (2022); Impink et al. (2024)) utilizes rich internal datasets that involve highly confidential information such as meeting transcripts, emails and hierarchical details, for a limited number of firms. Such evidence is hard to obtain at a large enough scale for quantifying macroeconomic consequences. In contrast, my methodology relies on managerial expectation, which is increasingly available for a wider range of firms.[2] Another strand of literature proxies communication friction by distance or travel time and reports evidence for a wider range of firms (e.g. Mian (2006); Giroud (2013)). My methodology does not rely on geographical variations and can be more flexibly applied to same-location teams with multiple decision-makers.

Lastly, my paper is closely related to the literature that studies the linkage between information friction/uncertainty and misallocation. Previous work, such as David, Hopenhayn, and Venkateswaran (2016) and David and Venkateswaran (2019), quantifies the contribution of information friction to misallocation under two types of internal information structure: either there is a firm-level noisy signal that is commonly known by the managers, or the labor decision is made ex post and only capital decisions are subject to information frictions. My paper generalizes the analysis and can accommodate any given information structure within the firm. Previous works on uncertainty-driven firm dynamics, such as Bloom, Floetotto, Jaimovich, Saporta-Eksten, and Terry (2018), focus on firm-level uncertainty. My paper shows the importance of subordinates' uncertainty in explaining firm dynamics (which echoes the large within-firm variation in management practices documented by Bloom, Brynjolfsson, Foster, Jarmin, Patnaik, Saporta-Eksten, and Van Reenen (2019)) and highlights the difference between information and communication friction (Bloom, Garicano, Sadun, and Van Reenen, 2014) in firm dynamics.

[2] For a recent survey, see Bachmann, Tapa, and van der Klaauw (2023).

Layout The remainder of the paper is organized as follows. Section 2 presents the theory. Section 3 describes the empirical test for within-firm communication friction and applies it to U.S. public firms. Section 4 quantifies the aggregate implications of communication frictions, discussing parameter identification and calibration results. Section 5 concludes.

2 Model

In this section, I describe the model of dynamic input choices under communication friction. Throughout the paper, there is no aggregate uncertainty, so that all the aggregate variables are fixed at their steady-state values in equilibrium and are common knowledge to all the agents in the economy.

Household There is a representative household in the economy, who has standard preference on final-good consumption, discounts future consumption utility by a discount factor β and supplies a fixed amount of labor $\bar{N}$ to the firms at a wage rate W_t. He can save in a risk-free bond at a real interest rate r in each period. The household chooses how much to save and consume subject to a budget constraint. In a stationary equilibrium, consumption is constant over time, and by the Euler equation, we have $\beta(1+r) = 1$. The only role of the household sector is to close the model, and it is not essential to the analysis.

Final-Good Producer There is a final good producer who combines intermediate goods $\{Y_{it}\}$ into final good Y_t by a CES technology

$$Y_t = \left(\int_0^1 (A_{it} Y_{it})^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}} \tag{1}$$

where A_{it} can be interpreted as an idiosyncratic productivity shock on intermediate-good producer i or a demand shock on the product i, and $\epsilon > 1$ is the parameter for elasticity of substitution.

The price of final good is normalized to one. Given price $\{P_{it}\}$, the final good producers choose how much intermediate goods to purchase from intermediate-good producers. This yields a standard demand curve

$$P_{it} = \left(\frac{Y_t}{Y_{it}}\right)^{\frac{1}{\epsilon}} A_{it}^{1-\frac{1}{\epsilon}} \tag{2}$$

Intermediate-Good Producer There is a continuum $i \in [0,1]$ of intermediate-good producers in the economy. Each producer i has a constant-return-to-scale production function to combine capital and labor inputs into intermediate goods:

$$Y_{it} = K_{it}^{\alpha} N_{it}^{1-\alpha} \tag{3}$$

where $\alpha \in (0,1)$ is the parameter for capital share. Each producer i operates in a standard monopolistic competition environment: it internalizes the demand function (2), so that its revenue can be expressed as a decreasing-return-to-scale function of capital and labor, denoted as $R(K_{it}, N_{it}; A_{it})$:

$$R(K_{it}, N_{it}; A_{it}) = P_{it} Y_{it} = Y_t^{\frac{1}{\epsilon}} A_{it}^{1-\frac{1}{\epsilon}} Y_{it}^{1-\frac{1}{\epsilon}} = Y_t^{\frac{1}{\epsilon}} A_{it}^{1-\frac{1}{\epsilon}} K_{it}^{\hat{\alpha}_1} N_{it}^{\hat{\alpha}_2}$$

where $\hat{\alpha}_1 \equiv \alpha(1 - 1/\epsilon)$ and $\hat{\alpha}_2 \equiv (1-\alpha)(1 - 1/\epsilon)$.

Producer i's cost of production includes two parts. First, the firm pays the rental cost of capital and hiring cost of labor. Additionally, the firm faces convex adjustment costs

$$\frac{\xi_k}{2}\left(\frac{I_{it}}{K_{i,t-1}} - \delta\right)^2 K_{i,t-1} \text{ and } \frac{\xi_n}{2}\left(\frac{H_{it}}{N_{i,t-1}}\right)^2 N_{i,t-1}$$

where $I_{it} \equiv K_{it} - (1-\delta)K_{i,t-1}$ is investment and $H_{it} \equiv N_{it} - N_{i,t-1}$ is hiring. ξ_k and ξ_n are parameters that control the slope of marginal adjustment cost for capital and labor respectively. Putting everything together, the cost function $\mathcal{C}(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1})$ is

$$\mathcal{C}(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1}) = R_t K_{it} + W_t N_{it} + \frac{\xi_k}{2}\left(\frac{I_{it}}{K_{i,t-1}} - \delta\right)^2 K_{i,t-1} + \frac{\xi_n}{2}\left(\frac{H_{it}}{N_{i,t-1}}\right)^2 N_{i,t-1}$$

where δ is the deprecation rate, and R_t is the rental rate of capital, which is equal to $r + \delta$ in a stationary equilibrium.

Denote profit function $\Pi(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1}, A_{it}) \equiv R(K_{it}, N_{it}; A_{it}) - \mathcal{C}(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1})$.

Under full information, the firm's input choices are characterized by the Bellman equation

$$V(K_{i,t-1}, N_{i,t-1}; A_{it}) = \max_{K_{it}, N_{it}} \Pi(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1}, A_{it}) + \frac{1}{1+r}\mathbb{E}[V(K_{it}, N_{it}; A_{i,t+1})|A_{it}] \tag{4}$$

To facilitate my later discussions, it is useful to think of the firm as a team of two managers, one in charge of capital decisions (I call him the Capital Manager) and the other in charge of labor decisions (I call her the Labor Manager), who share the common objective of maximizing the firm's lifetime profit. Hence, the capital and labor policy function that solves the *firm's* Bellman equation (4) are equivalent to a Markov Perfect Equilibrium (MPE) of a dynamic team production game (Marschak and Radner (1972), Chapter 7), in which

- Given the labor manager's policy N_{it}, the capital manager's strategy K_{it} is the one that solves the Bellman equation

$$V^k(K_{i,t-1}, N_{i,t-1}; A_{it}) = \max_{K} \Pi(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1}, A_{it}) + \frac{1}{1+r}\mathbb{E}[V^k(K, N_{it}; A_{i,t+1})|A_{it}] \tag{5}$$

- Given the capital manager's policy K_{it}, the capital manager's strategy N_{it} is the one that solves the Bellman equation

$$V^n(K_{i,t-1}, N_{i,t-1}; A_{it}) = \max_{N} \Pi(K_{it}, N_{it}; K_{i,t-1}, N_{i,t-1}, A_{it}) + \frac{1}{1+r}\mathbb{E}[V^n(K_{it}, N; A_{i,t+1})|A_{it}] \tag{6}$$

Equation (5) and (6) are the *manager*-level Bellman equations that characterize the firm's input decisions. This reformulation gives us more flexibility to accommodate incomplete/asymmetric information between managers within the firm and makes it easier for us to analyze a richer internal information structure within the firm.

Information The only source of uncertainty in this simple model is the idiosyncratic firm fundamentals A_{it}. I assume that the log of A_{it}, denoted as a_{it}, follows an AR(1) process

$$a_{it} = \rho a_{i,t-1} + \mu_{it}, \text{ with } \mu_{it} \overset{iid}{\sim} N(0, \sigma_\mu^2) \tag{7}$$

The firm managers know the data generating process in (7) and treat the objective distribution of μ_{it} or a_{it} as the prior. However, the true value of a_{it} is never fully revealed to the firm and its managers.[3] Hence, when making decisions, the capital and labor manager of firm i need to learn dynamically from the Gaussian noisy signals they receive to form expectations about the "latent variable" a_{it} over time using Kalman filters, which I will specify later in detail.

The key friction is that I allow each manager's decision to be dependent on different, non-nested information sets. Formally, let $\mathcal{I}_{it}^k$ denote capital manager's information set in period t, and $\mathcal{I}_{it}^n$ denote the labor manager's information set in period t. In particular, I allow the possibility that $\mathcal{I}_{it}^k \neq \mathcal{I}_{it}^n$ and are not a subset to each other. This captures the notion of imperfect communication within the organization: with imperfect information sharing, the managers' information sets do not perfectly overlap.

I now specify the elements in the managers' information sets $\mathcal{I}_{it}^k$ and $\mathcal{I}_{it}^n$. I assume that there is a piece of firm-level information a_{it}^p, which is a noisy signal of the firm's fundamental a_{it} and is commonly known by the managers (i.e. $a_{it}^p \in \mathcal{I}_{it}^k \cup \mathcal{I}_{it}^n$):

$$a_{it}^p = a_{it} + \epsilon_{it}^p, \text{ with } \epsilon_{it}^p \overset{iid}{\sim} N(0, \sigma_{\epsilon,p}^2) \tag{8}$$

For the non-nested part of the information set, I assume that each manager $m \in \{k, n\}$ is endowed with a private, noisy signal a_{it}^m about a_{it}:

$$a_{it}^m = a_{it} + \epsilon_{it}^m, \text{ with } \epsilon_{it}^m \overset{iid}{\sim} N(0, \sigma_{\epsilon,m}^2) \text{ for } m \in \{k, n\}. \tag{9}$$

The signal noises $\epsilon_{it}^k, \epsilon_{it}^n, \epsilon_{it}^p$ and the innovation in the firm fundamental μ_{it} are independent to each other. In each period, the realization of a_{it}^k and a_{it}^n are private to the capital and labor manager respectively, but the data generating process of the signals, (8) and (9), are commonly known by the managers.

The standard deviations of the signal noises, $(\sigma_{\epsilon,p}, \sigma_{\epsilon,k}, \sigma_{\epsilon,n})$, define the internal information structure of the firm and are the key parameters to pin down in the quantitative analysis. This parametrization gives us a parsimonious way to model the degree of information segmentation within the firm. For example, if we make the private signals to be extremely noisy, i.e. $\sigma_{\epsilon,m} \to \infty$ for $m \in \{k, n\}$, then the managers will put essentially zero weight on private signals when they apply Bayes's rule to update their beliefs, and the firm's behavior is as if the firm is endowed with a public information only and the managers' information set

[3]This is different from the setup in David, Hopenhayn, and Venkateswaran (2016) and David and Venkateswaran (2019), where the lagged firm fundamental $a_{i,t-1}$ is fully revealed to the firm in period t.

perfectly overlaps. On the other hand, if $\sigma_{\epsilon,p} \to \infty$ while $\sigma_{\epsilon,m} > 0$, then the managers will not rely on the firm-level signal to update their beliefs, and it's as if the firm is operating on a non-nested information structure.

Since the realization of signal a_{it}^m is private to manager m and both managers choose their decisions simultaneously, the action that manager m takes in period t is not known by the other manager m', and vice versa. However, I assume that at the end of each period t, the actions taken by the managers in period t, K_{it} and N_{it}, are revealed to both managers and become public information. This assumption allows the managers to use the observed actions in the previous period to update their beliefs about the unknown fundamental a_{it} and the other manager's action in the current period, which simplifies the characterization of the equilibrium concept. Formally, putting everything together, we can express the law of motion for the managers' information sets $\mathcal{I}_{it}^k$ and $\mathcal{I}_{it}^n$ as

$$\mathcal{I}_{it}^k = \mathcal{I}_{i,t-1}^k \cup \{a_{it}^p, a_{it}^k, N_{i,t-1}\} \text{ and } \mathcal{I}_{it}^n = \mathcal{I}_{i,t-1}^n \cup \{a_{it}^p, a_{it}^n, K_{i,t-1}\} \tag{10}$$

Equilibrium Under communication friction, the firm faces a dynamic team production problem with incomplete information. The firm's input decisions can therefore be interpreted as a Markov Perfect Bayesian Equilibrium (MPBE), in which

- Given labor manager's policy function $N_{it}(\cdot)$, capital manager's policy function $K_{it} : \mathcal{I}_{it}^k \to \mathbb{R}^+$ solves Bellman equation

  $$V^k(K_{i,t-1}, N_{i,t-1}; \mathcal{I}_{it}^k) = \max_K \mathbb{E}\left[\Pi(K, N_{it}(s_{it}^n); K_{i,t-1}, N_{i,t-1}, A_{it}) + \beta V^k(K, N_{it}(s_{it}^n); \mathcal{I}_{i,t+1}^k) \middle| \mathcal{I}_{it}^k\right] \tag{11}$$

  where s_{it}^n is the signal realization of the labor manager in period t, and

- Given capital manager's policy function $K_{it}(\cdot)$, labor manager's policy function $N_{it} : \mathcal{I}_{it}^n \to \mathbb{R}^+$ solves Bellman equation

  $$V^n(K_{i,t-1}, N_{i,t-1}; \mathcal{I}_{it}^n) = \max_N E_t\left[\Pi(K_{it}(s_{it}^k), N; K_{i,t-1}, N_{i,t-1}, A_{it}) + \beta V^n(K_{it}(s_{it}^k), N; \mathcal{I}_{i,t+1}^n) \middle| \mathcal{I}_{it}^n\right] \tag{12}$$

  where s_{it}^k is the signal realization of the capital manager in period t, and

- In period t, each manager updates his/her belief about a_{it} using his/her observed signals and the past actions $K_{i,t-1}, N_{i,t-1}$, consistent with the policy function $K_{it}(\cdot), N_{it}(\cdot)$ and the Bayes's rule.

The advantage of using MPBE in (11) and (12) to characterize firm's behavior lies in its flexibility of accommodating essentially any possible internal information structure of the firm. This flexibility makes it a more general model and nests two classes of extensively studied models in the firm dynamics literature. One is the models based on full-information rational expectation (FIRE), as in (4) or (5) and (6) above. They can be regarded as a special case of (11) and (12), with $\mathcal{I}_{it}^k = \mathcal{I}_{it}^n = \{a_{it}, a_{i,t-1}, \cdots, a_{i0}\}$, i.e. the true firm fundamental is always revealed to both managers. Using our parameterization (8) and (9) above, this full-information benchmark maps to the limit $\sigma_{\epsilon,p} \to 0$ and $\sigma_{\epsilon,m} \to \infty$ for $m \in \{k, n\}$. Another class of models are the ones based on common, noisy information, as in David, Hopenhayn, and Venkateswaran (2016).[4] Although it is a useful benchmark to analyze the effect of uncertainty and imperfect information on firm dynamics, it presumes that that imperfectly observed information is frictionlessly shared and communicated within the organization. From the above discussion, this is again a special case of the current framework, with $\mathcal{I}_{it}^k = \mathcal{I}_{it}^n = \{a_{it}^p, \cdots, a_{i0}^p\}$. Using our parameterization (8) and (9) above, this noisy, common information benchmark maps to the limit $\sigma_{\epsilon,p} \in (0, \infty)$ and $\sigma_{\epsilon,m} \to \infty$ for $m \in \{k, n\}$. Hence, the firm dynamics implied by noisy, common information models is also contained in (11) and (12).

Equation (11) and (12) illustrate that the managers face two types of uncertainty when they are making their own decisions. For the capital manager, apart from the fundamental uncertainty about a_{it} or A_{it}, he faces a strategic uncertainty about the action of the labor manager N_{it}, which is measurable to labor manager's information set only. The source of this strategic uncertainty is the fact that information is not perfectly shared within the firm, so that there is labor-specific information that is only known to the labor manager and not to the capital manager. This strategic uncertainty affects capital manager's decision because labor manager's action affects both the marginal revenue product of capital and the continuation value of the capital manager, and therefore capital manager has an incentive to form expectation about N_{it} to guide his own decision about K_{it}.

Log-Quadratic Approximation The third requirement of an MPBE requires that the belief dynamics of both managers are consistent with the policy function and Bayes's rule.

[4]More precisely, David et al. (2016) studies two extreme cases, where (1) imperfectly observed information is perfectly shared among capital and labor managers, and (2) labor decisions are made *ex post* and frictionlessly, and only capital decisions are subject to information friction (which is also the main specification of David and Venkateswaran (2019)). The latter case is also a special case of the current framework, with the capital manager's information set is a subset of the labor manager's information set. My paper can be regarded as providing a smooth version of information friction between these two extreme cases.

It is a difficult task to analytically characterize or compute the evolution of beliefs since it is an infinite-dimensional object.

Here, I get around this technical challenge by doing a Woodford (2003)-type log-quadratic approximation on the profit function with respect to $(K_{it}, N_{it}, K_{i,t-1}, N_{i,t-1}, A_{it})$ around their frictionless steady-state values.[5] I can therefore express the log-quadratic approximated profit function π_{it} as a quadratic function in terms of the log deviations of capital, labor and firm fundamental from their own frictionless steady-state value. Denote these log-deviation values as k_{it}, n_{it} and a_{it} respectively, and define investment (rate) as $\iota_{it} \equiv k_{it} - k_{i,t-1}$ and hiring (rate) as $h_{it} \equiv n_{it} - n_{i,t-1}$. The approximated profit function can be written as

$$\pi(\iota_{it}, h_{it}; x_{it}) = x_{it}' P x_{it} + x_{it}' Q \iota_{it} + x_{it}' R h_{it} + H_\iota \iota_{it}^2 + H_{\iota h} \iota_{it} h_{it} + H_h h_{it}^2 \tag{13}$$

where $x_{it} = [k_{i,t-1}, n_{i,t-1}, a_{it}]'$ is the state vector of the firm with a law of motion

$$x_{it} = A x_{i,t-1} + B \iota_{it} + C h_{it} + D \mu_{it} \tag{14}$$

and matrices $A, B, C, D, P, Q, R, H_\iota, H_h$ and $H_{\iota h}$ are constants imputed from the frictionless steady-state values and model parameters. The derivations of (13) are left to Appendix A.1.

Characterization of Linear MPBE The log-quadratic approximation of profit function allows us to characterize a linear MPBE analytically.

Plugging (13) and (14) back into the manager-level Bellman equations (11) and (12) transforms the original problem into a linear-quadratic (LQ) control problem, with investment ι_{it} and hiring h_{it} as control variables and x_{it} as the state vector. An immediate implication of the LQ control is that the policy function is linear and certainty equivalence holds:

Lemma 1. *With the log-quadratic-approximated profit function (13) and the law of motion for state (14), we have*

1. *The policy function for the full-information MPE in (5) and (6) takes the form*

$$\begin{aligned} \iota_{it} &= F_\iota x_{it} = F_\iota^k k_{i,t-1} + F_\iota^n n_{i,t-1} + F_\iota^a a_{it} \\ h_{it} &= F_h x_{it} = F_h^k k_{i,t-1} + F_h^n n_{i,t-1} + F_h^a a_{it} \end{aligned}$$

[5] By frictionless, I mean the case where there is no information friction and adjustment friction, i.e. a_{it} is fully revealed to the firm, and $\xi_k = \xi_n = 0$.

where the policy matrix F_ι, F_h can be solved by iterating forward a pair of Ricatti equations;

2. *The policy function for (11) and (12) under communication friction takes the form*

$$\begin{aligned} \iota_{it} &= F_\iota \mathbb{E}[x_{it}|\mathcal{I}_{it}^k] = F_\iota^k k_{i,t-1} + F_\iota^n n_{i,t-1} + F_\iota^a \hat{a}_{it}^k && (15) \\ h_{it} &= F_h \mathbb{E}[x_{it}|\mathcal{I}_{it}^n] = F_h^k k_{i,t-1} + F_h^n n_{i,t-1} + F_h^a \hat{a}_{it}^n && (16) \end{aligned}$$

where F_ι, F_h are the same policy matrix as in the full-information MPE, and $\hat{a}_{it}^k \equiv \mathbb{E}[a_{it}|\mathcal{I}_{it}^k]$ and $\hat{a}_{it}^n \equiv \mathbb{E}[a_{it}|\mathcal{I}_{it}^n]$ are the posterior mean of capital and labor manager's first-order belief on a_{it}.

Proof. See Appendix A.3. □

The linearity of (15) and (16), together with the Gaussian structure of μ_{it} and signal noises, significantly simplifies the characterization of the belief updating process for two reasons. First, the Gaussian shocks and noises imply that tracking the first and second moment is sufficient for the characterization of belief dynamics, which significantly lowers the dimensions needed for the computation. Moreover, the linearity of policy functions makes actions jointly Gaussian with firm fundamentals and signals, and the managers can simply treat the other manager's lagged action as a noisy signal of the firm fundamentals and learn from it in the same manner as from other signals. For instance, in period t, with (16), the capital manager forms belief about the labor manager's action h_{it} by

$$\hat{h}_{it}^k \equiv \mathbb{E}[h_{it}|\mathcal{I}_{it}^k] = F_h^k k_{i,t-1} + F_h^n n_{i,t-1} + F_h^a \mathbb{E}[\hat{a}_{it}^n|\mathcal{I}_{it}^k] \tag{17}$$

Equation (17) is an important equation. It sends us two important messages. First, it tells us that the capital manager's belief about the labor manager's action is linear on his *high-order belief* about labor manager's nowcast on a_{it}, and therefore it is necessary for him to track the evolution of both first-order beliefs $a_{it}|\mathcal{I}_{it}^k \sim N(\hat{a}_{it}^k, \hat{\Sigma}_t^k)$ and high-order beliefs $\hat{a}_{it}^n|\mathcal{I}_{it}^k \sim N(\hat{a}_{it}^{(k,n)}, \hat{\Sigma}_t^{(k,n)})$ along the belief updating process. Second, it tells us that the elasticity of hiring to firm fundamental, F_h^a, directly enters the belief formation process by affecting the nowcast error $h_{it} - \hat{h}_{it}^k$. This channel of policy matrix affecting beliefs will be in effect if and only if there is information asymmetry between the managers: if the information is symmetric within the firm, the capital manager knows that the labor manager shares the same information with him, which allows him to know the labor manager's action exactly, i.e. $h_{it} = \hat{h}_{it}^k$. Similarly, the labor manager updates her beliefs using capital manager's past action and policy function (15) in a MPBE. For that purpose, she will need to track her first-order belief $a_{it}|\mathcal{I}^n_{it} \sim N(\hat{a}^n_{it}, \hat{\Sigma}^n_t)$ and high-order belief $\hat{a}^k_{it}|\mathcal{I}^n_{it} \sim N(\hat{a}^{(n,k)}_{it}, \hat{\Sigma}^{(n,k)}_t)$ dynamically.

The following lemma characterizes the evolution of the posterior mean $\hat{a}^k_{it}, \hat{a}^{(k,n)}_{it}$ and $\hat{a}^n_{it}, \hat{a}^{(n.k)}_{it}$ for a given pair of policy matrices (F_ι, F_h).

Lemma 2. *For a given pair of policy matrices F_ι and F_h, the posterior first-order and high-order mean $\hat{a}^k_{it}, \hat{a}^{(k,n)}_{it}$, $\hat{a}^n_{it}, \hat{a}^{(n.k)}_{it}$ evolves by*

$$\begin{aligned}
\hat{a}^k_{it} &= \hat{a}^{(k,n)}_{it} = (\hat{a}^k_{it})^- + G^k_t \left(z^k_{it} - H(\hat{a}^k_{it})^-\right) && (18)\\
\hat{a}^n_{it} &= \hat{a}^{(n,k)}_{it} = (\hat{a}^n_{it})^- + G^n_t \left(z^n_{it} - H(\hat{a}^n_{it})^-\right) && (19)
\end{aligned}$$

where $z^k_{it} = [a^p_{it}, a^k_{it}]'$, $z^n_{it} = [a^p_{it}, a^n_{it}]'$, $H = [1, 1]'$, with the Kalman gains being

$$G^k_t = (\hat{\Sigma}^k_t)^- H' \left[H(\hat{\Sigma}^k_t)^- H' + W^k\right]^{-1} \text{ and } G^n_t = (\hat{\Sigma}^n_t)^- H' \left[H(\hat{\Sigma}^n_t)^- H' + W^n\right]^{-1}$$

and the pre-estimates $(\hat{a}^k_{it})^-, (\hat{a}^n_{it})^-$ are given by

$$\begin{aligned}
(\hat{a}^k_{it})^- &= \rho(\hat{a}^k_{i,t-1})^+ = \rho\left[\hat{a}^k_{i,t-1} + J^k_{t-1}(h_{i,t-1} - \hat{h}^k_{i,t-1})\right] && (20)\\
(\hat{a}^n_{it})^- &= \rho(\hat{a}^n_{i,t-1})^+ = \rho\left[\hat{a}^n_{i,t-1} + J^n_{t-1}(\iota_{i,t-1} - \hat{\iota}^n_{i,t-1})\right] && (21)
\end{aligned}$$

with the Kalman gain being

$$J^k_{t-1} = (\hat{\Sigma}^k_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1}) \left[\hat{\Sigma}^{(k,n)}_{t-1} F^a_h\right]^{-1} \text{ and } J^n_{t-1} = (\hat{\Sigma}^n_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1}) \left[\hat{\Sigma}^{(k,n)}_{t-1} F^a_\iota\right]^{-1}$$

The evolution of the first-order and higher order uncertainty $(\hat{\Sigma}^k_t, \hat{\Sigma}^n_t, \hat{\Sigma}^{(k,n)}_t)$ is defined by a triplet of Riccati equations (A6), (A7) and (A8) in Appendix A.2.

Proof. See Appendix A.2. ◻

In Lemma 2, the belief updating from period $t-1$ to t is done in two steps. In the first step, the managers observe the other manager's action in the previous period, $h_{i,t-1}$ and $\iota_{i,t-1}$, and use them to form a pre-estimate $(\hat{a}^k_{it})^-$ and $(\hat{a}^n_{it})^-$, according to (20) and (21). In the second step, the managers use the signal realization in period t to update their beliefs on top of these pre-estimates, following equation (18) and (19). Note that the first step of belief updating exists if and only if the managers' information sets are non-nested. Absent asymmetric information, there would be no strategic uncertainty between managers, so that

$h_{i,t-1} = \hat{h}^k_{i,t-1}$ and the equation (20) degenerates to $(\hat{a}^k_{it})^- = \rho \hat{a}^k_{i,t-1}$. In that case, the labor manager's action $h_{i,t-1}$ is already contained in the capital manager's information set in period $t-1$, and hence revealing $h_{i,t-1}$ to the capital manager doesn't give him additional information about the unobserved firm fundamental.

Now we have all the ingredients to characterize a linear MPBE associated with (13) and (14).

Lemma 3. *A linear Markov Perfect Bayesian Equilibrium is a set of policy (F_ι, F_h) and beliefs $(\hat{a}^k_{it}, \hat{\Sigma}^k_t)$, $(\hat{a}^n_{it}, , \hat{\Sigma}^n_t)$, $(\hat{a}^{(k,n)}_{it}, \hat{\Sigma}^{(k,n)}_t)$, and $(\hat{a}^{(n,k)}_{it}, \hat{\Sigma}^{(n,k)}_t)$ such that*

- *Beliefs are consistent with strategies: Given policy (F_ι, F_h), the capital and labor managers' update their beliefs $(\hat{a}^k_{it}, \hat{\Sigma}^k_t)$, $(\hat{a}^n_{it}, \hat{\Sigma}^n_t)$, $(\hat{a}^{(k,n)}_{it}, \hat{\Sigma}^{(k,n)}_t)$, and $(\hat{a}^{(n,k)}_{it}, \hat{\Sigma}^{(n,k)}_t)$ by the Bayes rule, according to Equations (18) to (19) in Lemma 2;*
- *Strategies are consistent with beliefs: Given the beliefs $(\hat{a}^k_{it}, \hat{\Sigma}^k_t)$, $(\hat{a}^n_{it}, , \hat{\Sigma}^n_t)$, $(\hat{a}^{(k,n)}_{it}, \hat{\Sigma}^{(k,n)}_t)$, and $(\hat{a}^{(n,k)}_{it}, \hat{\Sigma}^{(n,k)}_t)$, the policy matrices (F_ι, F_h) solves the Bellman equations (5) and (6), and the policy function for investment and hiring satisfy (15) and (16), as in Lemma 1.*

The linear MPBE in Lemma 3 implies a law of motion for the joint distribution of action, expectation and state, namely, the vector $\equiv [k_{it}, n_{it}, \hat{a}^k_{it}, \hat{a}^n_{it}, a_{it}]'$. A stationary equilibrium requires this distribution to be at its stationary distribution. The computation of this stationary distribution is essential to the quantitative analysis in Section 4 and the details of computation are left to Appendix **??**.

Communication Friction and Investment Forecast Error The key object that maps the theory to data is the investment forecast error, defined as the difference between actual capital investment ι_{it} and the capital manager's one-period-ahead forecast of capital investment, $\mathbb{F}^k_{i,t-1}[\iota_{it}]$. Note that the expectation operator $\mathbb{F}^k_{i,t-1}$ is allowed to deviate from rational expectation, but for now let's assume that $\mathbb{F}^k_{i,t-1}$ is the same as rational expectation $\mathbb{E}^k_{i,t-1} \equiv \mathbb{E}[\cdot|\mathcal{I}^k_{i,t-1}]$, in the sense that the capital manager updates his beliefs following the Bayes's rule with correct perception of model parameters (such as the persistence of productivity ρ). From the linear investment decision rule (15) and the belief updating equation (18) and (20) for the capital manager, we have the following proposition that characterize the structure of investment forecast error:

Proposition 1. *In a MPBE defined in Lemma 3, the investment forecast error of firm i takes the form*

$$\iota_{it} - \mathbb{E}^{k}_{i,t-1}[\iota_{it}] = \beta_1(a_{i,t-1} - \hat{a}^{k}_{i,t-1}) + \beta_2(h_{i,t-1} - \hat{h}^{k}_{i,t-1}) + \beta_3(H\mu_{it} + w^{k}_{it}) \tag{22}$$

where $\beta_1, \beta_2 \in \mathbb{R}$ and $\beta_3 \in \mathbb{R}^{1\times 2}$ are constants in terms of the policy matrices F_ι, F_h as in Lemma 1 and Kalman gain matrices G^k, G^n, J^k, J^n in Lemma 2.

Proof. See Appendix A.4 □

Proposition 1 shows that the investment forecast error has three components. The first term captures the fundamental uncertainty that the capital manager faces. It is the gap between actual productivity $a_{i,t-1}$ and the capital manager's expectation in period $t-1$ about the productivity in that period. As long as the capital manager is uncertain about the true productivity in period $t-1$, this gap will not be zero. The third term $\beta_3(H\mu_{it} + w^{k}_{it})$ collects all the time-t productivity innovations and signal noise, which are orthogonal to information on or before period $t-1$.

The second term, $\beta_2(h_{i,t-1} - \hat{h}^{k}_{i,t-1})$, speaks directly to the communication friction. It is the gap between actual hiring in period $t-1$ and the capital manager's expectation in period $t-1$ about $h_{i,t-1}$. Absent communication friction, this term will always equal to 0, as the capital manager knows that the labor manager's decision is based on the same information as his, so that he has no problem of knowing the labor manager's decision in $t-1$. His expectation $\hat{h}^{k}_{i,t-1}$ will therefore be exactly the same as $h_{i,t-1}$. If communication is imperfect within the organization, the capital manager does not observe the exact $h_{i,t-1}$ in period $t-1$ because of the labor manager's private information, and as a result, his belief about the labor manager's action in that period is generally not the same as what the labor manager's actual decision, making this strategic uncertainty contribute to the investment forecast error.

Summary and Extension In summary, my model provides a convenient way to build communication friction in the dynamic choice of capital and labor in presence of adjustment frictions. The main result is Proposition 1, which establishes the connection between communication friction and investment forecast error. Under communication friction, the capital manager does not know exactly the labor manager's private information and hence her decision, and this notion of strategic uncertainty contributes to the investment forecast error.

The decomposition in (22) motivates the empirical strategy of identifying communication friction in the next section.

Before proceeding, I shall discuss an important extension of the model. So far, I have been assuming that the beliefs are rational, in the sense that the managers update their beliefs following the Bayes's rule, and that the managers have the correct perception about the model parameters, such as discount factor β and persistence of productivity ρ. It is easy to incorporate other types of behavioral frictions into the existing framework. For instance, if we replace the Bayesian Kalman gains G^k, G^n, J^k, J^n in equation (18)-(21) in Lemma 2 with some arbitrary weights, we can nest the mechanisms with non-Bayesian belief updating, such as diagnostic expectation (Bordalo et al. (2020)). An important behavioral distortion is extrapolative belief as in Angeletos et al. (2021), which allows the managers' perceived persistence of productivity shock $\hat{\rho}$ to be different from the true persistence ρ. Let $\lambda \equiv \frac{\hat{\rho}}{\rho}$ parameterize the degree of extrapolation, so that $\lambda > 1$ captures over-extrapolation (i.e. the manager believes the productivity shock is more persistent than the truth) and $\lambda < 1$ captures under-extrapolation. The following corollary provides a modification to Proposition 1 after incorporating extrapolative beliefs:

Corollary 1. *Under extrapolative belief, the investment forecast error of firm i takes the form*

$$\iota_{it}-\mathbb{F}^k_{i,t-1}[\iota_{it}] = \beta_1(a_{i,t-1}-\mathbb{F}^k_{i,t-1}[a_{i,t-1}])+\beta_2(h_{i,t-1}-\mathbb{F}^k_{i,t-1}[h_{i,t-1}])+\beta_\lambda\mathbb{F}^k_{i,t-1}[a_{i,t-1}]+\beta_3(H\mu_{it}+w^k_{it}) \tag{23}$$

where the coefficient $\beta_\lambda = 0$ if and only if $\lambda = 1$, i.e. there is no over- or under-extrapolation in manager's beliefs.

Compared with (22), there is an extra term, $\beta_\lambda\mathbb{F}^k_{i,t-1}[a_{i,t-1}]$, in the investment forecast error under extrapolative belief. It turns out that this term creates an important confounding factor to the identification of communication friction, which relies heavily on the predictability of investment forecast error.

3 Identifying Communication Friction

In this section, I describe the empirical strategy of identifying communication friction from expectation data and implement the methodology to U.S. public firms.

3.1 Identification Strategy

Equation (22) motivates the empirical strategy of identifying communication friction from the predictability of investment forecast error. If we regress investment forecast error on lagged hiring and a proxy of the fundamental uncertainty (the first term on the right-hand side of (22)), we should not expect the lagged hiring to help predict the investment forecast error if communication is frictionless within the firm. Conversely, if the investment forecast error can be predicted by lagged hiring with the control for fundamental uncertainty, then communication friction is necessary to match the pattern.

Since the productivity expectations are not available in the data, we use the capital manager's sales nowcast error $y_{i,t-1} - \hat{y}^k_{i,t-1}$ to replace $a_{i,t-1} - \hat{a}^k_{i,t-1}$ in (22). We know that the logged sales $y_{i,t-1}$ satisfy

$$y_{i,t-1} = (1 - 1/\epsilon)a_{i,t-1} + \hat{\alpha}_1 k_{i,t-1} + \hat{\alpha}_2 n_{i,t-1}$$

and therefore the sales nowcast error in period $t-1$ can be written as

$$y_{i,t-1} - \hat{y}^k_{i,t-1} = (1 - 1/\epsilon)(a_{i,t-1} - \hat{a}^k_{i,t-1}) + \hat{\alpha}_2(h_{i,t-1} - \hat{h}^k_{i,t-1})$$

i.e. it is a linear combination of the fundamental nowcast error and the hiring nowcast error. Plug it into (22), we get a model-implied relationship between investment forecast error, lag fundamental nowcast error, lag fundamental hiring nowcast error, and time-t innovations/noises:

$$\iota_{it} - \mathbb{E}^k_{i,t-1}[\iota_{it}] = \beta_h(h_{i,t-1} - \hat{h}^k_{i,t-1}) + \beta_y(y_{i,t-1} - \hat{y}^k_{i,t-1}) + \nu_{it} \tag{24}$$

Equation (24) gives us a model-implied reduced-form relationship between $\iota_{it} - \mathbb{E}^k_{i,t-1}[\iota_{it}]$, h_{it} and $y_{i,t-1} - \hat{y}^k_{i,t-1}$, all of which are available or can be proxied from the firm-level expectation and outcome data. I start by considering the following regression

$$\iota_{it} - \mathbb{E}^k_{i,t-1}[\iota_{it}] = \tilde{\beta}_h h_{i,t-1} + \beta_y(y_{i,t-1} - \hat{y}^k_{i,t-1}) + \nu_{it} \tag{25}$$

The coefficient of interest is $\tilde{\beta}_h$. Colloquially, from the lens of the model-implied reduced-form equation (24), the regression coefficient $\tilde{\beta}_h$ picks up both the direct effect of $h_{i,t-1}$ on the investment forecast error and the indirect effect of $h_{i,t-1}$ via the expectation $\hat{h}^k_{i,t-1}$, holding

$y_{i,t-1} - \hat{y}^k_{i,t-1}$ fixed:

$$\tilde{\beta}_h = \frac{\partial(\iota_{it} - \mathbb{E}^k_{t-1}[\iota_{it}])}{\partial h_{it}}\bigg|_{y_{i,t-1}\,-\,\hat{y}^k_{i,t-1}\text{ is fixed}} = \beta_h \left(1 - \underbrace{\frac{\partial \hat{h}^k_{i,t-1}}{\partial h_{i,t-1}}}_{\equiv \zeta_h}\right) \tag{26}$$

The following corollary is immediate.

Corollary 2. *With rational expectation, regression (25) tests whether capital and labor decisions are based on the same information, as long as $\beta_h \neq 0$ in Equation (26):*

- *If the investment and hiring decisions are based upon same information in period $t-1$, then there is no strategic uncertainty for the capital manager, so that $\hat{h}^k_{i,t-1} = h_{i,t-1}$ and hence $\zeta_h = 1$ and $\tilde{\beta}_h = 0$.*
- *If the investment and hiring decisions are based on different, non-nested information in period $t-1$, then $\frac{\partial \hat{h}^k_{i,t-1}}{\partial h_{i,t-1}} \neq 1$ since the capital manager's expectation about the labor manager's action in period $t-1$ is not perfect, and, as a result $\zeta_h \neq 1$ and $\tilde{\beta}_h \neq 0$.*

Hence, under rational expectation, if we are able to reject the null hypothesis that $\tilde{\beta}_h = 0$, we can conclude that the investment and hiring decisions are made under different, non-nested information within the firm.

However, there might be reasons other than communication friction that makes $\tilde{\beta}_h \neq 0$. For instance, if the managers have extrapolative beliefs, then the extrapolative term $\beta_\lambda \mathbb{F}^k_{i,t-1}[a_{i,t-1}]$ in equation (23) will make $\tilde{\beta}_h$ in regression (25) nonzero in general, even under perfect communication. Hence, with other behavioral distortions as confounding factors, a nonzero regression coefficient $\tilde{\beta}_h$ alone cannot identify communication friction.

To address this confounding factor, I consider an alternative regression

$$\iota_{it} - \mathbb{E}^k_{i,t-1}[\iota_{it}] = \tilde{\beta}_\iota \iota_{i,t-1} + \beta_y(y_{i,t-1} - \hat{y}^k_{i,t-1}) + \nu_{it} \tag{27}$$

with $\tilde{\beta}_\iota$ being the coefficient of interest.

It turns out the sign patterns of $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ identify communication frictions from other types of behavioral distortions. To build some intuition, let's revisit our theoretical results to see the pattern $\tilde{\beta}_\iota$ under different frictions. From equation (22) in Proposition 1, we know that under rational expectation, the regression coefficient $\tilde{\beta}_\iota = 0$ since the investment forecast error does not depend on lag investment, no matter whether there is communication friction. From equation (23) in Corollary 1, we can see that since the extrapolative bias term $\beta_\lambda \mathbb{F}^k_{i,t-1}[a_{i,t-1}]$ is correlated with both lag hiring and lag investment, the regressions coefficients $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ are both nonzero, even under perfect communication. Furthermore, I can show that without communication friction, the sign of $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ tend to be the same. The following proposition characterize the sign patterns of $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ in a simplified version without adjustment cost:

Proposition 2. *If there is neither adjustment cost ($\xi_k = \xi_n = 0$) nor communication friction, the signs of $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ are always the same as the sign of $\lambda - 1$.*

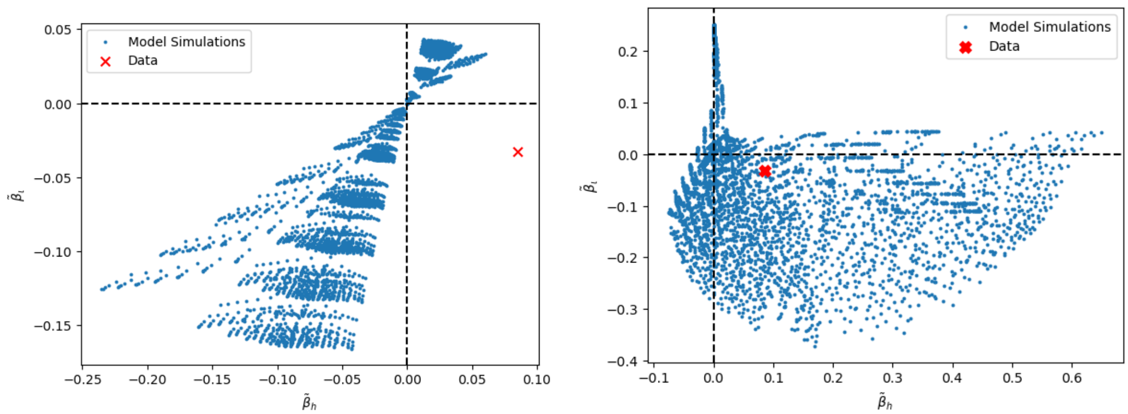
Proof. See Appendix A.5. □

The intuition of Proposition 2 is straightforward. Under perfect communication, both the investment $\iota_{i,t-1}$ and hiring $h_{i,t-1}$ are known by the capital manager in period $t-1$; additionally, since capital and labor are complements, the past investment and past hiring convey similar message about the productivity to the capital manager by themselves. As a result, the belief extrapolation of the capital manager, which generates the predictability by past investment and hiring, will bias the capital manager's interpretation of investment and hiring towards the same direction.

Communication friction makes it possible to break the same-sign pattern of predictability. Consider the case where the capital manager's belief is under-extrapolative, i.e. $\lambda < 1$. According to Proposition 2, absent communication friction, we expect the coefficients $\tilde{\beta}_\iota$ and $\tilde{\beta}_h$ to be both negative, and investment forecast over-reacts to both past hiring and past investment. The communication friction provides an additional channel to reverse the overreaction from extrapolative belief: due to labor manager's private information, the capital manager cannot distinguish whether a large hiring in the past means an increase in productivity or a favorable realization of labor manager's private signal. Therefore, his expectation responds sluggishly and under-reacts to variations in labor. If the force of under-reaction due to communication friction dominates that of over-reaction due to under-extrapolative beliefs, the sign of $\tilde{\beta}_h$ will be flipped to positive, while the negative sign of $\tilde{\beta}_\iota$ remains unchanged.

Note that the analytical result in Proposition 2 assumes no adjustment costs. With adjustment frictions, the regression coefficients can no longer be solved analytically. The left panel

Figure 1: $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ with Adjustment Cost, Common Noisy Info & Extrapolation



Note: In the left panel, blue dots are plotted from 4,096 model simulations with parameters $(\xi_k, \xi_n, \sigma_{\epsilon,p}, \lambda) \in [0, 1] \times [0, 1] \times [0.1, 5] \times [0.8, 1.07]$, equally spaced on each dimension. In each simulation, I simulate a panel of 1,000 firms for 500 periods and run the regressions (25) and (27). In the right panel, blue dots are plotted from 4.096 model simulations with parameters $(\sigma_{\epsilon,p}, \sigma_{\epsilon,k}, \sigma_{\epsilon,n}, \lambda) \in [0.1, 5] \times [0.1, 5] \times [0.1, 5] \times [0.8, 1.07]$, with adjustment cost parameters $\xi_k = \xi_n = 0.1$.

of Figure 1 plots model predictions on $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ by simulating the model with adjustment cost, common noisy information and extrapolative beliefs. We can see that all of the 4,096 simulations lie in the first and third quadrant, with $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ having the same signs. This reaffirms the point that opposite signs in $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ are unlikely to be reconciled with models under perfect communication. For comparison, the right panel of Figure 1 evokes communication friction and plots the model-implied regression coefficients $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$. We can see that most of the simulated coefficients $(\tilde{\beta}_h, \tilde{\beta}_\iota)$ fall on the fourth quadrant with $\tilde{\beta}_h > 0$ and $\tilde{\beta}_\iota < 0$, the quadrant where the US data lies at. It's much easier for models with communication frictions to match the opposite-sign pattern.

In summary, to identify communication friction, I propose to run both regression (25) and regression (27) and test if the estimated coefficients $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ have the same or opposite sign. An opposite sign of predictability by past actions is difficult to be reconciled by models without communication friction and indicative of the existence of imperfect communication between capital and labor decisions.

3.2 Application to Data

I now apply the test (25) and (27) to a merged firm-level expectation-outcome dataset for US public firms.

Dataset For the expectation data, I use the IBES Guidance dataset, which contains all the guidance activities (i.e. the firm's public announcement on its own expectations about sales, capital expenditure, earnings per share, among many others, for the current fiscal year) of US public firms since 1992. These guidances can be regarded as the self-reported expectation of the firm on its own financial performances. Since all measurements in the guidance are financial indicators, I assume that the expectations posted in IBES Guidance mainly reflect the capital manager's view about the firm's performance.

According to (25), the firm-level expectation that we are interested in is (1) the forecast of investment $\mathbb{E}^k_{i,t-1}[\iota_{it}]$ and (2) the nowcast of sales in period $t-1$, $\mathbb{E}^k_{i,t-1}[y_{i,t-1}]$. These two metrics can be proxied by the sales guidance and capital expenditure[6] guidance, both of which are widely covered in the IBES dataset as they are the most frequently reported guidances by the US public firms. However, since guidance activities are self-reported and can happen in any time of the fiscal year, it is difficult to get a notion for the time horizon of the forecast from the announcement date. Here, I restrict my attention to the firms that have multiple guidance activities within the same year, and in that given year, I take the first posted capital expenditure guidance as a proxy for the one-period-ahead forecast of investment $\mathbb{E}^k_{i,t-1}[\iota_{it}]$, and I take the last posted sales guidance as a proxy for the nowcast of sales $\mathbb{E}^k_{i,t-1}[y_{i,t-1}]$. I then merge the IBES Guidance dataset to CRSP/Compustat Fundamental Annual and get a firm-level dataset with both expectations and realized outcomes. The details of data cleaning and sample construction are laid out in Appendix B.1.

		CRSP-Compustat (2000-2024)		
Variables	Units	Mean	Median	No. Obs
Sales	Millions USD	3512.01	366.11	81937
Capital Stock (PPEGT)	Millions USD	2464.81	140.32	78478
Capital Expenditure	Millions USD	210.11	11.59	81937
Employment	Thousands	10.78	1.18	81937
		IBES-Compustat (2000-2024)		
Variables	Units	Mean	Median	No. Obs
Sales	Millions USD	6544.89	1811.80	8455
Capital Stock (PPEGT)	Millions USD	3491.86	771.47	8450
Capital Expenditure	Millions USD	281.56	64.29	8455
Employment	Thousands	22.45	6.90	8455

Table 1: Summary Statistics (2000-2024)

[6]Here, I regard capital expenditure (CPX) as a proxy for investment, although it is not exactly the same as ι_{it} in the model.

The resulting IBES-Compustat merged dataset spans over 2000 and 2024 with 9,102 firm-year observations. Table B1 shows that the IBES-Compustat merged dataset is a sample for gigantic firms in the US economy: on average, firms in the merged sample is about two times larger than those in the CRSP-Compustat dataset in terms of sales and employment. Although the merged sample covers only 11% of total observations in the CRSP-Compustat Fundamental Annual Dataset between 2000 and 2024, the total sales of firms in the merged IBES-Compustat sample take up about 20% of US GDP in a sample year.

For each firm-year observation in the IBES-Compustat merged sample, I take the realized sales and capital expenditure from Compustat and impute the forecast/nowcast error as the arc-percent difference between the predicted value and the actual value, as in Davis and Haltiwanger (1992) and Bloom, Codreanu, and Fletcher (2025):

$$\text{CapEX Forecast Error}_{it} = \frac{\text{CapEX}_{it} - \mathbb{E}^k_{i,t-1}[\text{CapEX}_{it}]}{1/2(\text{CapEX}_{it} + \mathbb{E}^k_{i,t-1}[\text{CapEX}_{it}])}$$

and

$$\text{Sales Nowcast Error}_{it} = \frac{\text{Sales}_{it} - \mathbb{E}^k_{i,t}[\text{Sales}_{it}]}{1/2(\text{Sales}_{it} + \mathbb{E}^k_{i,t}[\text{Sales}_{it}])}$$

The arc-difference definition of forecast/nowcast errors has two advantages over the log difference. First, it is bounded between $[-2, 2]$, so that outliers in the realized values and expectations would not affect the forecast/nowcast error and the regression results by much. Additionally, it avoids the problems of taking logs when the actual values or expectations are close to 0. This may not be a big issue for the large-firm samples like IBES or Compustat, and I verify that the regression results are robust to alternative definitions of forecast/nowcast errors (e.g. the log difference between predicted values and actual values).

Regression Results Table 2 tabulates the regression result for the joint test of (25) and (27). The first column shows the result for regression (25). We can see that $\tilde{\beta}_h$ is positive and statistically significant, which implies that the investment expectation of US public firms systematically under-react to their own past hiring. The second column shows the result for regression (27). We can see that $\tilde{\beta}_\iota$ is negative and statistically significant, which implies that the investment expectation of US public firms systematically over-react to their own past investment. The third column considers including both lagged hiring and lagged investment in the regression, and these multivariate estimates are similar to the separate estimates in the first two columns. It shows that the correlation between past investment and past hiring does not bias the separate regression results in the first two columns.

Table 2: Predictability of Investment Forecast Error by Past Hiring and Investment

	Dependent variable:		
	$\iota_t - \mathbb{E}^k_{t-1}[\iota_t]$		
	(1)	(2)	(3)
h_{t-1}	0.096^{***}		0.104^{***}
	(0.021)		(0.021)
ι_{t-1}		-0.033^{***}	-0.037^{***}
		(0.008)	(0.008)
$y_{t-1} - \hat{y}^k_{t-1}$	0.089^{***}	0.113^{***}	0.095^{***}
	(0.031)	(0.031)	(0.031)
Firm Fixed Effect	Yes	Yes	Yes
Sector-Year Fixed Effect	Yes	Yes	Yes
Observations	8,455	8,455	8,455
R^2	0.562	0.562	0.563
Note:			*p<0.1; **p<0.05; ***p<0.01

The sign pattern of $\tilde{\beta}_h > 0$ and $\tilde{\beta}_\iota < 0$ is consistent with an extended version of model in Corollary 1, with the capital manager holding under-extrapolative beliefs about the productivity process (i.e. his perceived persistence of productivity shock is smaller than the actual persistence) and imperfect communication between the capital and labor managers. Under-extrapolative belief makes both $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ negative, and communication friction is necessary to flip the sign of $\tilde{\beta}_h$ to positive. There has been no consensus in the literature so far about whether managerial beliefs exhibit under- or over-extrapolative beliefs. Barrero (2022) documents that sales forecasts overreacts to past sales and concludes over-extrapolation for a sample of US firms, while Ma et al. (2020) documents that sales forecasts under-reacts to past sales, alluding to under-extrapolation for a sample of Italian firms. The regression specification I consider here is strikingly different from their papers: here, over-reaction in beliefs to past information (a negative sign in $\tilde{\beta}$'s) results from under-extrapolation instead of over-extrapolation. Future work is needed to reconcile these facts with a new theory of managerial expectation formation.

I now address two robustness concerns. The first one is regarding the timing of expectations. Since firms post their guidance in a discrete manner, the first capital expenditure guidance of a year may be posted in the second, third or even the fourth quarter and therefore may not reflect the notion of "forecast". Similarly, the last sales expenditure of a year may be posted way earlier than the fourth quarter of the year and does not capture the notion of "nowcast". To mitigate the concern, I re-run the regressions in Table 2 on a subsample that contains only the first-quarter forecast of capital expenditure in year t and the last-quarter forecast of sales in year $t-1$. I tabulates the result in Table 3. We can see that the main takeaway remains the same as in the full sample: the estimates of $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ are similar to those in the full sample, statistically significant, and have the same sign pattern as in Table 2. This partially address the concern that the results are driven by not insufficient controls of fundamental uncertainty or invalid proxies. That being said, I shall mention that even in this restricted sample, the timing in data is not entirely the same as in the model: the Q4-expectation of sales in year $t-1$ and the Q1-expectation of forecasts in year t are treated as if they are conditional on the same information set $\mathcal{I}^k_{i,t-1}$, which implicitly assumes that there is no new information flows between the Q4 of year $t-1$ and Q1 of year t. This inconsistency will be better addressed using a firm sample that elicit forecast and nowcast/backcast expectations under a fixed time frame.

Table 3: Predictability under Strict Timing

	Dependent variable:		
	$\iota_t - \mathbb{E}^k_{t-1}[\iota_t]$		
	(1)	(2)	(3)
h_{t-1}	0.083***		0.090***
	(0.031)		(0.031)
ι_{t-1}		−0.035***	−0.038***
		(0.012)	(0.012)
$y_{t-1} - \hat{y}^k_{t-1}$	0.193***	0.214***	0.200***
	(0.057)	(0.057)	(0.057)
Firm Fixed Effect	Yes	Yes	Yes
Sector-Year Fixed Effect	Yes	Yes	Yes
Observations	4,327	4,327	4,327
R^2	0.656	0.656	0.657
Note:			*p<0.1; **p<0.05; ***p<0.01

Another concern is whether the results in Table 2 are driven by outliers. To address the concern, I trim the 1 percent tails for investment forecast error and sales nowcast error in the full sample and rerun the regression. The results are reported in Table 4. We can see that the estimated coefficients $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ are both statistically significant at 1% level and have the same sign pattern as in Table 2. The magnitude of $\tilde{\beta}_\iota$ in the trimmed sample is almost identical to the estimates in the full sample, but the magnitude of $\tilde{\beta}_h$ is about 25% smaller than the full-sample estimates. Since the dropped observations are mostly from very large companies, this is likely indicating that communication friction is weaker for smaller companies.

Table 4: Predictability with Trimmed Sample

	Dependent variable:		
	$\iota_t - \mathbb{E}^k_{t-1}[\iota_t]$		
	(1)	(2)	(3)
h_{t-1}	0.067***		0.076***
	(0.021)		(0.021)
ι_{t-1}		−0.034***	−0.036***
		(0.007)	(0.007)
$y_{t-1} - \hat{y}^k_{t-1}$	0.241***	0.287***	0.246***
	(0.063)	(0.062)	(0.063)
Firm Fixed Effect	Yes	Yes	Yes
Sector-Year Fixed Effect	Yes	Yes	Yes
Observations	8,099	8,099	8,099
R^2	0.517	0.518	0.519
Note:	*p<0.1; **p<0.05; ***p<0.01		

In summary, in this section I describe the strategy of identifying communication friction from the expectation data. The US public firm's managerial expectation data shows that the investment forecast simultaneously over-reacts to the firm's past investment and under-reacts to the firm's past hiring. I show that this sign pattern cannot be easily reconciled by models without communication friction and is more consistent with a model with communication friction and under-extrapolative beliefs.

4 Quantitative Analysis

In this section, I quantify the aggregate effect of communication friction. I will first extend the model in Section 2 to incorporate other well-documented distortions and frictions in the firm dynamics literature. Then, I will discuss the calibration strategy that allows me to identify the parameters for each mechanism in the extended model. After estimating the extended model, I will use the calibrated parameters to investigate the role communication friction plays in accounting for the patterns in the data.

4.1 Setup, Calibration Strategy and Data

An Extended Model for Quantification To obtain a credible quantification of aggregate effect of communication friction, I extend our baseline model so that it incorporates a number of well-documented frictions and distortions in the firm dynamics literature. These frictions include time-to-build, convex adjustment costs, cross-factor distortions and factor-specific distortions with correlated and uncorrelated parts, and firm characteristics.

Specifically, I assume that the firm's profit function becomes

$$\begin{aligned}\Pi_{it} &= (1-\tau^y_{it})R(K_{i,t-1}, N_{it}; A_{it}) - (1+\tau^k_{it})R_t K_{it} - W_t N_{it} \\ &\quad - \frac{\xi_k}{2}\left(\frac{K_{it}}{K_{i,t-1}} - 1\right)^2 K_{i,t-1} - \frac{\xi_n}{2}\left(\frac{N_{it}}{N_{i,t-1}} - 1\right)^2 N_{i,t-1} \end{aligned} \tag{28}$$

where $R(K_{i,t-1}, N_{it}; A_{it})$ is the revenue function with capital input in period t being predetermined so as to capture the time-to-build friction, and τ^y_{it} and τ^k_{it} are reduced-form firm-specific distortions as in Hsieh and Klenow (2009). τ^y_{it} captures the mechanisms that distort all factors symmetrically without changing the capital-labor ratio (such as price markup), while τ^k_{it} captures the capital-specific distortions (such as financial frictions) that create variations the capital-labor ratio. Following David and Venkateswaran (2019), I assume that these reduced-form distortions have the following structure:

$$\tau^y_{it} = \gamma^y a_{it} + \delta^y_i + \eta^y_{it} \tag{29}$$

$$\tau^k_{it} = \gamma^k a_{it} + \delta^k_i + \eta^k_{it} \tag{30}$$

where $\gamma^y, \gamma^k \in (0,1)$ captures the proportion of distortions correlated with productivity. They are referred to as "correlated distortion" as in Bartelsman et al. (2013) and are used to model size-dependent policies (such as antitrust regulations). $\eta_{it}^y \sim N(0, \sigma_{\eta,y}^2)$ and $\eta_{it}^k \sim N(0, \sigma_{\eta,k}^2)$ are the proportion of distortions uncorrelated to productivity. Lastly, $\delta_i^y \sim N(0, \sigma_{\delta,y}^2)$ and $\delta_i^k \sim N(0, \sigma_{\delta,k}^2)$ are firm characteristics that do not change over time. I assume that the correlated distortion γ^y, γ^k, firm characteristic δ_i^y, δ_i^k and the uncorrelated distortions η_{it}^y, η_{it}^k are common knowledge within the firm, so that the only source of uncertainty about τ^y and τ^k is the uncertainty about productivity a_{it}.

Calibration Strategy In total, there are 12 parameters to be calibrated in the extended model: convex adjustment costs ξ_k and ξ_n, correlated distortions γ^k and γ^y, dispersion of firm characteristics $\sigma_{\delta,k}$ and $\sigma_{\delta,y}$, dispersion of uncorrelated distortions $\sigma_{\eta,k}$ and $\sigma_{\eta,y}$, and the behavioral/information parameters that govern the expectation formation process $(\sigma_{\epsilon,p}, \sigma_{\epsilon,k}, \sigma_{\epsilon,n}, \lambda)$.

The extended model will be estimated by a Simulated Method of Moments (SMM) approach. I pick 12 empirical moments as targets in the calibration:

- The regression coefficient from (25) and (27), $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$:
  As I have discussed in the previous section, the predictability of investment forecast errors by past investment and by past hiring are indicative of communication friction and belief extrapolation and hence help us identify the information and behavioral friction parameters. To match $\tilde{\beta}_h > 0$ and $\tilde{\beta}_\iota < 0$, we expect $\lambda < 1$ (under-extrapolation) and $\sigma_{\epsilon,k}, \sigma_{\epsilon,n} < \infty$ (communication friction).

- The input-productivity correlation, $corr(\iota, \Delta a)$ and $corr(h, \Delta a)$:
  They capture the responsiveness of input choices to variations in productivity, and each of the mechanism I consider in the extended model (information friction, adjustment cost, correlated distortion, uncorrelated distortion) is able to attenuate the responsiveness of inputs to productivity shocks.

- Serial correlation of investment and hiring, $corr(\iota, \iota_{-1})$ and $corr(h, h_{-1})$:
  They are sensitive to adjustment frictions and help identify adjustment frictions from other mechanisms that attenuate the input response to productivity. High adjustment costs create an incentive for intertemporally smoothing the adjustment, making serial correlation high and responsiveness low. Other mechanisms attenuate the input response concurrently and are less sensitive to serial correlations.

- Volatility of investment and hiring, $var(\iota)$ and $var(h)$:
  These moments help separate uncorrelated distortions from other attenuation forces.

Uncorrelated distortion attenuates the responsiveness of inputs by increasing the volatility of inputs, while adjustment costs and information frictions tend to lower the volatility.

- Correlation between MRPK and MPRN with productivity, $corr(MRPK, a)$ and $corr(MRPN, a)$: The correlation between productivity and marginal revenue product of capital and labor is sensitive to correlated distortion. Low input responsiveness requires high γ^k and γ^y, which makes the MPRK and MPRN become more correlated with productivity.

- Observed dispersion in MRPK and MRPN, $var(MRPK)$ and $var(MRPN)$: Matching the dispersion in MPRK and MPRN allows the across-firm variation by the firm characteristic terms to absorb the proportion of observed capital and labor misallocation that can't be captured by within-firm variations.

The Simulated Method of Moments estimation is done in the following steps:

1. For each parameter $\Theta = (\xi_k, \xi_n, \sigma_p, \sigma_{\epsilon,k}, \sigma_{\epsilon,n}, \lambda, \gamma^k, \gamma^y, \sigma_{\delta,k}, \sigma_{\delta,y}, \sigma_{\eta,k}, \sigma_{\eta,y})$, solve the MPBE according to Lemma 1, 2 and 3;

2. Simulate a panel of 1,000 firms for 500 periods and compute the simulated moments $m(\Theta)$ from the panel;

3. Estimate the parameter by choosing the one that minimizes the sum of quadratic deviations of simulated moments from the data moments m:

$$\hat{\Theta} = \arg\min_{\Theta}(m(\Theta) - m)'(m(\Theta) - m)$$

The targeted moments and estimated parameter values are summarized in Table 5.

Data The data moments are computed from a merged IBES/Compustat sample for US public firms between 2000 to 2024. From the Compustat sample, I take the book value of property, plant and equipment (PPE) as the measure for capital stock.[7] Investment rate ι and hiring rate h are obtained by first-differencing the logged capital and employment series

[7]The quantitative results are robust to alternative definitions of capital stock, e.g. the ones constructed by the perpetual inventory method.

Table 5: Estimated Parameter Values and Targets

Description		Value	Source/Target
Elasticity of substitution	ε	4	Standard Value
Capital share	α	0.47	Cost share in IBES/Compustat
Persistence of Productivity	ρ	0.94	Estimates of $a_{it} = \rho a_{i,t-1} + \mu_{it}$.
Std.dev. of Productivity shock	σ_μ	0.24	
Capital convex adj. cost	ξ_k	0.22	corr(ι, ι_{-1})
Labor convex adj. cost	ξ_n	2.24	corr(h, h_{-1})
Noise on the common signal	$\sigma_{\epsilon,p}$	130.50	Reg. Coef. $\tilde{\beta}_\iota$ from (27)
Noise on K manager's signal	$\sigma_{\epsilon,k}$	1.71	Reg. Coef. $\tilde{\beta}_h$ from (25)
Noise on N manager's signal	$\sigma_{\epsilon,n}$	0.23	corr($h, \Delta a$)
Degree of Extrapolation	λ	0.95	corr($\iota, \Delta a$)
Correlated Distortion on capital	γ_k	0.00	corr(MRPN, a)
Correlated Distortion on revenue	γ_y	0.11	corr(MRPK, a)
Transitory Distortion η^k	$\sigma_{\eta,k}$	1.18	var(ι)
Transitory Distortion η^y	$\sigma_{\eta,y}$	0.91	var(h)
Permanent Distortion δ^k	$\sigma_{\delta,k}$	0.30	var(MRPK)
Permanent Distortion δ^y	$\sigma_{\delta,y}$	0.68	var(MRPN)

respectively. Value-added is constructed by sales minus the material expenditure.[8] I then construct the total revenue factor productivity (TFPR) or sales Solow residual as the firm fundamental a_{it} by

$$a_{it} = \log(\text{Value-added}) - \alpha(1 - 1/\epsilon)\log(\text{PPE}) - (1-\alpha)(1-1/\epsilon)\log(\text{Employment})$$

where the capital share α is chosen to match the cost share of labor in the sample, and the elasticity of substitution $\varepsilon = 4$ is taken from the literature (Hsieh and Klenow (2009)). As standard in the literature, the measure for capital and labor misallocation is the dispersion of the average/marginal revenue product of capital and labor (MRPK or MRPN). I define MRPK as the difference between logged value-added and logged capital and MRPN as the difference between logged value-added and logged employment. Following David and Venkateswaran (2019), I regress the MPRK, MRPN, investment, hiring and firm fundamental series on a sector-year fixed effect and keep the residual to isolate the firm-specific variation in each series.[9] I then merge the IBES Guidance, which contains firms' self-reported

[8]Material expenditure is defined as Cost of Goods Sold (COGS) + Selling, General & Administrative Expenses (XSGA) - Depreciation (DP) - Wage Bill (XLR), following the approach by Keller and Yeaple (2009) and Flynn and Sastry (2020). Since wage bill is missing for about 90% of the firm, I use the two-digit industry-level wages in the Census Bureau County Business Patterns dataset to impute the wage bill for the firm. See Appendix for details.

[9]Here, sector or industry is defined as 2-digit NAICS code for non-manufacturing industries and 3-digit NAICS code for manufacturing industries.

expectations, to the Compustat sample. Same as before, I use the first capital expenditure guidance of the year as "investment forecasts" and the last sales guidance of the year as "sales nowcast". I trim 1% tail of each variable in the merged sample, and then keep the observations that have investment, lagged investment, lagged hiring, lagged sales, capital expenditure forecast and lagged sales nowcast. The resulting sample has 6,492 firm-year observations, which I use to compute the data moments in the calibration. I relegate further details of sample construction and production function estimation to Appendix B.3.

4.2 Results

Communication Friction and Regression Coefficients Table 6 shows that our model can match the data moments well. In particular, the model-implied regression coefficients $\tilde{\beta}_h$ and $\tilde{\beta}_\iota$ are close to the data moments, and communication friction plays a key role in matching these moments. As comparison, I re-calibrate a simplified version of the model to the same set of targets, muting communication frictions by setting manager's private information to be extremely noisy ($\sigma_{\epsilon,k}, \sigma_{\epsilon,n} \to \infty$). The results are tabulated on the third column of Table 6. We can see that the model without communication friction is not able to match the magnitude and the sign pattern of the two regression coefficients. Under the common-information calibration, the model-implied predictability by lag hiring and lag investment are on the same direction: the investment forecast under-reacts to both past hiring and past investment. Our full model with communication friction can easily match the opposite-sign pattern in the data. This is consistent with the takeaway in Proposition 2 and Figure 1.

Sources of Learning The estimated parameters shed light on the learning patterns of the capital and labor managers. Table 7 shows how much the managers learn from multiple information resources. In total, there is a massive reduction in uncertainty after learning repeatedly from the noisy signals about the productivity for both managers. The posterior uncertainty is below 20% for the capital manager and below 10% for the labor manager. Regarding the source of learning, since the estimated noise of public signal is large, both managers put much more weight on their private information than the public signal in decision-making. Both managers put a substantial weight on their prior beliefs, which are heterogeneous across managers because the true productivity is never fully revealed to the managers and they reach their own priors using distinct signals. This is a subtle difference from the existing literature on firm learning (e.g. David et al. (2016) and David et al.

Table 6: Targeted Moments: Data vs Model

Moment	Data	Model (wl. Communication Friction)	Model (w/o. Communication Friction)
$\text{corr}(\iota, \iota_{-1})$	0.274	0.277	0.264
$\text{corr}(h, h_{-1})$	0.168	0.176	0.181
$\text{corr}(\iota, \Delta a)$	-0.030	-0.030	0.088
$\text{corr}(h, \Delta a)$	0.261	0.259	0.207
$\tilde{\beta}_h$	**0.085**	**0.079**	**0.008**
$\tilde{\beta}_\iota$	**-0.032**	**-0.029**	**0.013**
$\text{var}(\iota)$	0.019	0.046	0.037
$\text{corr}(\text{MRPK}, a)$	0.425	0.435	0.510
$\text{var}(\text{MRPK})$	0.419	0.424	0.465
$\text{var}(h)$	0.020	0.056	0.062
$\text{corr}(\text{MRPN}, a)$	0.366	0.371	0.352
$\text{var}(\text{MRPN})$	0.199	0.205	0.185

(2016)), which typically assumes that productivity are fully revealed and becomes public information to the managers, so that they start with the same prior in each period. This subtlety stemming from communication friction turns out to have important quantitative implications for firm dynamics.

Table 7: Learning Patterns of the Managers

	Uncertainty		Source of Learning		
	Posterior	Posterior/Prior	Public	Private	Prior
Capital Manager	0.082	17.9%	0%	2.8%	97.2%
Labor Manager	0.032	7.0%	0%	62.1%	37.9%

Sources of Low Investment-hiring Correlation The calibrated model allows us to investigate the contributions of each mechanism to several key features of data. In Table 8, I report the contribution of each friction to the low input-productivity correlation and the low investment-hiring correlation, two key features in the data. The numbers in the table are computed under the assumption that each time only one friction is active (i.e. the relevant parameters are fixed at the estimated level) while all other frictions are muted (i.e. the relevant parameters are set to the frictionless level).

In Compustat, the correlation between investment and hiring in the US public firms is low. The third column of Table 8 investigates the sources of this investment-hiring disconnect. We can see that the low investment-hiring correlation is primarily driven by communication friction and time-to-build. Each of other frictions generate almost perfect correlation in investment and hiring. Additionally, the information friction is a much more powerful force in driving investment-hiring correlation down: with information friction only, we can even generate a negative correlation between investment and hiring. This is not entirely absurd: Hawkins, Michaels, and Oh (2015) documents that firms and plants have episodes when capital and labor adjustments are in opposite directions. The leading explanation is that the complementarity and substitutability between capital and labor vary with tasks (Acemoglu and Restrepo (2019)). Our result shows that communication friction may provide an alternative explanation.

Investment-hiring correlation highlights a crucial difference between firm-level public information and manager-level private information. We can see that if we mute communication friction and only allow for firm-level public information, the investment-hiring correlation is equal to 1. Manager-level private signals, on the other hand, lower the correlation significantly. This can be seen from the expressions of capital and labor choices. Without other frictions, capital and labor are both proportional to the managers' posterior beliefs about productivity, i.e. $k_{it} \propto \mathbb{E}_{it}^k[a_{it}]$ and $n_{it} \propto \mathbb{E}_{it}^n[a_{it}]$. Absent communication friction, managers share the same information, so that $\mathbb{E}_{it}^k = \mathbb{E}_{it}^n = \mathbb{E}_{it}$, and firm-level common noisy information attenuates the responsiveness of both inputs by the same proportion, without breaking the perfect correlation between them. The existence of private signals breaks this pattern as the managers' expectations are no longer perfectly correlated in general. Moreover, in my model, since the true productivity is never fully revealed to the managers, the disagreement between managers on productivity is preserved over time, and hence the prior beliefs of the managers are different from each other. This largely amplifies the effect of imperfect communication on the capital-labor correlation.

The distinction between firm-level public information and manager-level private information is akin to the difference between revenue distortion τ^y and capital distortion τ^k. We can see from the table that uncorrelated capital distortion η^k lowers the capital-investment correlation, while the the uncorrelated revenue distortion η^y doesn't. Common, noisy information acts like the revenue distortion τ^y as they both affect all the inputs symmetrically, while the private information introduced by communication friction acts like the capital-specific distortion τ^k, as they both affect different inputs disproportionately and generate variations in capital-labor ratio.

Quantitatively, ignoring communication friction will lead to an exaggeration of the role capital distortion τ^k plays in the firm dynamics. In the table, we can see that the role of capital distortion η^k is limited in accounting for the capital-labor disconnect. This is because of two reasons. First, the calibration explicitly targets the low investment and hiring volatility in the data. This largely restricts the attenuation force of η^k, which lowers the investment-hiring correlation by increasing their volatility. Additionally, the inclusion of communication friction in the model largely absorbs the effect of capital-specific distortions in the calibration.

Table 8: Sources of Low Responsiveness & Capital-labor Disconnect

Mechanism	corr($\iota, \Delta a$)	corr($h, \Delta a$)	corr(ι, h)
Time-to-build	1.000	0.631	0.631
Convex Costs	0.734	0.669	0.995
- Capital Adjustment Cost ξ_k	0.801	0.970	0.922
- Labor Adjustment Cost ξ_n	0.964	0.718	0.877
Information frictions	-0.000	0.634	-0.026
- Public Signal $\sigma_{\epsilon,p}$	-0.001	-0.001	1.000
- Capital Manager's Private Signal $\sigma_{\epsilon,k}$	0.114	-0.022	-0.034
- Labor Manager's Private Signal $\sigma_{\epsilon,n}$	-0.034	0.634	-0.033
- Extrapolative Belief λ	1.000	1.000	1.000
Correlated distortions	1.000	1.000	1.000
- Correlated Capital Distortion γ^k	1.000	1.000	1.000
- Correlated Revenue Distortion γ^y	1.000	1.000	1.000
Firm Characteristics δ	1.000	1.000	1.000
- Firm Characteristics $\sigma_{\delta,k}$	1.000	1.000	1.000
- Firm Characteristics $\sigma_{\delta,y}$	1.000	1.000	1.000
Transitory, Uncorrelated Distortion η	0.179	0.183	0.995
- Transitory, Uncorrelated Capital Distortion $\sigma_{\eta,k}$	0.582	0.773	0.966
- Transitory, Uncorrelated Revenue Distortion $\sigma_{\eta,y}$	0.185	0.185	1.000
All mechanisms	-0.030	0.259	0.416
Data	-0.030	0.261	0.521

Sources of Low Input Responsiveness to Productivity The first two columns of Table 8 report the contribution of each friction to the low input-productivity correlation in the data. The first takeaway is that transitory, uncorrelated distortion plays an important role in attenuating the input responsiveness to productivity. With uncorrelated distortion only, the investment-productivity correlation drops to 0.179 and the hiring-productivity correlation drops to 0.183, both of which are already reasonably close to the responsiveness in the data. The second takeaway is that information frictions plays a more prominent role in attenuating capital responsiveness than in labor responsiveness. Information friction alone drives the investment-productivity correlation to 0, which is close to the -0.03 in the data. It plays a more modest role in dampening labor responses. This is mainly because the capital manager's signals are estimated to be less informative than the labor manager's. Lastly, convex costs also help dampen the labor and capital responsiveness but the effect is modest. Both investment- and hiring-productivity correlation is about around 0.7, which is higher than those implied by transitory distortion and information friction. The effect of other frictions on attenuating input responsiveness are negligible.

Table 9: Misallocation Decompositions

Mechanism	Capital misalloc.	Labor misalloc.	Aggregate TFP loss
Convex costs	0.023 (5.50%)	0.048 (23.25%)	-7.01%
Information frictions	0.211 (49.84%)	0.076 (37.10%)	-10.73%
Correlated distortions	0.005 (1.26%)	0.005 (2.60%)	-1.05%
Firm Characteristics (δ's)	0.285 (67.18%)	0.125 (60.97%)	-17.65%
Uncorrelated Distortions (η's)	0.009 (2.22%)	0.015 (7.46%)	-0.89%
All mechanisms	0.424 (100.00%)	0.205 (100.00%)	-33.34%
Data	0.419	0.199	-

Sources of Misallocation Lastly, in Table 9, I tabulate the contribution of each friction on the capital and labor misallocation and quantify the aggregate TFP loss generated by each friction. Again, these numbers are computed under the assumption that only one friction is active each time. We can see that information friction accounts for almost half of observed MRPK dispersion and about 37% of MRPN dispersion. This is mainly because the capital manager's private information is noisier than the labor manager's private information. The MPRK and MPRN dispersion resulting from information friction generate a sizable 10% of aggregate TFP loss, making it the most prominent source of misallocation among the within-firm frictions.

My quantification of information-driven misallocation is comparable to the efficiency loss suggested by the literature. David et al. (2016) estimate the US aggregate TFP loss from information friction to be between 4% (in the case of frictionless labor) and 40% (when labor is subject to the same noisy information as capital). David and Venkateswaran (2019)'s estimate is around 7%. Their quantification is based on two benchmarks: either capital and labor managers share information frictionlessly, or the labor manager has full information while the capital manager has noisy information. In other words, either the managers perfectly communicate their information with each other, or the managers do not communicate at all. My model explicitly incorporates communication friction and allows for a more sophisticated internal information structure. Hence, it can be regarded as a smooth version of the two benchmark cases, and the quantification is indeed within the range and closer to the no-communication benchmark. Another important paper by Ma et al. (2020) claims that removing imperfect expectations in sales improves the aggregate TFP by less than 1%. My quantification is significantly larger than theirs. Apart from the fact that we are using different expectational moments to discipline our models, my model differs from theirs in a crucial way: in their model, past productivity is fully revealed to the firm in the periods that follow, while in my setup it is never fully revealed and needs to be gradually learned by managers. This slow learning by Kalman filters amplifies the aggregate implication of information friction because the signal noise in earlier periods persists in future period's expectations.

The takeaways for other frictions are largely consistent with the literature. Overall, convex cost plays a modest role in misallocation and TFP loss. In my parameterization, it is a more important friction for labor misallocation than capital misallocation. It accounts for about a quarter of observed MRPN dispersion but explains only 5.5 percent of capital misallocation. Correlated distortions and transitory, uncorrelated distortions are not able to produce large capital and labor misallocation by themselves and have small implications on aggregate TFP. Lastly, as in David and Venkateswaran (2019), across-firm dispersion generated by firm characteristics δ^k and δ^y is a large contributor to the capital and labor misallocation and aggregate TFP loss. My current framework focuses on within-firm friction and assumes that the internal information structure is same across firms. Since the management practices (including information processing techniques) are highly sticky over time and heterogeneous across firms Bloom et al. (2019), one could imagine that the degrees of communication friction are likely to be different across firms. A promising future direction is to allow for across-firm heterogeneity on their practices of information sharing within the firm using richer firm surveys and open the black box of the across-firm dispersion.

5 Conclusion

This paper develops a parsimonious framework to incorporate within-firm communication friction into the analysis of firm's dynamic input choices. The framework makes it possible for us to provide a direct test on whether multiple input decisions of the firm is made under same information. Using a merged expectation-outcome dataset for the US public firms, I reject the null hypothesis and conclude that the investment decisions are not based on the same information as the hiring decisions in the firm. In the quantitative exercise, I discipline the information friction parameters using the expectation evidence and confirm that communication friction is quantitatively important. Modeling the information friction as a combination of noisy firm-level and manager-level information is the key to match both the low action-fundamental correlations and the low investment-hiring correlation in the data. Hence, information friction should not only be considered as distorting capital and labor jointly but also an input-specific distortion that alters the capital-labor ratio. Controlling for convex adjustment costs, I find that communication friction generates strong aggregate effects, explaining about a substantial portion of MRPK and MRPN dispersion in the data, and implying a sizable 10% aggregate TFP loss.

References

Acemoglu, D. and P. Restrepo (2019): "Automation and New Tasks: How Technology Displaces and Reinstates Labor," *Journal of Economic Perspectives*, 33, 3–30.

Afrouzi, H., A. Drenik, and R. Kim (2023): "Concentration, Market Power, and Misallocation: The Role of Endogenous Customer Acquisition," *NBER Working Paper No. 31415*.

Angeletos, G.-M. and Z. Huo (2021): "Myopia and Anchoring," *American Economic Review*, 111, 1166–1200.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2021): "Imperfect Macroeconomic Expectations: Evidence and Theory," *NBER Macroeconomics Annual*, 35, 1–86, _eprint: https://doi.org/10.1086/712313.

Angeletos, G.-M. and C. Lian (2016): "Incomplete Information in Macroeconomics," in *Handbook of Macroeconomics*, Elsevier, vol. 2, 1065–1240.

——— (2018): "Forward Guidance without Common Knowledge," *American Economic Review*, 108, 2477–2512.

Asriyan, V. and A. N. Kohlhas (2024): "The Macroeconomics of Data," .

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020): "The Fall of the Labor Share and the Rise of Superstar Firms*," *The Quarterly Journal of Economics*, 135, 645–709.

Bachmann, R. and C. Bayer (2014): "Investment Dispersion and the Business Cycle," *THE AMERICAN ECONOMIC REVIEW.*

Bachmann, R., G. Tapa, and W. van der Klaauw (2023): *Handbook of Economic Expectations*, Academic Press.

Barrero, J. M. (2022): "The micro and macro of managerial beliefs," *Journal of Financial Economics*, 143, 640–667.

Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2013): "Cross-Country Differences in Productivity: The Role of Allocation and Selection," *American Economic Review*, 103, 305–334.

Becker, R. A., J. C. Haltiwanger, R. S. Jarmin, S. D. Klimek, and D. J. Wilson (2005): "Micro and Macro Data Integration: The Case of Capital," *SSRN Electronic Journal.*

Bloom, N., E. Brynjolfsson, L. Foster, R. Jarmin, M. Patnaik, I. Saporta-Eksten, and J. Van Reenen (2019): "What Drives Differences in Management Practices?" *American Economic Review*, 109, 1648–1683.

Bloom, N., M. Codreanu, and R. Fletcher (2025): "Rationalizing Firm Forecasts," Tech. Rep. w33384, National Bureau of Economic Research, Cambridge, MA.

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

Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen (2014): "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science*, 60, 2859–2885.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110, 2748–82.

Coibion, O. and Y. Gorodnichenko (2012): "What can survey forecasts tell us about information rigidities?" *Journal of Political Economy*, 120, 116–159.

——— (2015): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 105, 2644–2678.

Cooper, R. W. and J. Haltiwanger (2006): "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 73, 611–633.

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

David, J. M. and V. Venkateswaran (2019): "The Sources of Capital Misallocation," *American Economic Review*, 109, 2531–2567.

Davis, S. J. and J. C. Haltiwanger (1992): "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics*, 107, 819–863.

De Loecker, J., J. Eeckhout, and G. Unger (2020): "The Rise of Market Power and the Macroeconomic Implications*," *The Quarterly Journal of Economics*, 135, 561–644.

Dessein, W., D. H.-F. Lo, and C. Minami (2022): "Coordination and Organization Design: Theory and Micro-Evidence," *American Economic Journal: Microeconomics*, 14, 804–843.

Flynn, J. P. and K. Sastry (2020): "Attention Cycles," *SSRN Electronic Journal*.

Gabaix, X. (2011): "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 79, 733–772.

——— (2020): "A Behavioral New Keynesian Model," *American Economic Review*, 110, 2271–2327.

Giroud, X. (2013): "Proximity and Investment: Evidence from Plant-Level Data*," *The Quarterly Journal of Economics*, 128, 861–915.

Hawkins, W. B., R. Michaels, and J. Oh (2015): "The Joint Dynamics of Capital and Employment at the Plant Level," *Working Paper*.

Hsieh, C.-T. and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124, 1403–1448.

Impink, S. M., A. Prat, and R. Sadun (2024): "Communication Within Firms: Evidence from CEO Turnovers," *Management Science*.

Keller, W. and S. R. Yeaple (2009): "MULTINATIONAL ENTERPRISES, INTERNATIONAL TRADE, AND PRODUCTIVITY GROWTH: FIRM LEVEL EVIDENCE FROM THE UNITED STATES," *THE REVIEW OF ECONOMICS AND STATISTICS*, 91, 821–831.

Kohlhas, A. N. and A. Walther (2021): "Asymmetric Attention," *American Economic Review*, 111, 2879–2925.

Lian, C. (2021): "A Theory of Narrow Thinking," *The Review of Economic Studies*, 88, 2344–2374.

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020): "A quantitative analysis of distortions in managerial forecasts," Tech. rep., National Bureau of Economic Research.

Malenko, N. (2022): "Information Flows, Organizational Structure, and Corporate Governance," *SSRN Electronic Journal*.

Marschak, J. and R. Radner (1972): *Economic Theory of Teams*, Yale University Press.

Mian, A. (2006): "Distance Constraints: The Limits of Foreign Lending in Poor Economies," *The Journal of Finance*, 61, 1465–1505.

Ottonello, P. and T. Winberry (2020): "Financial Heterogeneity and the Investment Channel of Monetary Policy," *Econometrica*, 88, 2473–2502.

Woodford, M. (2001): "Imperfect Common Knowledge and the Effects of Monetary Policy," Tech. Rep. w8673, National Bureau of Economic Research, Cambridge, MA.

——— (2003): *Interest and prices: Foundations of a theory of monetary policy*, Princeton University Press Princeton, NJ and Oxford.

A Theory Appendix

A.1 Second-order Approximation of the Profit Function

Denote the frictionless steady state value of variable X_t as $\bar{X}$, and denote the log deviation of X_t from $\bar{X}$ as x_t.

Frictionless Steady State Fix $\bar{A} = 1$. The profit maximization problem without information and adjustment frictions is to choose K and N to maximize

$$\Pi(K, N) = Y^{1/\epsilon} A^{1-1/\epsilon} (K^\alpha N^{1-\alpha})^{1-1/\epsilon} - RK - WN$$

FOC:

$$Y^\epsilon A^{1-1/\epsilon} K^{\alpha(1-1/\epsilon)-1} N^{(1-\alpha)(1-1/\epsilon)} \alpha(1 - 1/\epsilon) = R$$

$$Y^\epsilon A^{1-1/\epsilon} K^{\alpha(1-1/\epsilon)} N^{(1-\alpha)(1-1/\epsilon)-1} (1 - \alpha)(1 - 1/\epsilon) = W$$

In the quantitative exercise, $\alpha, \epsilon, \rho, \sigma_u, \delta, \beta$ are calibrated outside the model. Given these parameters, the frictionless steady-state values for rental rate, capital, labor and wage can be expressed analytically. First, rental price $\bar{R}$ is

$$\bar{R} = 1/\gamma - 1 + \delta = \alpha(1 - 1/\epsilon) \bar{K}^{\alpha(1-1/\epsilon)-1} N^{(1-\alpha)(1-1/\epsilon)}$$

Normalize $\bar{Y} = 1, \bar{A} = 1$, we get

$$\bar{K}^{\alpha(1-1/\epsilon)} \bar{N}^{(1-\alpha)(1-1/\epsilon)} = 1$$

We can use the above two equations, together with FOCs, to back out steady state $\bar{K}, \bar{N}, \bar{W}$:

$$\bar{K} = \frac{\alpha(1 - 1/\epsilon)}{1/\gamma - 1 + \delta}$$

$$\bar{N} = \left(\frac{1/\gamma - 1 + \delta}{\alpha(1 - 1/\epsilon)} \right)^{\alpha/(1-\alpha)}$$

$$\bar{W} = \bar{K}^\alpha \bar{N}^{-\alpha} (1 - \alpha)(1 - 1/\epsilon)$$

Second-order Approximation We have

$$
\begin{aligned}
\Pi \approx\ & \tilde{A}^{1/\epsilon}\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)} + \bar{R}\bar{K}k_t + \bar{W}\bar{N}n_t \\
& + \frac{1}{2}\bar{R}\bar{K}(1-1+\alpha(1-1/\epsilon))k_t^2 + \frac{1}{2}\bar{W}\bar{N}(1-1+(1-\alpha)(1-1/\epsilon))n_t^2 \\
& + \tilde{A}^{1/\epsilon}\alpha(1-\alpha)(1-1/\epsilon)^2\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)}k_t n_t \\
& - \bar{R}\bar{K} - \bar{R}\bar{K}k_t - \bar{R}\bar{K}\frac{1}{2}k_t^2 \\
& - \bar{W}\bar{N} - \bar{W}\bar{N}n_t - \bar{W}\bar{N}\frac{1}{2}n_t^2 \\
& - \frac{\xi_k}{2}\bar{K}(k_t-k_{t-1})^2 - \frac{\xi_n}{2}\bar{N}(n_t-n_{t-1})^2 \\
& + \frac{1}{\epsilon}\tilde{A}^{1/\epsilon}\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)}(a_t+\frac{1}{2}a_t^2) + \frac{1}{2}\frac{1}{\epsilon}(\frac{1}{\epsilon}-1)\tilde{A}^{1/\epsilon}\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)}a_t^2 \\
& + \frac{1}{\epsilon}\tilde{A}^{1/\epsilon}\alpha(1-1/\epsilon)\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)}k_t a_t + \frac{1}{\epsilon}\tilde{A}^{1/\epsilon}(1-\alpha)(1-1/\epsilon)\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)}n_t a_t
\end{aligned}
$$

where $\tilde{A} \equiv \bar{Y}\bar{A}^{\epsilon-1}$, and the sixth line is an approximation of adjustment cost for capital:

$$
\begin{aligned}
\frac{\xi_k}{2}\left(\frac{K_t-(1-\delta)K_{t-1}}{K_{t-1}}-\delta\right)^2 K_{t-1} &= \frac{\xi_k}{2}\left(\exp(k_t-k_{t-1})-1\right)^2\bar{K}\exp(k_{t-1}) \\
&= \frac{\xi_k\bar{K}}{2}(k_t-k_{t-1})^2(1+k_{t-1}+\ldots) = \frac{\xi_k\bar{K}}{2}(k_t-k_{t-1})^2
\end{aligned}
$$

Define investment $\iota_t \equiv k_t - k_{t-1}$ and hiring $h_t \equiv n_t - n_{t-1}$.

Replace $k_t = k_{t-1} + \iota_t$ and $n_t = h_t + n_{t-1}$:

$$
\begin{aligned}
\pi =\ & \frac{1}{\epsilon}\bar{A}^{1/\epsilon}\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)} \\
& + 0k_t + 0n_t + \frac{1}{\epsilon}\bar{A}^{1/\epsilon}\bar{K}^{\alpha(1-1/\epsilon)}\bar{N}^{(1-\alpha)(1-1/\epsilon)}a_t \\
& + \frac{1}{2}\bar{R}\bar{K}(\alpha(1-1/\epsilon)-1)k_t^2 + \frac{1}{2}\bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1)n_t^2 + \frac{1}{2}\frac{1}{\epsilon^2}\bar{Y}a_t^2 \\
& + \alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y}k_t n_t + \frac{1}{\epsilon}\alpha(1-1/\epsilon)\bar{Y}k_t a_t + \frac{1}{\epsilon}(1-\alpha)(1-1/\epsilon)\bar{Y}n_t a_t \\
& - \frac{\xi_k}{2}\bar{K}(k_t-k_{t-1})^2 - \frac{\xi_n}{2}\bar{N}(n_t-n_{t-1})^2 \qquad \text{(A1)} \\
=\ & \frac{1}{\epsilon}\bar{Y} + \frac{1}{\epsilon}\bar{Y}a_t + \frac{1}{2}\bar{R}\bar{K}(\alpha(1-1/\epsilon)-1)(k_{t-1}+\iota_t)^2 + \frac{1}{2}\bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1)(n_{t-1}+h_t)^2 \\
& + \frac{1}{2}\frac{1}{\epsilon^2}\bar{Y}a_t^2 \\
& + \alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y}(k_{t-1}+\iota_t)(n_{t-1}+h_t) + \frac{1}{\epsilon}\alpha(1-1/\epsilon)\bar{Y}(k_{t-1}+\iota_t)a_t
\end{aligned}
$$

$$+\frac{1}{\epsilon}(1-\alpha)(1-1/\epsilon)\bar{Y}(n_{t-1}+h_t)a_t-\frac{\xi_k\bar{K}}{2}\iota_t^2-\frac{\xi_n\bar{N}}{2}h_t^2$$

Hence

$$\begin{aligned}\pi =& \frac{1}{\epsilon}\bar{Y}+\frac{1}{\epsilon}\bar{Y}a_t+\frac{1}{2}\frac{1}{\epsilon^2}\bar{Y}a_t^2\\&+\frac{1}{2}\bar{R}\bar{K}(\alpha(1-1/\epsilon)-1)\left(k_{t-1}^2+2k_{t-1}\iota_t+\iota_t^2\right)\\&+\frac{1}{2}\bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1)(n_{t-1}^2+2n_{t-1}h_t+h_t^2)\\&+\alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y}(k_{t-1}n_{t-1}+\iota_t n_{t-1}+k_{t-1}h_t+\iota_t h_t)\\&+\frac{1}{\epsilon}\alpha(1-1/\epsilon)\bar{Y}(k_{t-1}a_t+\iota_t a_t)+\frac{1}{\epsilon}(1-\alpha)(1-1/\epsilon)\bar{Y}(n_{t-1}a_t+h_t a_t)\\&-\frac{\xi_k}{2}\bar{K}\iota_t^2-\frac{\xi_n}{2}\bar{N}h_t^2\end{aligned}\tag{A2}$$

Hence The log-quadratic profit function takes the form

$$x_t'Px_t+x_t'Q\iota_t+x_t'Rh_t+H_\iota\iota_t^2+H_{\iota h}\iota_t h_t+H_h h_t^2\tag{A3}$$

where state vector $x_t=[1,k_{t-1},n_{t-1},a_t]$, and the matrices are

$$P=\begin{bmatrix}\frac{1}{\epsilon}\bar{Y} & 0 & 0 & \frac{\epsilon-1}{2\epsilon}\bar{Y}\\ * & \frac{1}{2}\bar{R}\bar{K}(\alpha(1-1/\epsilon)-1) & \frac{1}{2}\alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y} & \frac{\epsilon-1}{2\epsilon}\alpha(1-1/\epsilon)\bar{Y}\\ * & * & \frac{1}{2}\bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1) & \frac{\epsilon-1}{2\epsilon}(1-\alpha)(1-1/\epsilon)\bar{Y}\\ * & * & * & \frac{1}{2}\frac{(\epsilon-1)^2}{\epsilon^2}\bar{Y}\end{bmatrix}$$

$$Q=\begin{bmatrix}0\\ \bar{R}\bar{K}(\alpha(1-1/\epsilon)-1)\\ \alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y}\\ \frac{1}{\epsilon}\alpha(1-1/\epsilon)\bar{Y}\end{bmatrix}$$

$$R=\begin{bmatrix}0\\ \alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y}\\ \bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1)\\ \frac{1}{\epsilon}(1-\alpha)(1-1/\epsilon)\bar{Y}\end{bmatrix}$$

and $H_\iota = \frac{1}{2}\bar{R}\bar{K}(\alpha(1-1/\epsilon)-1) - \frac{1}{2}\xi_k\bar{K}$, $H_h = \frac{1}{2}\bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1) - \frac{1}{2}\xi_n\bar{N}$, $H_{\iota h} = \alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y}$.

The law of motion for the state vector x_t is given by

$$x_{t+1} \equiv \begin{bmatrix} 1 \\ k_{t-1}+\iota_t \\ n_{t-1}+h_t \\ \rho a_t + \mu_{t+1} \end{bmatrix} = Ax_t + B\iota_t + Ch_t + D\mu_{t+1}$$

where

$$A \equiv \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \rho \end{bmatrix}, B \equiv \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, C \equiv \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, D \equiv \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

A.2 Proof of Lemma 2

To save some notations, let

$$\hat{a}_t^k \equiv E\left[a_{it}\middle|\mathcal{I}_{it}^k\right], \hat{\Sigma}_t^k \equiv E\left[(a_{it}-\hat{a}_t^k)(a_{it}-\hat{a}_t^k)'\right]$$

$$\hat{a}_t^n \equiv E\left[a_{it}\middle|\mathcal{I}_{it}^n\right], \hat{\Sigma}_t^n \equiv E\left[(a_{it}-\hat{a}_t^n)(a_{it}-\hat{a}_t^n)'\right]$$

and

$$\hat{a}_t^{(k,n)} \equiv E\left[\hat{a}_t^n\middle|\mathcal{I}_{it}^k\right], \hat{\Sigma}_t^{(k,n)} \equiv E\left[(\hat{a}_t^{(k,n)}-\hat{a}_t^n)(\hat{a}_t^{(k,n)}-\hat{a}_t^n)'\right]$$

$$\hat{a}_t^{(n,k)} \equiv E\left[\hat{a}_t^k\middle|\mathcal{I}_{it}^n\right], \hat{\Sigma}_t^{(n,k)} \equiv E\left[(\hat{a}_t^{(n,k)}-\hat{a}_t^k)(\hat{a}_t^{(n,k)}-\hat{a}_t^k)'\right]$$

so that

$$a_{it}\middle|\mathcal{I}_{it}^k \sim N(\hat{a}_t^k, \hat{\Sigma}_t^k) \text{ and } a_{it}\middle|\mathcal{I}_{it}^n \sim N(\hat{a}_t^n, \hat{\Sigma}_t^n) \tag{A4}$$

characterize the first-order beliefs of capital and labor manager on the state a_{it}, and

$$\hat{a}_t^n\middle|\mathcal{I}_{it}^k \sim N(\hat{a}_t^{(k,n)}, \hat{\Sigma}_t^{(k,n)}) \text{ and } \hat{a}_t^k\middle|\mathcal{I}_{it}^n \sim N(\hat{a}_t^{(n,k)}, \hat{\Sigma}_t^{(n,k)}) \tag{A5}$$

characterize the high-order beliefs of capital and labor manager on each other's beliefs about a_{it}.

Lemma 2 states that the first-order and higher-order uncertainty are given by the following triplet of Riccati equations:

$$\begin{aligned}
\hat{\Sigma}_t^k &= (\hat{\Sigma}_t^k)^- - (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k \right]^{-1} H(\hat{\Sigma}_t^k)^- && \text{(A6)} \\
\hat{\Sigma}_t^n &= (\hat{\Sigma}_t^n)^- - (\hat{\Sigma}_t^n)^- H' \left[H(\hat{\Sigma}_t^n)^- H' + W^n \right]^{-1} H(\hat{\Sigma}_t^n)^- && \text{(A7)} \\
\hat{\Sigma}_t^{(k,n)} &= \hat{\Sigma}_t^k + \hat{\Sigma}_t^n - 2\tilde{\Sigma}_t^{(k,n)} && \text{(A8)}
\end{aligned}$$

where the pre-estimate uncertainty $(\hat{\Sigma}_t^k)^-, (\hat{\Sigma}_t^n)^-$ are given by

$$(\hat{\Sigma}_{t-1}^k)^+ = \hat{\Sigma}_{t-1}^k - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)} \right]^{-1} (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})' \tag{A9}$$

$$\begin{aligned}
(\hat{\Sigma}_t^k)^- &= \rho^2 \left\{ \hat{\Sigma}_{t-1}^k - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)} \right]^{-1} (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})' \right\} + \sigma_\mu^2 \\
(\hat{\Sigma}_t^n)^- &= \rho^2 \left\{ \hat{\Sigma}_{t-1}^n - (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)} \right]^{-1} (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)})' \right\} + \sigma_\mu^2
\end{aligned}$$

and the covariance between forecast errors $\tilde{\Sigma}_t^{(k,n)}$ is defined recursively by

$$\begin{aligned}
\tilde{\Sigma}_t^{(k,n)} = &\left(I - (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k \right]^{-1} H \right) \Lambda_t \left(I - (\hat{\Sigma}_t^n)^- H' \left[H(\hat{\Sigma}_t^n)^- H' + W^n \right]^{-1} H \right)' \\
&+ (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k \right]^{-1} \Omega_t \left[H(\hat{\Sigma}_t^n)^- H' + W^n \right]^{-1} H(\hat{\Sigma}_t^n)^- && \text{(A10)}
\end{aligned}$$

with

$$\Lambda_t = \rho^2 \left\{ \tilde{\Sigma}_{t-1}^{(k,n)} + (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)} \right]^{-1} (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)}) \right\} + \sigma_\mu^2$$

Initial Beliefs To fully specify the belief dynamics, I assume that the capital manager believes that the initial state $a_{i,0}$ is drawn from a Gaussian distribution

$$a_{i,0} \sim N(\hat{a}_0^k, W_0^k)$$

and that the initial information set $\mathcal{I}_0^k = \{\hat{a}_0^k\}$. Similarly, for the labor manager, she believes that the initial state is drawn from a Gaussian distribution

$$a_{i,0} \sim N(\hat{a}_0^n, W_0^n)$$

and her initial information set is $\mathcal{I}_0^n = \{\hat{a}_t^n\}$. For simplicity, I assume that the two initial distributions are independent to each other.

For simplicity, I assume that W_0^k and W_0^n are common knowledge. To characterize the initial higher-order belief $\hat{a}_0^n|\mathcal{I}_0^k$ and $\hat{a}_0^k|\mathcal{I}_0^n$, it is perhaps more convenient to define the initial forecast errors e_0^k, e_0^n as

$$e_0^k \equiv \hat{a}_0^k - a_{i,0} \text{ and } e_0^n \equiv \hat{a}_0^n - a_{i,0}$$

with $e_0^k \sim N(0, W_0^k)$ and $e_0^n \sim N(0, W_0^n)$ being independent to each other. Therefore

$$\hat{a}_0^n = \hat{a}_0^k - e_0^k + e_0^n$$

From the capital manager's perspective,

$$E\left[\hat{a}_0^n|\mathcal{I}_0^k\right] = \hat{a}_0^k$$

$$var\left[\hat{a}_0^n|\mathcal{I}_0^k\right] = var(-e_0^k + e_0^n) = W_0^n + W_0^k$$

Therefore the initial high-order beliefs are

$$\hat{a}_0^n|\mathcal{I}_0^k \sim N(\hat{a}_0^k, W_0^n + W_0^k) \text{ and } \hat{a}_0^k|\mathcal{I}_0^n \sim N(\hat{a}_0^n, W_0^n + W_0^k)$$

A first-glance at HOB at period t The same method can be used to characterize the first moment of high-order beliefs at period t. At period t, from the capital manager's perspective,

$$\hat{a}_t^n = \hat{a}_t^k - e_t^k + e_t^n$$

where e_t^k and e_t^n are forecast errors of the capital and labor manager. Hence

$$\hat{a}_t^{(k,n)} = E\left[\hat{a}_t^n|\mathcal{I}_t^k\right] = \hat{a}_t^k \quad \text{(A11)}$$

$$\hat{a}_t^{(n,k)} = E\left[\hat{a}_t^k|\mathcal{I}_t^n\right] = \hat{a}_t^n \quad \text{(A12)}$$

Characterizing the second moment of high-order belief is a bit more involved since the forecast errors e_t^k and e_t^n are correlated with each other. For simplicity, I denote the correlation $E[e_t^k e_t^n]$ as $\tilde{\Sigma}_t^{(k,n)}$. Immediately, we have

$$\hat{\Sigma}_t^{(k,n)} = E\left[(\hat{a}_t^n - \hat{a}_t^{(k,n)})^2\right] = var\left[-e_t^k + e_t^n\right] = \hat{\Sigma}_t^k + \hat{\Sigma}_t^n - 2\tilde{\Sigma}_t^{(k,n)} \tag{A13}$$

Partitioning the State Vector Note that our state vector $x_t \equiv [1, k_{t-1}, n_{t-1}, a_{it}]'$ can be partitioned into a perfectly observed part $s_t \equiv [1, k_{i,t-1}, n_{i,t-1}]'$ and an imperfectly observed part a_{it}. Hence, the linear policy

$$\iota_t = F_\iota E\left[x_t | \mathcal{I}_t^k\right], h_t = F_h E\left[x_t | \mathcal{I}_t^n\right]$$

can be rewritten as

$$\iota_t = F_\iota^s s_t + F_\iota^a \hat{a}_t^k \tag{A14}$$
$$h_t = F_h^s s_t + F_h^a \hat{a}_t^n \tag{A15}$$

Now I specify the observation equations for the partitioned state vector. The capital manager's signals about a_{it} are summarized by

$$z_t^k \equiv \begin{bmatrix} a_t^p \\ a_t^k \end{bmatrix} = Ha_t + w_t^k \tag{A16}$$

where $H = [1, 1]'$ and $w_t^k = [\epsilon_t^p, \epsilon_t^k]' \sim N(0, diag(\sigma_p^2, \sigma_{\epsilon,k}^2))$. Similarly, the labor manager's signals about a_{it} are summarized by

$$z_t^n \equiv \begin{bmatrix} a_t^p \\ a_t^n \end{bmatrix} = Ha_t + w_t^n \tag{A17}$$

where $H = [1, 1]'$ and $w_t^k = [\epsilon_t^p, \epsilon_t^n]' \sim N(0, diag(\sigma_p^2, \sigma_{\epsilon,n}^2))$.

Updating FOB and HOBs Now, suppose we are at the end of period $t-1$, with information set $\mathcal{I}_{t-1}^k$ and $\mathcal{I}_{t-1}^n$. From now on I will only focus on how the capital manager updates beliefs, given he knows that the labor manager's action takes the form

$$h_t = F_h^s s_t + F_h^a \hat{a}_t^n$$

the arguments for the labor manager are basically the same.

At period t, the belief updating takes two steps:

1. The capital manager observes the labor manager's action $h_{t-1} = F_h^s s_{t-1} + F_h^a \hat{a}_{t-1}^n$ and updates the estimates about a_{t-1} as well as a_t. Formally, after observing h_{t-1}, the capital manager computes

$$a_{t-1} \big| \mathcal{I}_{t-1}^k \cup \{h_{t-1}\} \sim N\left((\hat{a}_{t-1}^k)^+, (\hat{\Sigma}_{t-1}^k)^+\right)$$

   and

$$a_t \big| \mathcal{I}_{t-1}^k \cup \{h_{t-1}\} \sim N\left((\hat{a}_t^k)^-, (\hat{\Sigma}_t^k)^-\right)$$

2. Then the capital manager uses the private signal z_t^k to further update the state estimate a_t on top of the pre-estimate $a_t \big| \mathcal{I}_{t-1}^k \cup \{h_{t-1}\}$ and get

$$a_t \big| \mathcal{I}_{t-1}^k \cup \{h_{t-1}\} \cup \{z_t^k\} = a_t \big| \mathcal{I}_t^k \sim N\left(\hat{a}_t^k, \hat{\Sigma}_t^k\right)$$

We do it step-by-step.

The first step: note that

$$\begin{bmatrix} a_{t-1} \\ h_{t-1} \end{bmatrix} \Bigg| \mathcal{I}_{t-1}^k = \begin{bmatrix} a_{t-1} \\ F_h^s s_{t-1} + F_h^a \hat{a}_{t-1}^n \end{bmatrix} \Bigg| \mathcal{I}_{t-1}^k \sim N\left(\begin{bmatrix} \hat{a}_{t-1}^k \\ F_h^s s_{t-1} + F_h^a \hat{a}_{t-1}^{(k,n)} \end{bmatrix}, \begin{bmatrix} \hat{\Sigma}_{t-1}^k & * \\ * & F_h^a \hat{\Sigma}_{t-1}^{(k,n)} (F_h^a)' \end{bmatrix} \right) \tag{A18}$$

where $*$ is equal to

$$\begin{aligned} * &= E\left[(a_{t-1} - \hat{a}_{t-1}^k)(\hat{a}_{t-1}^n - \hat{a}_{t-1}^{(k,n)})'(F_h^a)'\right] = E\left[(a_{t-1} - \hat{a}_{t-1}^k)(\hat{a}_{t-1}^n - \hat{a}_{t-1}^k)'(F_h^a)'\right] \\ &= E\left[-e_{t-1}^k(\hat{a}_{t-1}^n - a_{t-1} + a_{t-1} - \hat{a}_{t-1}^k)'(F_h^a)'\right] = E\left[-e_{t-1}^k(e_{t-1}^n - e_{t-1}^k)'(F_h^a)'\right] \\ &= (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})(F_h^a)' \end{aligned} \tag{A19}$$

Therefore the capital manager updates his belief about a_{t-1} using the labor manager's action by

$$(\hat{a}_{t-1}^k)^+ = \hat{a}_{t-1}^k + (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})(F_h^a)' \left[F_h^a \hat{\Sigma}_{t-1}^{(k,n)} (F_h^a)'\right]^{-1} F_h^a (\hat{a}_{t-1}^n - \hat{a}_{t-1}^k) \tag{A20}$$

$$(\hat{\Sigma}_{t-1}^k)^+ = \hat{\Sigma}_{t-1}^k - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})(F_h^a)' \left[F_h^a \hat{\Sigma}_{t-1}^{(k,n)} (F_h^a)'\right]^{-1} F_h^a (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})' \tag{A21}$$

Since F_h^a is a scalar, these can be simplified as

$$(\hat{a}_{t-1}^k)^+ = \hat{a}_{t-1}^k + (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1} (\hat{a}_{t-1}^n - \hat{a}_{t-1}^k) \tag{A22}$$

$$(\hat{\Sigma}_{t-1}^k)^+ = \hat{\Sigma}_{t-1}^k - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1} (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)})' \tag{A23}$$

and therefore, the pre-estimates for a_t after observing the opponent's action is given by

$$(\hat{a}_t^k)^- = \rho(\hat{a}_{t-1}^k)^+ \tag{A24}$$

$$(\hat{\Sigma}_t^k)^- = \rho^2(\hat{\Sigma}_{t-1}^k)^+ + \sigma_\mu^2 \tag{A25}$$

Now we incorporate the signal z_t^k into the information set and update the belief about a_t. We have

$$\begin{bmatrix} a_t \\ z_t^k \end{bmatrix} \Big| \mathcal{I}_{t-1}^k \cup \{h_{t-1}\} \sim N\left(\begin{bmatrix} (\hat{a}_t^k)^- \\ H(\hat{a}_t^k)^- \end{bmatrix}, \begin{bmatrix} (\hat{\Sigma}_t^k)^- & (\hat{\Sigma}_t^k)^- H' \\ H(\hat{\Sigma}_t^k)^- & H(\hat{\Sigma}_t^k)^- H' + W^k \end{bmatrix} \right)$$

where $W^k \equiv diag(\sigma_p^2, \sigma_{\epsilon,k}^2)$. Therefore,

$$\hat{a}_t^k = (\hat{a}_t^k)^- + (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k\right]^{-1} \left(z_t^k - H(\hat{a}_t^k)^-\right) \tag{A26}$$

$$\hat{\Sigma}_t^k = (\hat{\Sigma}_t^k)^- - (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k\right]^{-1} H(\hat{\Sigma}_t^k)^- \tag{A27}$$

To close the analysis, I need to specify how high-order beliefs evolve. According to (A13), the uncertainty of the capital manager on the labor manager's expectation about the productivity is a function of (1) the capital manager's posterior uncertainty $\hat{\Sigma}_t^k$ on the productivity at period t, (2) the labor manager's posterior uncertainty $\hat{\Sigma}_t^n$ on the productivity at period t, and (3) the covariance between capital manager and labor manager's forecast errors $\tilde{\Sigma}_t^{(k,n)}$. We already have the law of motion for $\hat{\Sigma}_t^k$ given by (A27) and a similar law of motion for $\hat{\Sigma}_t^n$. It remains to characterize the covariance for forecast errors $\tilde{\Sigma}_t^{(k,n)}$.

According to (A26), the forecast error at period t is given by

$$\hat{a}_t^k - a_t = \left(I - (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k\right]^{-1} H\right) \left((\hat{a}_t^k)^- - a_t\right) + (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k\right]^{-1} w_t^k \tag{A28}$$

and similarly

$$\hat{a}_t^n - a_t = \left(I - (\hat{\Sigma}_t^n)^- H' \left[H(\hat{\Sigma}_t^n)^- H' + W^n\right]^{-1} H\right) \left((\hat{a}_t^n)^- - a_t\right) + (\hat{\Sigma}_t^n)^- H' \left[H(\hat{\Sigma}_t^n)^- H' + W^n\right]^{-1} w_t^n \tag{A29}$$

Therefore

$$\begin{aligned}\tilde{\Sigma}_t^{(k,n)} = &\left(I - (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k\right]^{-1} H\right) \Lambda_t \left(I - (\hat{\Sigma}_t^n)^- H' \left[H(\hat{\Sigma}_t^n)^- H' + W^n\right]^{-1} H\right)' \\ &+ (\hat{\Sigma}_t^k)^- H' \left[H(\hat{\Sigma}_t^k)^- H' + W^k\right]^{-1} \Omega_t \left[H(\hat{\Sigma}_t^n)^- H' + W^n\right]^{-1} H(\hat{\Sigma}_t^n)^- \end{aligned} \tag{A30}$$

where $\Lambda_t \equiv E\left[(a_t - (\hat{a}_t^k)^-)(a_t - (\hat{a}_t^n)^-)'\right]$ and $\Omega_t = E\left[w_t^k (w_t^n)'\right] = diag(\sigma_p^2, 0)$.

We know that

$$a_t - (\hat{a}_t^k)^- = \rho a_{t-1} + \mu_t - \rho(\hat{a}_{t-1}^k)^+ = \rho(a_{t-1} - (\hat{a}_{t-1}^k)^+) + \mu_t$$

$$a_t - (\hat{a}_t^n)^- = \rho a_{t-1} + \mu_t - \rho(\hat{a}_{t-1}^n)^+ = \rho(a_{t-1} - (\hat{a}_{t-1}^n)^+) + \mu_t$$

and therefore

$$\Lambda_t = E\left[(a_t - (\hat{a}_t^k)^-)(a_t - (\hat{a}_t^n)^-)'\right] = \rho^2 E\left[(a_{t-1} - (\hat{a}_{t-1}^k)^+)(a_{t-1} - (\hat{a}_{t-1}^n)^+)'\right] + \sigma_\mu^2$$

It remains to compute $E\left[(a_{t-1} - (\hat{a}_{t-1}^k)^+)(a_{t-1} - (\hat{a}_{t-1}^n)^+)'\right]$. Note that

$$a_{t-1} - (\hat{a}_{t-1}^k)^+ = -\left(I - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1}\right) e_{t-1}^k - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1} e_{t-1}^n$$

$$a_{t-1} - (\hat{a}_{t-1}^n)^+ = -\left(I - (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1}\right) e_{t-1}^n - (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1} e_{t-1}^k$$

Hence

$$\begin{aligned}&E\left[(a_{t-1} - (\hat{a}_{t-1}^k)^+)(a_{t-1} - (\hat{a}_{t-1}^n)^+)'\right] \\ &= \left(I - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1}\right) \tilde{\Sigma}_{t-1}^{(k,n)} \left(I - (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1}\right)' \\ &+ (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1} \hat{\Sigma}_{t-1}^n \left(I - (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1}\right)' \\ &+ \left(I - (\hat{\Sigma}_{t-1}^k - \tilde{\Sigma}_{t-1}^{(k,n)}) \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1}\right) \hat{\Sigma}_{t-1}^k \left[\hat{\Sigma}_{t-1}^{(k,n)}\right]^{-1} (\hat{\Sigma}_{t-1}^n - \tilde{\Sigma}_{t-1}^{(k,n)})'\end{aligned}$$

$$+ (\hat{\Sigma}^k_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1}) \left[\hat{\Sigma}^{(k,n)}_{t-1}\right]^{-1} \tilde{\Sigma}^{(k,n)}_{t-1} \left[\hat{\Sigma}^{(k,n)}_{t-1}\right]^{-1} (\hat{\Sigma}^n_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1})'$$

$$= \tilde{\Sigma}^{(k,n)}_{t-1} + (\hat{\Sigma}^k_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1}) \left[\hat{\Sigma}^{(k,n)}_{t-1}\right]^{-1} (\hat{\Sigma}^n_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1}) \tag{A31}$$

and therefore

$$\Lambda_t = \rho^2 \left\{\tilde{\Sigma}^{(k,n)}_{t-1} + (\hat{\Sigma}^k_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1}) \left[\hat{\Sigma}^{(k,n)}_{t-1}\right]^{-1} (\hat{\Sigma}^n_{t-1} - \tilde{\Sigma}^{(k,n)}_{t-1})\right\} + \sigma^2_\mu \tag{A32}$$

and we finally complete our characterization of belief updates.

A.3 Proof of Lemma 1

A.3.1 Part I: linearity of MPE

Guess that ι and h are linear:

$$\iota_t = F_\iota x_t, h_t = F_h x_t$$

Hence, the MPE is defined as

- Given F_h, the capital manager's strategy $\iota_t = F_\iota x_t$ solves the Bellman equation

$$\begin{aligned} V^k(x_t) &= \max_{\iota_t} x_t' P x_t + x_t' Q \iota_t + x_t' R F_h x_t + H_\iota \iota_t^2 + H_{\iota h} \iota_t F_h x_t + x_t' F_h' H_h F_h x_t + \gamma E\left[V^k(x_{t+1})\right] \\ &= \max_{\iota_t} x_t' \left(P + R F_h + F_h' H_h F_h\right) x_t + x_t' \left(Q + F_h' H_{\iota h}\right) \iota_t + H_\iota \iota_t^2 + \gamma E\left[V^k(x_{t+1})\right] \end{aligned} \tag{A33}$$

  subject to the law of motion (14) with h_t replaced by $F_h x_t$:

$$x_{t+1} = (A + C F_h) x_t + B \iota_t + D \mu_{i,t+1} \tag{A34}$$

- Given F_ι, the labor manager's strategy $h_t = F_h x_t$ solves the Bellman equation

$$\begin{aligned} V^n(x_t) &= \max_{h_t} x_t' P x_t + x_t' Q F_\iota x_t + x_t' R h_t + x_t' F_\iota' H_\iota F_\iota x_t + x_t' F_\iota' H_{\iota h} h_t + H_h h_t^2 + \gamma E\left[V^n(x_{t+1})\right] \\ &= \max_{h_t} x_t' \left(P + Q F_\iota + F_\iota' H_\iota F_\iota\right) x_t + x_t' \left(R + F_\iota' H_{\iota h}\right) h_t + H_h h_t^2 + \gamma E\left[V^n(x_{t+1})\right] \end{aligned} \tag{A35}$$

subject to the law of motion (14), with ι_t replaced by $F_\iota x_t$:

$$x_{t+1} = (A + BF_\iota)\, x_t + Ch_t + D\mu_{i,t+1} \tag{A36}$$

With the linear-quadratic setting, we know that the value functions V^k and V^n takes the form

$$V^k(x_t) = x_t'\hat{P}_k x_t + \hat{Q}_k, V^n(x_t) = x_t'\hat{P}_n x_t + \hat{Q}_n$$

Plug it into the Bellman equations: capital manager's problem becomes

$$\begin{aligned}
&x_t' \left[P + RF_h + F_h' H_h F_h + \gamma(A + CF_h)'\hat{P}_k(A + CF_h)\right] x_t + i_t'(H_i + \gamma B'\hat{P}_k B)i_t \\
&+ i_t'(H_{ih}F_h + Q' + \gamma B'\hat{P}_k(A + CF_h))x_t + \gamma x_t'(A + CF_h)'\hat{P}_k B i_t \\
&= \cdots + i_t' \left[H_{ih}F_h + Q' + \gamma B'(\hat{P}_k + \hat{P}_k')(A + CF_h)\right] x_t
\end{aligned}$$

take first order condition on ι:

$$2(H_i + \gamma B'\hat{P}_k B)\iota_t = -H_{ih}F_h + Q' + \gamma B'(\hat{P}_k + \hat{P}_k')(A + CF_h)$$

we get the policy matrix F_ι that solves (A33):

$$F_\iota = -\frac{1}{2}\left(H_\iota + \gamma B'\hat{P}_k B\right)^{-1}\left[Q' + H_{\iota h}F_h + \gamma B'(\hat{P}_k + \hat{P}_k')(A + CF_h)\right] \tag{A37}$$

where $\hat{P}_k$ is defined by Riccati equation

$$\begin{aligned}
\hat{P}_k &= P + RF_h + F_h' H_h F_h + \gamma(A + CF_h)'\hat{P}_k(A + CF_h) \\
&\quad - \frac{1}{4}\left[Q' + H_{\iota h}F_h + 2\gamma B'\hat{P}_k(A + CF_h)\right]'\left(H_\iota + \gamma B'\hat{P}_k B\right)^{-1}\left[Q' + H_{\iota h}F_h + 2\gamma B'\hat{P}_k(A + CF_h)\right]
\end{aligned} \tag{A38}$$

and similarly the solution to (A35) is

$$F_h = -\frac{1}{2}\left(H_h + \gamma C'\hat{P}_n C\right)^{-1}\left[R' + H_{\iota h}F_\iota + 2\gamma C'\hat{P}_n(A + BF_\iota)\right] \tag{A39}$$

where $\hat{P}_n$ is defined by Riccati equation

$$\begin{aligned}
\hat{P}_n &= P + QF_\iota + F_\iota' H_\iota F_\iota + (A + BF_\iota)'\hat{P}_n(A + BF_\iota) \\
&\quad - \frac{1}{4}\left[R' + H_{\iota h}F_\iota + 2\gamma C'\hat{P}_n(A + BF_\iota)\right]'\left(H_h + \gamma C'\hat{P}_n C\right)^{-1}\left[R' + H_{\iota h}F_\iota + 2\gamma C'\hat{P}_n(A + BF_\iota)\right]
\end{aligned} \tag{A40}$$

A.3.2 Part II: linearity of MPBE

Now that we have characterized beliefs given policy F_ι, F_h, I go back to the Bellman equations and derive these policy matrices F_ι, F_h. Under incomplete information, the capital manager's Bellman equation is

$$V^k(x_t) = \max_{\iota_t} E\left[x_t'\left(P + RF_h + F_h'H_hF_h\right)x_t + x_t'\left(Q + F_h'H_{\iota h}\right)\iota_t + H_\iota \iota_t^2 + \gamma V^k(x_{t+1})\middle|\mathcal{I}_t^k\right] \tag{A41}$$

To save some notation, let $\tilde{P}^k \equiv P + RF_h + F_h'H_hF_h$ and $\tilde{Q}^k \equiv Q + F_h'H_{\iota h}$.

From the previous section, we know that $x_t = [s_t, a_t]'$ where $s_t \equiv [1, k_{i,t-1}, n_{i,t-1}]'$ is the perfectly observable part, and a_t is the partially observed productivity. We know that the time-t estimate of x_t for the capital manager is

$$\hat{x}_t^k \equiv [s_t, \hat{a}_t^k]'$$

To see how the noisy information changes the Bellman equation, we first look at the first term $x_t'\left(P + RF_h + F_h'H_hF_h\right)x_t$. Put it in the conditional expectation, we get

$$\begin{aligned}
E\left[x_t'\left(P + RF_h + F_h'H_hF_h\right)x_t\middle|\mathcal{I}_t^k\right] &= E\left[x_t'\tilde{P}^kx_t\middle|\mathcal{I}_t^k\right] \\
&= E\left[[s_t', a_t']\begin{bmatrix}\tilde{P}_{11}^k & \tilde{P}_{12}^k \\ \tilde{P}_{21}^k & \tilde{P}_{22}^k\end{bmatrix}\begin{bmatrix}s_t \\ a_t\end{bmatrix}\middle|\mathcal{I}_t^k\right] \\
&= E\left[s_t'\tilde{P}_{11}^ks_t + s_t'\tilde{P}_{12}^ka_t + a_t'\tilde{P}_{21}^ks_t + a_t'\tilde{P}_{22}^ka_t\middle|\mathcal{I}_t^k\right] \\
&= s_t'\tilde{P}_{11}^ks_t + s_t'\tilde{P}_{12}^k\hat{a}_t^k + (\hat{a}_t^k)'\tilde{P}_{21}^ks_t + E\left[a_t'\tilde{P}_{22}^ka_t\middle|\mathcal{I}_t^k\right]
\end{aligned}$$

We knwo that

$$\begin{aligned}
E\left[a_t'\tilde{P}_{22}^ka_t\middle|\mathcal{I}_t^k\right] &= E\left[tr\left(a_t'\tilde{P}_{22}^ka_t\right)\middle|\mathcal{I}_t^k\right] \\
&= E\left[tr\left(\tilde{P}_{22}^ka_ta_t'\right)\middle|\mathcal{I}_t^k\right] = tr\left(\tilde{P}_{22}^kE\left[a_ta_t'\middle|\mathcal{I}_t^k\right]\right) \\
&= tr\left(\tilde{P}_{22}^k\left[\hat{a}_t^k(\hat{a}_t^k)' + \hat{\Sigma}_t^k\right]\right) = tr\left(\tilde{P}_{22}^k\hat{\Sigma}_t^k\right) + tr\left(\tilde{P}_{22}^k\hat{a}_t^k(\hat{a}_t^k)'\right)
\end{aligned}$$

$$= tr\left(\tilde{P}^k_{22}\hat{\Sigma}^k_t\right) + (\hat{a}^k_t)'\tilde{P}^k_{22}\hat{a}^k_t$$

Plug it back:

$$\begin{aligned}
E\left[x_t'\left(P + RF_h + F_h'H_hF_h\right)x_t\middle|\mathcal{I}^k_t\right] &= s_t'\tilde{P}^k_{11}s_t + s_t'\tilde{P}^k_{12}\hat{a}^k_t + (\hat{a}^k_t)'\tilde{P}^k_{21}s_t + E\left[a_t'\tilde{P}^k_{22}a_t\middle|\mathcal{I}^k_t\right] \\
&= s_t'\tilde{P}^k_{11}s_t + s_t'\tilde{P}^k_{12}\hat{a}^k_t + (\hat{a}^k_t)'\tilde{P}^k_{21}s_t + tr\left(\tilde{P}^k_{22}\hat{\Sigma}^k_t\right) + (\hat{a}^k_t)'\tilde{P}^k_{22}\hat{a}^k_t \\
&= tr\left(\tilde{P}^k_{22}\hat{\Sigma}^k_t\right) + [s_t', (\hat{a}^k_t)']\begin{bmatrix}\tilde{P}^k_{11} & \tilde{P}^k_{12} \\ \tilde{P}^k_{21} & \tilde{P}^k_{22}\end{bmatrix}\begin{bmatrix}s_t \\ \hat{a}^k_t\end{bmatrix} \\
&= tr\left(\tilde{P}^k_{22}\hat{\Sigma}^k_t\right) + (\hat{x}^k_t)'\tilde{P}^k\hat{x}^k_t
\end{aligned}$$

The rest of terms in the quadratic profit function is straight forward. For the second term,

$$E\left[x_t'\left(Q + F_h'H_{\iota h}\right)\iota_t\middle|\mathcal{I}^k_t\right] = (\hat{x}^k_t)'\tilde{Q}^k\iota_t$$

and the third term can be taken out of the expectation directly since ι_t is a function of elements in $\mathcal{I}^k_t$.

As a result, we have

$$\begin{aligned}
V^k(x_t) = \max_{\iota_t} tr\left(\tilde{P}^k_{22}\hat{\Sigma}^k_t\right) &+ (\hat{x}^k_t)'\left(P + RF_h + F_h'H_hF_h\right)\hat{x}^k_t + (\hat{x}^k_t)'\left(Q + F_h'H_{\iota h}\right)\iota_t + H_\iota\iota_t^2 \\
&+ \gamma E\left[V^k(x_{t+1})\middle|\mathcal{I}^k_t\right] \qquad \text{(A42)}
\end{aligned}$$

Compared with the complete-information case, we can see that the expected profit under incomplete information is different in the following aspects:

- replacing the true state variable x_t by the belief $\hat{x}^k_t$;
- adding a constant term that contains the posterior uncertainty $\hat{\Sigma}^k_t$.

Notably, the second moment of state estimate $\hat{\Sigma}^k_t$ enters the value function only via the constant term (in the trace) and does not become part of the coefficient on terms with $\hat{x}^k_t$ or ι_t. Hence, $\hat{\Sigma}^k_t$ will not affect the first-order conditions of the Bellman equation — this should not be surprising since we are doing linear-quadratic control and certainty equivalence holds.

Again, I guess that the value function takes the form

$$(\hat{x}_t^k)'\hat{P}_k\hat{x}_t^k + \hat{Q}_k$$

so that

$$\begin{aligned}
E\left[V^k(x_{t+1})\middle|\mathcal{I}_t^k\right] &= E\left[(\hat{x}_{t+1}^k)'\hat{P}_k\hat{x}_{t+1}^k + \hat{Q}_k\middle|\mathcal{I}_t^k\right] \\
&= tr\left(\hat{P}_k\hat{\Sigma}_{t+1}^k\right) + (\hat{x}_{t+1}^k)'\hat{P}_k\hat{x}_{t+1}^k + \hat{Q}_k
\end{aligned}$$

We plug it into the (A42) and take first order condition, we will end up with exactly the same F_ι that we had in (A37), the full-information benchmark. The undertermined coefficient matrix $\hat{P}_k$ in the value function is pinned down by exactly the same Riccatti equation (A38) in the full-information benchmark. The constant matrix $\hat{Q}^k$ will be different, since now it has to contain the trace term with posterior uncertainty.

Similarly, for the labor manager, the policy matrix F_h will be exactly the same as the (A39), with Riccatti equation defined by (A40).

A.4 Proof of Proposition 1

We know that

$$\begin{aligned}
\iota_t &= F_\iota^k k_{t-1} + F_\iota^n n_{t-1} + F_\iota^a \hat{a}_t^k \\
\mathbb{E}_{t-1}^k[\iota_t] &= F_\iota^k k_{t-1} + F_\iota^n \mathbb{E}_{t-1}^k[n_{t-1}] + F_\iota^a \mathbb{E}_{t-1}^k[a_t]
\end{aligned}$$

Hence

$$\begin{aligned}
\iota_t - \mathbb{E}_{t-1}^k[\iota_t] &= F_\iota^n(n_{t-1} - \mathbb{E}_{t-1}^k[n_{t-1}]) + F_\iota^a(\hat{a}_t^k - \mathbb{E}_{t-1}^k[a_t]) \\
&= F_\iota^n(h_{t-1} - \hat{h}_{t-1}^k) + F_\iota^a((I - G^kH)\varrho J^k(h_{t-1} - \hat{h}_{t-1}^k) + G^k(z_t^k - H\varrho\hat{a}_{t-1}^k)) \\
&= \left[F_\iota^n + F_\iota^a(I - G^kH)\varrho J^k\right](h_{t-1} - \hat{h}_{t-1}^k) + F_\iota^a G^k(z_t^k - H\varrho\hat{a}_{t-1}^k) \qquad \text{(A43)}
\end{aligned}$$

Since

$$\begin{aligned}
F_\iota^a G^k(z_t^k - H\varrho\hat{a}_{t-1}^k) &= F_\iota^a G^k(Ha_t + w_t^k - H\varrho\hat{a}_{t-1}^k) \\
&= F_\iota^a G^k(H\varrho a_{t-1} + H\mu_t + w_t^k - H\varrho\hat{a}_{t-1}^k)
\end{aligned}$$

$$= \underbrace{F_\iota^a G^k (H\mu_t + w_t^k)}_{\text{orthogonal to } h_{t-1}} + \underbrace{F_\iota^a G^k H \varrho (a_{t-1} - \hat{a}_{t-1}^k)}_{\text{correlates with } h_{t-1} - \hat{h}_{t-1}^k}$$

we have

$$\begin{aligned} \iota_t - \mathbb{E}_{t-1}^k[\iota_t] &= \left[F_\iota^n + F_\iota^a (I - G^k H)\varrho J^k\right](h_{t-1} - \hat{h}_{t-1}^k) + F_\iota^a G^k (H\varrho a_{t-1} + H\mu_t + w_t^k - H\varrho \hat{a}_{t-1}^k) \\ &= \left[F_\iota^n + F_\iota^a (I - G^k H)\varrho J^k\right](h_{t-1} - \hat{h}_{t-1}^k) + F_\iota^a G^k H \varrho (a_{t-1} - \hat{a}_{t-1}^k) + F_\iota^a G^k (H\mu_t + w_t^k) \end{aligned} \tag{A44}$$

A.5 Proof of Proposition 2

First, note that under common info,

$$\begin{aligned} y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}] &= (1 - 1/\epsilon)(a_{i,t-1} - \mathbb{F}_{i,t-1}[a_{i,t-1}]) + \hat{\alpha}_1 (k_{i,t-1} - \mathbb{F}_{i,t-1}[k_{i,t-1}]) + \hat{\alpha}_2 (n_{i,t-1} - \mathbb{F}_{i,t-1}[n_{i,t-1}]) \\ &= (1 - 1/\epsilon)(a_{i,t-1} - \mathbb{F}_{i,t-1}[a_{i,t-1}]) \end{aligned} \tag{A45}$$

Write every object in the regression in its MA representation.

$$a_{it} = \sum_{j=0}^{\infty} \rho^j \mu_{i,t-j} \tag{A46}$$

and

$$\begin{aligned} \mathbb{F}_{it}[a_{it}] &= G \sum_{j=0}^{\infty} \lambda^j \rho^j (1-G)^j s_{i,t-j} \\ &= G \sum_{j=0}^{\infty} \lambda^j \rho^j (1-G)^j \left(\sum_{k=0}^{\infty} \rho^k \mu_{i,t-j-k} + \epsilon_{i,t-j} \right) \\ &= G \left\{ \sum_{j=0}^{\infty} \frac{\rho^j (1 - \lambda^{j+1}(1-G)^{j+1})}{1 - \lambda(1-G)} \mu_{i,t-j} + \sum_{j=0}^{\infty} \rho^j \lambda^j (1-G)^j \epsilon_{i,t-j} \right\} \end{aligned} \tag{A47}$$

Hence the nowcast error can be written as

$$\begin{aligned} & a_{it} - \mathbb{F}_{it}[a_{it}] \\ &= \sum_{j=0}^{\infty} \rho^j \mu_{i,t-j} - G \left\{ \sum_{j=0}^{\infty} \frac{\rho^j (1 - \lambda^{j+1}(1-G)^{j+1})}{1 - \lambda(1-G)} \mu_{i,t-j} + \sum_{j=0}^{\infty} \rho^j \lambda^j (1-G)^j \epsilon_{i,t-j} \right\} \end{aligned}$$

$$= \sum_{j=0}^{\infty} \frac{(1-G)(1-\lambda) + \lambda^{j+1}(1-G)^{j+1}}{1-\lambda(1-G)} \rho^j \mu_{i,t-j} - \sum_{j=0}^{\infty} G\rho^j \lambda^j (1-G)^j \epsilon_{i,t-j} \tag{A48}$$

and hence

$$\begin{aligned} & y_{it} - \mathbb{F}_{it}[y_{it}] = (1-1/\epsilon)(a_{it} - \mathbb{F}_{it}[a_{it}]) \\ & = \sum_{j=0}^{\infty} \frac{(1-G)(1-\lambda) + \lambda^{j+1}(1-G)^{j+1}}{1-\lambda(1-G)} \rho^j (1-1/\epsilon) \mu_{i,t-j} - \sum_{j=0}^{\infty} G(1-1/\epsilon)\rho^j \lambda^j (1-G)^j \epsilon_{i,t-j} \end{aligned} \tag{A49}$$

The investment forecast error is given by

$$\begin{aligned} & \iota_{it} - \mathbb{F}_{i,t-1}[\iota_{it}] = F_\iota^a \rho G(a_{i,t-1} - \mathbb{F}_{i,t-1}[a_{i,t-1}]) + F_\iota^a G\rho(1-\lambda)\mathbb{F}_{i,t-1}[a_{i,t-1}] + L(\mu_{it}, \epsilon_{it}) \\ & = F_\iota^a \rho G \left\{ \sum_{j=0}^{\infty} \frac{(1-G)(1-\lambda) + \lambda^{j+1}(1-G)^{j+1}}{1-\lambda(1-G)} \rho^j \mu_{i,t-1-j} - \sum_{j=0}^{\infty} G\rho^j \lambda^j (1-G)^j \epsilon_{i,t-1-j} \right\} \\ & + F_\iota^a G\rho(1-\lambda)G \left\{ \sum_{j=0}^{\infty} \frac{\rho^j(1-\lambda^{j+1}(1-G)^{j+1})}{1-\lambda(1-G)} \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \rho^j \lambda^j (1-G)^j \epsilon_{i,t-1-j} \right\} + L(\mu_{it}, \epsilon_{it}) \\ & = \sum_{j=0}^{\infty} \frac{1-\lambda + (1-(1-\lambda)G)\lambda^{j+1}(1-G)^{j+1}}{1-\lambda(1-G)} F_\iota^a G\rho^{j+1} \mu_{i,t-1-j} - \sum_{j=0}^{\infty} F_\iota^a G^2 \rho^{j+1} \lambda^{j+1} (1-G)^j \epsilon_{i,t-1-j} \\ & + L(\mu_{it}, \epsilon_{it}) \end{aligned} \tag{A50}$$

Lastly we derive the MA representation of ι_{it} and h_{it}. ι_{it} and h_{it} are given by

$$\iota_{it} = F_\iota^k k_{i,t-1} + F_\iota^n n_{i,t-1} + F_\iota^a \mathbb{F}_{it}[a_{it}] \text{ and } h_{it} = F_h^k k_{i,t-1} + F_h^n n_{i,t-1} + F_h^a \mathbb{F}_{it}[a_{it}]$$

and by definition, $k_{i,t-1}$ and $n_{i,t-1}$ can be written as

$$k_{i,t-1} = \sum_{j=0}^{\infty} \iota_{i,t-1-j} \text{ and } n_{i,t-1} = \sum_{j=0}^{\infty} h_{i,t-1-j}$$

I guess that the MA representation takes the form

$$\iota_{it} = \sum_{j=0}^{\infty} a_j^k \mu_{i,t-j} + \sum_{j=0}^{\infty} b_j^k \epsilon_{i,t-j} \text{ and } h_{it} = \sum_{j=0}^{\infty} a_j^n \mu_{i,t-j} + \sum_{j=0}^{\infty} b_j^n \epsilon_{i,t-j} \tag{A51}$$

our goal is to solve sequence $\{a_j^k\}, \{a_j^n\}, \{b_j^k\}, \{b_j^n\}$.

From (A51), we get an MA representation for $k_{i,t-1}$ and $n_{i,t-1}$:

$$\begin{aligned} k_{i,t-1} &= \sum_{j=0}^{\infty} \iota_{i,t-1-j} = \sum_{s=0}^{\infty} \left\{ \sum_{j=0}^{\infty} a_j^k \mu_{i,t-1-s-j} + \sum_{j=0}^{\infty} b_j^k \epsilon_{i,t-1-s-j} \right\} \\ &= \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} a_s^k \right) \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} b_j^k \right) \epsilon_{i,t-1-j} \end{aligned} \tag{A52}$$

and

$$n_{i,t-1} = \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} a_s^n \right) \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} b_j^n \right) \epsilon_{i,t-1-j} \tag{A53}$$

Plug them into the decision rules:

$$\begin{aligned} \iota_{it} &= \sum_{j=0}^{\infty} a_j^k \mu_{i,t-j} + \sum_{j=0}^{\infty} b_j^k \epsilon_{i,t-j} \\ &= F_\iota^a G \mu_{it} + F_\iota^a G \epsilon_{it} + F_\iota^k \left\{ \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} a_s^k \right) \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} b_j^k \right) \epsilon_{i,t-1-j} \right\} \\ &+ F_\iota^n \left\{ \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} a_s^n \right) \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} b_j^n \right) \epsilon_{i,t-1-j} \right\} \\ &+ F_\iota^a G \left\{ \sum_{j=0}^{\infty} \frac{\rho^j (1 - \lambda^{j+1} (1-G)^{j+1})}{1 - \lambda(1-G)} \mu_{i,t-j} + \sum_{j=0}^{\infty} \rho^j \lambda^j (1-G)^j \epsilon_{i,t-j} \right\} \end{aligned} \tag{A54}$$

Let

$$c_j \equiv \frac{\rho^j (1 - \lambda^{j+1} (1-G)^{j+1})}{1 - \lambda(1-G)}$$

and

$$d_j \equiv \rho^j \lambda^j (1-G)^j$$

We have

$$a_0^k = F_\iota^a G + F_\iota^a G c_0, \; b_0^k = F_\iota^a G + F_\iota^a G d_0, \text{ and}$$

$$a_j^k = F_\iota^k \sum_{s=0}^{j-1} a_s^k + F_\iota^n \sum_{s=0}^{j-1} a_s^n + F_\iota^a G c_j \text{ if } j \geq 1 \implies a_j^k - a_{j-1}^k = F_\iota^k a_{j-1}^k + F_\iota^n a_{j-1}^n + F_\iota^a G (c_j - c_{j-1}) \tag{A55}$$

and

$$b_j^k = F_\iota^k \sum_{s=0}^{j-1} b_s^k + F_\iota^n \sum_{s=0}^{j-1} b_s^n + F_\iota^a G d_j \text{ if } j \geq 1 \implies b_j^k - b_{j-1}^k = F_\iota^k b_{j-1}^k + F_\iota^n b_{j-1}^n + F_\iota^a G(d_j - d_{j-1}) \tag{A56}$$

Similarly, for hiring, we have

$$\begin{aligned} h_{it} &= \sum_{j=0}^{\infty} a_j^n \mu_{i,t-j} + \sum_{j=0}^{\infty} b_j^n \epsilon_{i,t-j} \\ &= F_h^a G \mu_{it} + F_h^a G \epsilon_{it} + F_h^k \left\{ \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} a_s^k \right) \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} b_j^k \right) \epsilon_{i,t-1-j} \right\} \\ &+ F_h^n \left\{ \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} a_s^n \right) \mu_{i,t-1-j} + \sum_{j=0}^{\infty} \left(\sum_{s=0}^{j} b_j^n \right) \epsilon_{i,t-1-j} \right\} \\ &+ F_h^a G \left\{ \sum_{j=0}^{\infty} \frac{\rho^j (1 - \lambda^{j+1} (1-G)^{j+1})}{1 - \lambda(1-G)} \mu_{i,t-j} + \sum_{j=0}^{\infty} \rho^j \lambda^j (1-G)^j \epsilon_{i,t-j} \right\} \end{aligned} \tag{A57}$$

and therefore we have

$$a_0^n = F_h^a G + F_h^a G c_0, \; b_0^n = F_h^a G + F_h^a G d_0, \text{ and}$$

$$a_j^n = F_h^k \sum_{s=0}^{j-1} a_s^k + F_h^n \sum_{s=0}^{j-1} a_s^n + F_h^a G c_j \text{ if } j \geq 1 \implies a_j^n - a_{j-1}^n = F_h^k a_{j-1}^k + F_h^n a_{j-1}^n + F_h^a G(c_j - c_{j-1}) \tag{A58}$$

and

$$b_j^n = F_h^k \sum_{s=0}^{j-1} b_s^k + F_h^n \sum_{s=0}^{j-1} b_s^n + F_h^a G d_j \text{ if } j \geq 1 \implies b_j^n - b_{j-1}^n = F_h^k b_{j-1}^k + F_h^n b_{j-1}^n + F_h^a G(d_j - d_{j-1}) \tag{A59}$$

We have

$$\begin{bmatrix} a_j^k \\ a_j^n \end{bmatrix} = \begin{bmatrix} 1 + F_\iota^k & F_\iota^n \\ F_h^k & 1 + F_h^n \end{bmatrix} \begin{bmatrix} a_{j-1}^k \\ a_{j-1}^n \end{bmatrix} + \begin{bmatrix} F_\iota^a G \\ F_h^a G \end{bmatrix} (c_j - c_{j-1}) \tag{A60}$$

and

$$\begin{bmatrix} b_j^k \\ b_j^n \end{bmatrix} = \begin{bmatrix} 1 + F_\iota^k & F_\iota^n \\ F_h^k & 1 + F_h^n \end{bmatrix} \begin{bmatrix} b_{j-1}^k \\ b_{j-1}^n \end{bmatrix} + \begin{bmatrix} F_\iota^a G \\ F_h^a G \end{bmatrix} (d_j - d_{j-1}) \tag{A61}$$

The sequence of coefficients can be solved from these recursions.

When there is no adjustment costs, $\xi_k = \xi_n = 0$, so that $F_\iota^a = F_h^a = \epsilon - 1$, $F_\iota^k = F_h^n = -1$ and $F_\iota^n = F_h^k = 0$. In this case, we have

$$a_j^k = a_j^n = F_\iota^a G \frac{\rho^j(1-\lambda^{j+1}(1-G)^{j+1}) - \rho^{j-1}(1-\lambda^j(1-G)^j)}{1-\lambda(1-G)}$$

$$b_j^k = b_j^n = F_\iota^a G(\rho^j\lambda^j(1-G)^j - \rho^{j-1}\lambda^{j-1}(1-G)^{j-1})$$

Assume that $\rho\lambda(1-G) < 1$, so that all the geometric sum exists. We have

$$\begin{aligned}
&\mathbb{E}[(y_{it} - \mathbb{F}_{it}[y_{it}])^2] \\
&= \left[\frac{(1-\lambda)^2}{(1-\rho^2)} + \frac{2\lambda(1-\lambda)}{(1-\rho^2\lambda(1-G))} + \frac{\lambda^2}{(1-\rho^2\lambda^2(1-G)^2)}\right]\frac{(1-1/\epsilon)^2(1-G)^2}{(1-\lambda(1-G))^2}\sigma_\mu^2 \\
&+ \frac{G^2(1-1/\epsilon)^2}{1-\rho^2\lambda^2(1-G)^2}\sigma_\epsilon^2 \\
&= \left[(1-\lambda)\left(\frac{1-\lambda}{x} + \frac{\lambda}{y}\right) + \lambda\left(\frac{1-\lambda}{y} + \frac{\lambda}{z}\right)\right]\frac{(1-1/\epsilon)^2(1-G)^2}{(1-\lambda(1-G))^2}\sigma_\mu^2 + \frac{G^2(1-1/\epsilon)^2}{1-\rho^2\lambda^2(1-G)^2}\sigma_\epsilon^2
\end{aligned} \tag{A62}$$

$$\begin{aligned}
&\mathbb{E}[(\iota_{it} - \mathbb{F}_{i,t-1}[\iota_{it}])(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])] \\
&= \sigma_\mu^2 \frac{F_\iota^a G(1-1/\epsilon)(1-G)\rho}{(1-\lambda(1-G))^2}\left[\frac{(1-\lambda)^2}{1-\rho^2} + \frac{(1-\lambda)\lambda((1-(1-\lambda)G)(1-G)+1)}{1-\rho^2\lambda(1-G)}\right. \\
&\left. + \frac{(1-(1-\lambda)G)\lambda^2(1-G)}{1-\rho^2\lambda^2(1-G)^2}\right] + \sigma_\epsilon^2 F_\iota^a G^3(1-1/\epsilon)\frac{\rho\lambda}{1-\rho^2\lambda^2(1-G)^2} \\
&= \sigma_\mu^2 \frac{F_\iota^a G(1-1/\epsilon)(1-G)\rho}{(1-\lambda(1-G))^2}\left[(1-\lambda)\left(\frac{1-\lambda}{x} + \frac{\lambda}{y}\right) + \lambda(1-G)(1-(1-\lambda)G)\left(\frac{1-\lambda}{y} + \frac{\lambda}{z}\right)\right] \\
&+ \sigma_\epsilon^2 \frac{F_\iota^a G^3(1-1/\epsilon)\rho\lambda}{1-\rho^2\lambda^2(1-G)^2}
\end{aligned} \tag{A63}$$

$$\begin{aligned}
&\mathbb{E}[(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])\iota_{i,t-1}] \\
&= \sigma_\mu^2 \frac{F_\iota^a G(1-G)(1-1/\epsilon)\rho^{-1}}{(1-\lambda(1-G))^2}\left[(1-\lambda)\left(\frac{\rho-1}{x} + \frac{1-\rho\lambda(1-G)}{y}\right) + \lambda\left(\frac{\rho-1}{y} + \frac{1-\rho\lambda(1-G)}{z}\right)\right] \\
&+ F_\iota^a G^2(1-1/\epsilon)\frac{\rho^{-1}\lambda^{-1}(1-G)^{-1}(1-\rho\lambda(1-G))}{1-\rho^2\lambda^2(1-G)^2}\sigma_\epsilon^2
\end{aligned} \tag{A64}$$

$$= \frac{F_\iota^a G(1-G)(1-1/\epsilon)}{(1-\lambda(1-G))^2} \left[1 - \lambda(1-G) + (\rho-1)\rho\left(\frac{1-\lambda}{x} + \frac{(1-G)\lambda^2}{y}\right) + \rho\lambda(1-G)(1-\rho\lambda(1-G))\left(\frac{1-\lambda}{y} + \frac{\lambda^2(1-G)}{z}\right)\right] \tag{A65}$$

$$\mathbb{E}[(\iota_{it} - \mathbb{F}_{i,t-1}[\iota_{it}])\iota_{i,t-1}]$$
$$= \frac{(F_\iota^a)^2 G^2}{(1-\lambda(1-G))^2} \left[(\rho(1-\lambda) + (1-(1-\lambda)G)(1-G)\rho\lambda)(1-\lambda(1-G)) + (1-\lambda)\rho^2\left(\frac{\rho-1}{x} + \frac{(1-G)\lambda(1-\rho\lambda(1-G))}{y}\right) + \rho^2\lambda^2(1-G)^2(1-(1-\lambda)G)\left(\frac{\rho-1}{y} + \frac{\lambda(1-G)(1-\rho\lambda(1-G))}{z}\right)\right] \tag{A66}$$

where

$$x \equiv 1-\rho^2, y \equiv 1-\rho^2\lambda(1-G), z \equiv 1-\rho^2\lambda^2(1-G)^2$$

We know that the regression coefficient is

$$\tilde{\beta}_\iota = \frac{\mathbb{E}[(y_{it-1} - \mathbb{F}_{it-1}[y_{it-1}])^2]\mathbb{E}[(\iota_{it} - \mathbb{F}_{i,t-1}[\iota_{it}])\iota_{i,t-1}] - \mathbb{E}[(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])\iota_{i,t-1}]\mathbb{E}[(\iota_{it} - \mathbb{F}_{i,t-1}[\iota_{it}])(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])]}{\mathbb{E}[(y_{it-1} - \mathbb{F}_{it-1}[y_{it-1}])^2]\mathbb{E}[\iota_{i,t-1}^2] - \mathbb{E}[(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])\iota_{i,t-1}]^2} \tag{A67}$$

To determine the sign of $\tilde{\beta}_\iota$, we investigate the sign of the numerator and the denominator. The numerator can be written in the form of

$$A\sigma_\mu^4 + B\sigma_\epsilon^4 + C\sigma_\mu^2\sigma_\epsilon^2$$

where A, B, C are functions of $\rho, \lambda, G, \epsilon, F_\iota^a$.

We analyze the structure of A, B, C.

First, since

$$A$$
$$= \frac{(F_\iota^a)^2(1-1/\epsilon)^2(1-G)^2G^2}{(1-\lambda(1-G))^4}\lambda(1-\lambda)G\left[(1-G)\lambda - (2-G)\right]\left[\lambda(\rho-1) - (1-\lambda)(1-\rho\lambda(1-G))\right]\left[\frac{1}{y^2} - \frac{1}{xz}\right]$$
$$\propto -\lambda(1-\lambda)\left[(1-G)\lambda - (2-G)\right]\left[\lambda(\rho-1) - (1-\lambda)(1-\rho\lambda(1-G))\right] \tag{A68}$$

We have the following:

- If $\lambda < 1$, then $1-\lambda > 0$, $(1-G)\lambda < 2-G$, and $\lambda(\rho-1) - (1-\lambda)(1-\rho\lambda(1-G)) < 0$, hence $A < 0$;

- If $\lambda = 1$, then $A = 0$;
- If $\lambda > 1$, it's a big complicated:
  - We know that $1 - \lambda < 0$ for sure;
  - The term $(1-G)\lambda - (2-G)$ will flip to positive if $\lambda > \frac{2-G}{1-G} > 2$.
  - $f(\lambda) \equiv \lambda(\rho - 1) - (1-\lambda)(1 - \rho\lambda(1-G))$ is a quadratic function on λ. We have $f(1) = \rho - 1 = f\left(\frac{1}{1-G}\right)$ and $f'(1) > 0$. Hence, there exists $\bar{M}_A > 1$ such that $\forall \lambda \in (1, \bar{M}_A)$, $f(\lambda) < 0$.

  Hence, if we take $1 < \lambda < \min\{\bar{M}_A, \frac{2-G}{1-G}\}$, we have $(1-G)\lambda - (2-G) < 0$ and $f(\lambda) < 0$, and therefore $sgn(A) = sgn(\lambda - 1) > 0$.

Second, we check the sign of B:

$$
\begin{aligned}
&B \\
&= \frac{G^2(1-1/\epsilon)^2}{1-\rho^2\lambda^2(1-G)^2} \frac{(F_\iota^a)^2 G^3(1-\rho\lambda(1-G))}{(1-G)(1-\rho^2\lambda^2(1-G)^2)} - \frac{F_\iota^a G^2(1-1/\epsilon)(1-\rho\lambda(1-G))}{(1-\rho^2\lambda^2(1-G)^2)\rho\lambda(1-G)} \frac{F_\iota^a G^3(1-1/\epsilon)\rho\lambda}{1-\rho^2\lambda^2(1-G)^2} \\
&= 0
\end{aligned} \tag{A69}
$$

Lastly, we check the sign of C:

$$
\begin{aligned}
C &= \frac{(1-\lambda)(F_\iota^a)^2(1-1/\epsilon)^2 G^3}{(1-\rho^2\lambda^2(1-G)^2)(1-\lambda(1-G))^2} \cdot \left[\frac{(1-(1-G)\lambda)^2\lambda^{-1}(\rho\lambda-1)}{1-\rho^2} - \frac{(1-\rho\lambda(1-G))(1-G)^3\lambda}{1-\rho^2\lambda^2(1-G)^2} + \frac{(1-G)(1-\lambda(1-G))[(1-G)(\rho\lambda-1)-1+\rho\lambda(1-G)]}{1-\rho^2\lambda(1-G)} \right] \\
&\propto (1-\lambda)(1-\lambda(1-G))(\rho\lambda-1)z\left[(1-\lambda(1-G))y + (1-G)^2 x\lambda\right]\lambda^{-1} - (1-\lambda)(1-\rho\lambda(1-G))(1-G)x\left[(1-G)^2\lambda y + (1-\lambda(1-G))z\right]
\end{aligned} \tag{A70}
$$

We have the following:

- If $\lambda < 1$, then the first term is negative, and the second term is positive, hence $C < 0$;
- If $\lambda = 1$, then $C = 0$;
- If $1 < \lambda < \min\{\rho^{-1}, (1-G)^{-1}\}$, then $[(1-\lambda(1-G))y + (1-G)^2 x\lambda] > 0$, $1 - \lambda < 0$, $1 - \lambda(1-G) > 0$, $\rho\lambda - 1 < 0$ so that the first term is positive, and $[(1-G)^2\lambda y + (1-\lambda(1-G))z] > 0$ so that the second term is negative. Hence $C > 0$ in this region.

Hence, if we take $\bar{M} = \min\{\bar{M}_A, \rho^{-1}, (1-G)^{-1}\}$, then as long as $\lambda \leq \bar{M}$, we have $sgn(\tilde{\beta}_\iota) = sgn(\tilde{\beta}_h) = sgn(\lambda - 1)$.

Lastly, the denominator of $\tilde{\beta}_\iota$ and $\tilde{\beta}_h$ is given by

$$\mathbb{E}[(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])^2]\mathbb{E}[h_{i,t-1}^2] - \mathbb{E}[(y_{i,t-1} - \mathbb{F}_{i,t-1}[y_{i,t-1}])h_{i,t-1}]^2$$

which is positive by Cauchy-Schwartz inequality.

B Data Appendix

This appendix shows the details of sample construction and variable definition for the expectation test and the calibration exercise.

B.1 Construction of IBES-Compustat Dataset

Prepare the CRSP-Compustat Dataset After downloading the CRSP-Compustat Annual Dataset from the WRDS website, I apply the following filters:

- Keep only the firms incorporated in the US (`fic = "USA"`);
- Keep only the firms using USD as accounting units (`curncd = "USD"`);
- Drop the observations with missing NAICS codes;
- Drop the observations with 2-digit NAICS codes = 52 (finance) and = 22 (utilities);
- Drop the observations with negative sales;
- Drop the observations with negative employment;
- Drop the observations with missing capital expenditure;
- Drop the observations with missing sales.
- Drop the observations with missing employment.
- Remove the duplicate (LPEMNO, fyear) combinations.

I end up getting a CRSP-Compustat Annual sample with 81,937 firm-year observations from 2000 to 2024.

Prepare the IBES-Compustat Joint Sample After downloading the IBES Guidance Dataset from WRDS, I apply the following filters and imputations:

- Keep only the annual guidance (`pdicity = "ANN"`);
- Keep only the sales and capital expenditure guidances (`measure = "CPX" or "SAL"`);
- Keep only the observations with range guidance (`range_desc = 1`) or point estimates (`range_desc = 2`). For the firms that adopt range guidance, drop the observations with missing lower or upper bound, and impute the forecast/nowcast by the midpoint of the lower and upper bound (`(val_1 + val_2)/2`).
- Merge the IBES-CRSP bridge file to the IBES sample by matching the IBES ticker (firm identifier in IBES datasets). This step allows us to map the IBES firm identifier (IBES ticker) to CRSP-Compustat firm identifier (PERMNO).
- Drop all the guidances with forecasting horizon $>$ 12 months.
- Up to now, the observations in the IBES Guidance dataset are at firm-year-guidance level. Drop the observation if it is the only guidance posted by the firm in that year. We are left with the guidances posted by the firms that make multiple guidances in a fiscal year.
- Drop the guidances that (1) are not the first guidance posted by the firm in a fiscal year, or (2) are not the last guidance posted by the firm in a fiscal year.
- Now, We are left with only the first and the last guidance posted by a firm in a fiscal year. I use the first guidance as "forecast" and the last guidance as "nowcast" for each firm and fiscal year. Up to now, each observation in the IBES Guidance dataset is at firm-year level.
- Merge the IBES Guidance dataset to the CRSP-Compustat dataset above and remove the observations with (1) missing capital expenditure forecasts, (2) missing lagged sales nowcasts, or (3) missing lagged hirings.

Summary Statistics We end up with a IBES-Compustat merged sample. This merged sample spans over 2000 and 2024 and contains 9201 firm-year observations. This is a sample for gigantic firms in the US. Table B1 compares the average and median size of firms in the IBES-Compustat sample with the whole CRSP-Compustat sample between 2000 and 2024. We can see that the firms in our IBES-Compustat sample is about two times larger in sales and employment on average. Median sales in the IBES-Compustat sample is about five times larger than the CRSP-Compustat firms. Table B2 tabulates the share of total sales in the US GDP for the firms in the samples. We can see that the total sales of the firms in our IBES-Compustat sample take up about 20 percent of US GDP, depite the fact that it only keeps about 10 percent of the CRSP-Compustat sample.

		CRSP-Compustat (2000-2024)		
Variables	Units	Mean	Median	No. Obs
Sales	Millions USD	3512.01	366.11	81937
Capital Stock (PPEGT)	Millions USD	2464.81	140.32	78478
Capital Expenditure	Millions USD	210.11	11.59	81937
Employment	Thousands	10.78	1.18	81937
		IBES-Compustat (2000-2024)		
Variables	Units	Mean	Median	No. Obs
Sales	Millions USD	6996.31	1848.92	9201
Capital Stock (PPEGT)	Millions USD	3609.12	784.72	9119
Capital Expenditure	Millions USD	286.34	64.54	9201
Employment	Thousands	24.92	7.19	9201

Table B1: Summary Statistics (2000-2024)

	CRSP-Compustat Sample			
	2005	2010	2015	2020
Total Sales (Billions USD)	8812	10057	11957	13413
Sales/GDP	68%	67%	65%	63%
	IBES-Compustat Sample			
	2005	2010	2015	2020
Total Sales (Billions USD)	1666	2954	3534	4146
Sales/GDP	13%	20%	19%	19%

Table B2: Total Sales and Share of Sales in US GDP

Construction of Variables for the Expectation Test The key variables in the expectation tests in Section 3 are constructed as follows. For the forecast and nowcast errors, I use the the arc-percent difference between the forecast/nowcast value and the actual value, as in Bloom, Codreanu, and Fletcher (2025):

$$\text{CapEX Forecast Error}_{it} = \frac{\text{CapEX}_{it} - \mathbb{E}^k_{i,t-1}[\text{CapEX}_{it}]}{1/2(\text{CapEX}_{it} + \mathbb{E}^k_{i,t-1}[\text{CapEX}_{it}])}$$

and

$$\text{Sales Nowcast Error}_{it} = \frac{\text{Sales}_{it} - \mathbb{E}^k_{i,t}[\text{Sales}_{it}]}{1/2(\text{Sales}_{it} + \mathbb{E}^k_{i,t}[\text{Sales}_{it}])}$$

The arc-difference definition of forecast/nowcast errors has two advantages. First, it is bounded between $[-2, 2]$, so that outliers in the realized values and expectations would not affect the forecast/nowcast error and the regression results by much. Additionally, it avoids the problems of taking logs when the actual values or expectations are close to 0. This may not be a big issue for the large-firm samples like IBES or Compustat, and I verify that the regression results are robust to alternative definitions of forecast/nowcast errors (e.g. the log difference between predicted values and actual values).

The definition of hiring is standard. For firm i in year t, hiring h_{it} is defined as

$$h_{it} \equiv \log(\text{emp}_{it}) - \log(\text{emp}_{i,t-1})$$

where `emp` is the employment of the firm in CRSP-Compustat. Alternatively, an arc-difference measure of hiring is widely adopted in the literature as well (e.g. Davis and Haltiwanger (1992)). The regression results do not change with alternative definitions of hiring.

B.2 Robustness Checks for the Expectation Test

In this section, I consider several robustness checks to the expectation test in Section 3.

Time-to-build In the capital adjustment literature, it is common to assume that the capital takes time to build (e.g. Cooper and Haltiwanger (2006), David and Venkateswaran (2019)): today's investment transforms into capital stock with a one-period lag. One may wonder whether time-to-build invalidates the power of empirical test: namely, whether incorporating time-to-build in capital alone makes the coefficient of interest $\tilde{\beta}_h \neq 0$.

I verify that time-to-build in capital does not invalidate the regression test. In our framework of Dynamic Narrow Framing, the only change is on the profit function: without time-to-build, we have

$$\Pi(K, N; A, K_{-1}, N_{-1}) = \underbrace{AK^{\alpha(1-1/\epsilon)}N^{(1-\alpha)(1-1/\epsilon)}}_{=\mathcal{R}(A,K,N)} - RK - WN - \text{adj. costs}$$

With time-to-build on capital, it seems that the only change is on the profit function:

$$\Pi(K, N; A, K_{-1}, N_{-1}) = A{K_{-1}}^{\alpha(1-1/\epsilon)}N^{(1-\alpha)(1-1/\epsilon)} - RK - WN - \text{adj. costs} \tag{B1}$$

In other words, with the time-to-build friction, capital for production today is predetermined, and the cost for tomorrow's capital is paid today. This friction changes the matrices in the log-quadratic profit function (A3): now the matrix $P, Q, R, H_\iota, H_h, H_{\iota h}$ becomes

$$P \equiv \begin{bmatrix} \frac{1}{2\epsilon}\bar{Y} & 0 & 0 & \frac{1}{2}(1-1/\epsilon)\bar{Y} \\ * & \frac{1}{2}\bar{R}\bar{K}(\alpha(1-1/\epsilon)-1) & \frac{1}{2}\alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y} & \frac{1}{2}\alpha(1-1/\epsilon)^2\bar{Y} \\ * & * & \frac{1}{2}\bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1) & \frac{1}{2}(1-\alpha)(1-1/\epsilon)^2\bar{Y} \\ * & * & * & \frac{1}{2}(1-1/\epsilon)^2\bar{Y} \end{bmatrix} \tag{B2}$$

$$Q \equiv \begin{bmatrix} -\bar{R}\bar{K} \\ -\bar{R}\bar{K} \\ 0 \\ 0 \end{bmatrix} \text{ and } R \equiv \begin{bmatrix} 0 \\ \alpha(1-\alpha)(1-1/\epsilon)^2\bar{Y} \\ \bar{W}\bar{N}((1-\alpha)(1-1/\epsilon)-1) \\ (1-\alpha)(1-1/\epsilon)^2\bar{Y} \end{bmatrix} \tag{B3}$$

and

$$H_\iota \equiv -\frac{1}{2}(\bar{R}\bar{K} + \xi_k\bar{K}), H_h \equiv -\frac{1}{2}([1-(1-\alpha)(1-1/\epsilon)]\bar{W}\bar{N} + \xi_n\bar{N}) \text{ and } H_{\iota h} = 0 \tag{B4}$$

The learning process of firm managers is not changed by the time-to-build friction, and therefore the intuition we built in Section 3 remains true: lagged hiring cannot forecast the investment forecast error.

Non-convex Adjustment Cost In Progress.

B.3 Construction of Compustat Sample for Calibration

After obtaining COMPUSTAT Fundamentals Annual dataset from WRDS, I applied the following filter:

1. Exclude the observations before 1998;
2. Exclude the firms incorporated or legally registered outside the US (filter by COMPUSTAT terminology `fic`);
3. Exclude the firms that do not use US dollar as their native currency (filter by COMPUSTAT terminology `curncd`);
4. Following the literature (e.g. De Loecker, Eeckhout, and Unger (2020), Flynn and Sastry (2020), Afrouzi, Drenik, and Kim (2023), among many others), exclude firms in the following sectors
   (a) financial sector (SIC code 6000-6999, or 2-digit NAICS 52);
   (b) utility sector (SIC code 4900-4999, or 2-digit NAICS 22);
   (c) non-operating sector (SIC code 9995).

   It is well documented that financial firms are different from non-financial firms in many aspects (e.g. production function, business cycle properties, etc). The input and output prices of the utility sector is heavily regulated.
5. Exclude observations with missing SIC or NAICS codes;
6. Exclude observations with missing total assets (COMPUSTAT terminology `at`);
7. Exclude observations with acquisitions (COMPUSTAT terminology `aq`) exceeding 5 percent of total assets (COMPUSTAT terminology `at`);
   This is a common practice in the literature because firms that went through large mergers and acquisitions tend to have very large changes in input choices and measured productivity. [10]
8. Exclude observations with non-positive sales (COMPUSTAT terminology `sale`);
9. Exclude observations with non-positive employments (COMPUSTAT terminology `emp`);

[10] I may slack this filter if one believes that major M&As are (exogenous) variations to the management style and hence information friction.

10. (Optional; for robustness check) Construct the capital stock series by perpetual inventory method. As pointed out in Becker et al. (2005) and Bachmann and Bayer (2014), the capital measure in COMPUSTAT, namely PPEGT, is book value. It cannot be directly used because it is valued at historical prices and the accounting depreciation in the balance sheet is usually higher than the economic depreciation for tax benefit. I construct the capital stock series as follows:

    (a) For each firm, the initial capital stock K_{i,t_0} is obtained by deflating the first observation of total gross property, plant and equipment (COMPUSTAT terminology `ppegt`), denoted BV_{i,t_0}; for the deflator P_{t_0}, I use the non-residential gross private domestic investment deflator (row 9) as in Implicit Price Deflators for Gross Domestic Product (NIPA Table 1.1.9) from BEA website [in this table, the base year is 2017]. Hence, initial capital stock is

    $$K_{i,t_0} = \frac{P_{2017}}{P_{t_0}} BV_{i,t_0}$$

    (b) For each firm, build up the capital recursively by

    $$K_{i,t} = I_t - \delta K_{i,t-1} + K_{i,t-1} = NI_{i,t} + K_{i,t-1} \tag{B5}$$

    where NI denotes net investment. There are multiple ways of measuring net investment. Here I use the approach in Ottonello and Winberry (2020): deflate the series of total net property, plant and equipment (COMPUSTAT terminlogy `ppent`) and use $\Delta PPENT_{i,t} \equiv PPENT_{i,t} - PPENT_{i,t-1}$ as a measure of (undeflated) net investment. Deflate $\Delta PPENT_{i,t}$ by the corresponding deflator P_t, we obtain

    $$K_{i,t} = K_{i,t-1} + \frac{P_{2017}}{P_t} \Delta PPENT_{i,t} \tag{B6}$$

11. Exclude observations with missing cost of goods sold.

12. The following steps are essential to backing out the capital share α:

    (a) merge compustat with County Business Pattern (CBP) from Census (which contains industry-level wages - here, industry = 3-digit NAICS for manufacturing and 2-digit for other industries);

    (b) fill up the missing values in wage bill for Compustat firms: wage bill = emp * CBP$_{wages}$;

(c) compute the material expenditure = cogs + xsga - dp - wage bill (see Keller and Yeaple (2009) and Flynn and Sastry (2020)).

(d) compute value-added = sales - material expenditure;

(e) compute the labor share in sales:

$$\text{rshare}_{\text{labor}} = \sum(wagebill)/\sum(\text{value-added})$$

Here, I sum up all the firm-year observations to get an aggregate share; in an alternative exercise I do the summation up to the industry level and get an industry-specific labor share for all the industries;

(f) back out the capital share of income:

$$\alpha = 1 - rshare_{labor}/(1 - 1/\epsilon)$$

13. I prepare the all the variables that will be used in the quantitative analysis:

(a) Define capital stock as the book value of PPE (PPEGT);

(b) take log of value-added, employment and capital stock, get the lagged and first-differenced log employment and log capital stock;

(c) Define the profit shock/Sales Solow residual as

$$a = \log(valueadded) - \alpha * (1 - 1/\epsilon)\log(capitalstock) - (1 - \alpha) * (1 - 1/\epsilon) * \log(emp)$$

(d) get the lagged and first-differenced series for a;

(e) define mrpk = log(value added) - log(capital stock) and mrpn = log(value added) - log (emp).

14. I append the cleaned IBES dataset to the cleaned Compustat sample. Each firm-year observation in the cleaned IBES dataset has the current year's first guidance of capital expenditure and the previous year's last guidance of sales. Call the resulting sample "IBES/Compustat Merged Quantitative" Sample.

15. In this "IBES/Compustat Merged Quantitative" Sample, I regress $k, n, a, \Delta k, \Delta n, \Delta a, mrpk, mrpn$ on sector-year fixed effects and keep the residuals (this is to remove the trends in the data).

16. I trim the 1 percent tail of the above residuals;

17. I trim the 1 percent tails of investment forecast errors and sales nowcast errors.

This results in a sammple that has 6492 firm-year observations. Table B3 reports the summary statistics.

Table B3: Summary Statistics of the Quantitative Sample

Variables	Units	Mean	Median	No. Obs
Sales	Millions USD	5319.80	1769.03	6492
Capital Stock (PPEGT)	Millions USD	2260.60	734.10	6492
Capital Expenditure	Millions USD	186.60	61.00	6492
Employment	Thousands	16.74	6.45	6492