ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

Attention and a Paradox of Uncertainty

Authors	Yu-Ting Chiang
Working Paper Number	2022-004D
Revision Date	March 2022
Citable Link	https://doi.org/10.20955/wp.2022.004
Suggested Citation	Chiang, Y.-T., 2022; Attention and a Paradox of Uncertainty, Federal Reserve Bank of St. Louis Working Paper 2022-004. URL https://doi.org/10.20955/wp.2022.004

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

Attention and a Paradox of Uncertainty

Yu-Ting Chiang[†]

October 8, 2024

Abstract

I show that macroeconomic uncertainty during recessions can arise from people paying more attention to aggregate events. When information is dispersed, people's attempts to acquire more information can lead to higher aggregate volatility, forecast dispersion, and uncertainty about aggregate output. Information rigidity is reduced, consistent with evidence in forecast surveys, and distinct from the prediction of exogenous volatility shocks. When the model is calibrated to U.S. data, endogenous attention accounts for half of the observed fluctuations in volatility, forecast dispersion, and uncertainty. I also provide a method to solve models with varying attention and uncertainty under an infinite regress problem due to dispersed information.

Keywords: business cycles, information frictions, uncertainty and volatility, expectation formation

JEL code: D8, E1, E3, E7

[†]Federal Reserve Bank of St. Louis, contact email: yu-ting.chiang@stls.frb.org. This paper is a revised version of the first chapter of my dissertation at the University of Chicago. I thank Mikhail Golosov, Lars Hansen, Fernando Alvarez, and Yueran Ma for their guidance and advice. I appreciate valuable discussions with George-Marios Angeletos, Anmol Bhandari, Harris Dellas, Zhen Huo, Ezra Karger, Stijn Van Nieuwerburgh, Hyunmin Park, B. Ravikumar, Edouard Schaal, Wenting Song, Mikayel Sukiasyan, Venky Venkateswaran, Kai-Hao Yang, Mark Wright, and Piotr Żoch. Views in this paper do not necessarily represent those of the Federal Reserve System.

1 Introduction

Macroeconomic uncertainty is central to the business cycle: It increases during downturns and crises and also plays a major role in driving economic activity and asset prices. The conventional view is that it can be mitigated when people have more information. While this seems true in many circumstances, there may be important exceptions. Identifying such exceptions could prove important for understanding not only the behavior of uncertainty but also the business cycle more broadly.

In this paper, I identify one such exception. During periods when uncertainty matters more (e.g., recessions), people have an incentive to acquire more information to mitigate its effects. However, if information is dispersed, an attempt to acquire more information can prove counterproductive: paying more attention to aggregate events can lead to an *increase* in aggregate volatility and uncertainty about aggregate output. This results from the uncertainty each individual faces about other people's actions, which manifests in higher forecast dispersion among individuals.

I consider a standard dispersed information model in which agents observe an unknown aggregate state with idiosyncratic noise and can increase the precision of their information with costly attention. They make production decisions based on their expectations of the aggregate state and beliefs about other agents' actions. Production features strategic complementarity: The more others produce, the more each agent wants to produce. The choice of attention is subject to an *income effect*: When agents expect low income, they pay more attention to avoid mistakes because the marginal cost of making mistakes is higher. Due to dispersed information, agents face uncertainty not only about the aggregate state but also about the endogenous actions of others.

I show that such countercyclical variation in attention can generate countercyclical uncertainty, volatility, and forecast dispersion over the business cycle. As agents pay more attention, their production decisions respond strongly to shocks, and their collective response increases aggregate volatility. If the economy exhibits strong strategic complementarity and a low initial level of attention, an increase in attention has counter-intuitive implications: When agents act on different information, their expectations diverge, and each one faces higher uncertainty about the endogenous responses of other agents. As a result, uncertainty about aggregate outcomes increases along with forecast dispersion. This is despite the fact that all agents have learned more about the exogenous aggregate state.

This attention-based mechanism accounts for a closely related business cycle phenomenon, namely, a reduction in *information rigidity* during recessions: Relative to the movements in aggregate variables, expectation updates are larger during recessions than in normal times. This implication of endogenous attention is useful because it contrasts with the implications of exogenous volatility shocks — the standard mechanism of generating uncertainty fluctuations. An exogenous increase in volatility, as long as it is not perfectly observed, shifts aggregate variables more than people's expectations. As a result, higher uncertainty due to exogenous volatility is associated with increased information rigidity. This is distinct from the case where uncertainty arises from endogenous attention.

I first use a static model to establish these results analytically and then a dynamic model to quantify these mechanisms. The dynamic model allows for both endogenous attention and exogenous volatility shocks to aggregate productivity. I calibrate the model so that: the size of volatility shocks corresponds to its empirical counterpart; variation in information rigidity matches that in forecast surveys, quantifying the extent to which attention fluctuates over the business cycle. In the calibrated model, the two mechanisms together fully account for the observed fluctuations in measures of uncertainty, volatility, and forecast dispersion. As in the data, standard deviations of these measures are 40% to 50% of their long-run averages. A decomposition of the two mechanisms shows that, without exogenous volatility shocks, agents' endogenous attention response can account for about half of the observed variation in uncertainty, volatility, and forecast dispersion.

Finally, this paper develops a new method for solving higher-order dynamics of dispersed information models. The persistence of dispersed information in these models is essential for generating key empirical features of macroeconomic expectation updates. However, these models feature an infinite regress problem as each agent's decision can potentially depend on the entire history of signals they received in the past. Existing solution methods rely on first-order approximations and miss the fluctuations in attention and uncertainty, as these fluctuations are higher-order features of the models. The method developed in this paper captures these nonlinear dynamics. More broadly, the method allows for the use of higher-order moments in the calibration and evaluation of models with dispersed information.

Literature

My framework builds on the dispersed information and rational inattention literature, following Phelps (1970), Lucas (1972), and Sims (2003). Most works in the literature feature a static information structure. Woodford (2001), Lorenzoni (2009, 2010), Angeletos and La'O (2010, 2013), Angeletos and Lian (2018), and Angeletos and Huo (2021) feature a static exogenous information structure. Information acquisition is endogenous in Maćkowiak and Wiederholt (2015, 2009), but due to linear-quadratic approximation, their information structure remains static. I depart from the static information structure and show that countercyclical information acquisition over the business cycle can explain a broad set of phenomena related to macroeconomic uncertainty.

The attention-based mechanism in this paper connects countercyclical uncertainty to the study of macroeconomic expectation formation. Coibion and Gorodnichenko (2015), and more recently Goldstein (2023), use data from the U.S. forecast surveys to provide evidence of reduced information rigidity during recessions, indicating countercyclical attention. This evidence is further supported by Song and Stern (2020) and Flynn and Sastry (2023) using a text-analysis approach. Flynn and Sastry (2023) proceeds to generate countercyclical volatility with a behavioral constraint, abstracting away from expectation formation. In this paper, I adhere to Bayesian rationality and establish that countercyclical volatility, forecast dispersion, and uncertainty about aggregate outcomes are intrinsically linked to how people update their expectations over the business cycle, consistent with evidence of reduced information rigidity during recessions.

Countercyclical uncertainty has been the focus of extensive literature following Bloom (2009). The literature has provided several mechanisms to generate uncertainty fluctuations. One strand relies on exogenous volatility shocks, such as Bloom et al. (2018) and, more closely related to this paper, Nimark (2014) and Kozeniauskas, Orlik, and Veldkamp (2018) in dispersed information models. Another strand generates countercyclical uncertainty through procyclical learning, including Van Nieuwerburgh and Veldkamp (2006), Fajgelbaum, Schaal, and Taschereau-Dumouchel (2017).[1] I show that countercyclical information acquisition, when quantified by the variation in information rigidity, can serve as a quantitatively important source of countercyclical uncertainty.

Finally, the method developed in this paper uses perturbation techniques to address issues that arise from solving dispersed information models with state-dependent attention choices. The technique is related to Lombardo and Uhlig (2018), Borovicka and Hansen (2013), and Bhandari et al. (2021) that utilize small shock expansions. A challenge unique to dispersed information models comes from the infinite regress problem studied by Townsend (1983), Kasa (2000), Lorenzoni (2009), Nimark (2017), Huo and Pedroni (2020), Chahrour and Jurado (2023), and Huo and Takayama (2023). Due to infinite regress, individual decisions depend on an infinite-dimension history of signals. Existing works focus on the linear dynamics of dispersed information models. I provide a method to solve higher-order dynamics in these models and show they are useful for understanding a broad set of phenomena related

[1] Benhabib, Liu, and Wang (2016) and Straub and Ulbricht (2023) also feature decreasing information during recessions, but focus on uncertainty about idiosyncratic shocks.

to macroeconomic uncertainty and expectation formation.

2 Model

The economy consists of a continuum of agents indexed by $i \in [0,1]$. Each agent produces a unique intermediate good with labor. A representative final good producer combines intermediate goods to produce a final good that agents consume. Agents and the final good producer are price-takers. The aggregate productivity of the final good producer is unknown to agents when they decide on labor inputs. Depending on agents' expectations about the aggregate productivity, they can pay attention and acquire information about it. The economy proceeds in three stages. Agents make their attention choices in stage 1. Labor inputs are decided in stage 2. Final goods are produced and consumed in stage 3.

Preferences and Technology

Agents derive utility from final good consumption $c_i \in \mathbb{R}^+$ and disutility from labor and attention $n_i, z_i \in \mathbb{R}^+$:

$$u(c_i, n_i) - \kappa z_i.$$

The payoff from c_i and n_i takes a modified Greenwood–Hercowitz–Huffman (GHH) form:

$$u(c_i, n_i) = \frac{1}{1-\tilde{\gamma}} \left(\max\left\{ c_i - \frac{n_i^{1+\nu}}{1+\nu},\ \underline{u} \right\} \right)^{1-\tilde{\gamma}},$$

where a lower bound $\underline{u} > 0$ ensures that preferences are well-defined for all realizations of consumption after any labor input choice. Marginal disutility from attention is given by a constant κ. Parameter ν is the inverse Frisch elasticity of labor supply. Parameter $\tilde{\gamma}$ plays a dual role: Besides governing the relative risk aversion over realizations of consumption-labor bundles, it controls how agents trade off attention z_i and payoffs from c_i and n_i.[2]

Each agent produces a unique intermediate good using labor with linear technology, $q_i = n_i$, where q_i denotes the quantity of intermediate good i. Agents face budget constraints $c_i \leq p_i q_i$, where p_i is the relative price of intermediate good i with the final good as the numeraire.

A representative final good producer produces final good Y with intermediate goods $\{y_i\}$ to maximize profit $Y - \int p_i y_i \, di$. The production function of the final good producer is given

[2]As I discuss in Section 3.3 and Appendix B, the main result does not rely on the GHH form.

by a constant elasticity of substitution (CES) production function:

$$Y = e^{\bar{\theta}+\theta} \left(\int y_i^{1-\eta} \, di \right)^{\frac{1}{1-\eta}},$$

where $\eta \in [0,1)$ is the inverse elasticity of substitution between intermediate goods. Aggregate productivity is stochastic and consists of two components, $\bar{\theta}$ and θ, representing the initial condition of the economy and its subsequent development. These components are independent and normally distributed with mean zero and variances $\sigma_{\bar{\theta}}^2$ and σ_{θ}^2. The sequence of events begins with the realization of $\bar{\theta}$, followed by agents' responses, and finally the realization of θ. The key exercise in the following analysis studies how variation in initial condition $\bar{\theta}$ affects agents' attention choices, and how that affects their production decisions. In essence, variation in $\bar{\theta}$ reflects how the economic condition fluctuates over the business cycle. This fluctuation over time will be modeled explicitly in the dynamic framework in Section 5.

Timing and Information

The economy proceeds in three stages. In stage 1, agents observe a signal $x = \bar{\theta} + \epsilon$ about the initial condition $\bar{\theta}$ with common noise $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. They form beliefs about $\bar{\theta}$ and choose attention z_i.

In stage 2, each agent receives an idiosyncratic signal x_i about shock θ with precision z_i:

$$x_i = \theta + \frac{\epsilon_i}{\sqrt{z_i}}, \quad \epsilon_i \overset{iid}{\sim} \mathcal{N}(0,1).$$

Given information set $\mathcal{F}_i := \boldsymbol{\sigma}(x_i, x)$, agents form beliefs about the price of their product, p_i, and choose labor input n_i.

In stage 3, the final good producer combines intermediate goods to produce the final good. Agents receive the final good as proceeds from selling intermediate goods, and they consume subject to budget constraints. Prices $\{p_i\}$ realize to clear the markets.

Definition of Equilibrium

An equilibrium is a collection of random variables $\{z_i, n_i, q_i, p_i, c_i, y_i, Y\}$ such that (i) z_i optimizes agents' expected utility, given signal x; (ii) n_i optimizes each agent's expected utility, given signals x_i, x; (iii) c_i is optimal subject to budget constraints; (iv) the final good producer chooses $\{y_i\}$ to maximize profit, given prices $\{p_i\}$; (v) productions of $\{q_i\}$ and Y are given by respective technologies; and (vi) markets clear: $y_i = q_i, \ \forall i$ and $Y = \int c_i \, di$.

Since agents are ex-ante identical, I focus on a symmetric equilibrium in which $\log z_i = \boldsymbol{z}(x)$ and $\log n_i = \boldsymbol{n}(x, x_i)$ for some functions $\boldsymbol{z}, \boldsymbol{n}$.

2.1 Equilibrium Characterization

Consider first the final good producer's decision in stage 3. The final good producer takes prices $\{p_i\}$ as given and chooses $\{y_i\}$ to maximize profit. Their maximization problem leads to the standard CES demand for intermediate goods: $p_i = e^{(1-\eta)(\bar{\theta}+\theta)}\ Y^{\eta} y_i^{-\eta}$.

Given labor input $\{n_i\}$ chosen by the agents, market clearing and production feasibility imply the equilibrium price of intermediate good i can be solved as a function of productivity, individual input n_i, and an aggregate input N:

$$p_i = e^{\bar{\theta}+\theta} N^{\eta} n_i^{-\eta}, \quad \text{where} \quad N := \left(\int n_i^{1-\eta}\ di \right)^{\frac{1}{1-\eta}}. \tag{1}$$

Combining the expression for p_i and the budget constraint, we have

$$c_i = e^{\bar{\theta}+\theta} N^{\eta} n_i^{1-\eta}.$$

The price of intermediate good p_i (relative to the final good) increases with productivity because more of the final good is produced when productivity is high; it increases with N because intermediate goods are complementary, and the value of good i is higher when other agents produce more; it decreases with n_i because the marginal value of good i decreases with the quantity produced. Consumption c_i increases with n_i despite a decrease in p_i because the elasticity of substitution between intermediate goods is greater than one.

In stage 2, agents form beliefs about price p_i based on their expectations about aggregate productivity and input. Optimality of input $n_i \in \mathcal{F}_i$ requires

$$\mathbb{E}\left[(1-\eta) e^{\bar{\theta}+\theta} N^{\eta} n_i^{-\eta} \frac{u_c(c_i, n_i)}{\mathbb{E}[\ u_c(c_i, n_i)\ |\mathcal{F}_i]} \middle| \mathcal{F}_i \right] = n_i^{\nu}, \tag{2}$$

The condition requires that agents equalize the expected marginal product of labor input, weighted by the normalized marginal utility of consumption, to their marginal disutility of labor.

In stage 1, given other agents' attention choices, each agent chooses z_i so that the marginal value of attention equals the marginal cost of attention κ:

$$\int u(c_i, n_i) \frac{\partial}{\partial z_i} \varphi(\bar{\theta}, \theta, x_i | x, z_i)\ d\bar{\theta}\, d\theta\, dx_i = \kappa, \tag{3}$$

where φ is the Gaussian density function of $\bar{\theta}, \theta, x_i$, given x and z_i. Each agent chooses z_i, understanding its effect on the precision of signal x_i and how it affects their subsequent decisions on labor input n_i and consumption $c_i = e^{\bar{\theta}+\theta} N^{\eta} n_i^{1-\eta}$. The equilibrium requires the optimality condition to hold when n_i is given by the equilibrium function $\boldsymbol{n}(\cdot)$, following the envelope theorem. The attention choice of each agent depends on other agents' attention through aggregate input N.

The following lemma summarizes the characterization above:

Lemma 1 *An equilibrium is given by functions* $\{\boldsymbol{z}, \boldsymbol{n}, \boldsymbol{N}\}$, *such that* $\log z_i = \boldsymbol{z}(x)$, $\log n_i = \boldsymbol{n}(x, x_i)$ *and* $\log N = \boldsymbol{N}(x, \theta)$ *solve Equations 1, 2, and 3.*

Proof. See Appendix A.1. □

2.2 Equilibrium Approximation

The equilibrium conditions in Lemma 1 constitute a fixed-point problem that does not generally have a closed-form solution. Therefore, I proceed with an approximation of the equilibrium. The approximation modifies the standard perturbation method to address issues emerging from the attention choice.

I consider a sequence of economies indexed by a perturbation parameter δ that scales the size of the shocks, noises, and attention cost:

$$\bar{\theta}(\delta) = \delta\bar{\theta}, \quad \theta(\delta) = \delta\theta, \quad \epsilon_i(\delta) = \delta\epsilon_i, \quad \kappa(\delta) = \delta^2\kappa.$$

When $\delta = 1$, the sequence corresponds to the economy to be approximated. The limit $\delta \to 0$ corresponds to a deterministic economy with a vanishing attention cost and no shocks, which can be solved easily.[3] While the scaling of shocks and noises is standard, scaling the attention cost at rate δ^2 is a crucial feature of the perturbation scheme. This particular scaling is useful for approximating attention choice problems because, as $\delta \to 0$, the marginal value of information is second order: Deviation of input from its deterministic optimal level has no first-order effect on agents' payoff. By scaling the marginal cost of attention $\kappa(\delta)$ at rate δ^2, the approximation scheme keeps the attention choice non-degenerate along the perturbation sequence as $\delta \to 0$.[4] This perturbation contains the common linear-quadratic approximation for information acquisition problems as a special case and generalizes to higher-order approximations.

[3] I focus on the case in which $\bar{c} - \frac{\bar{n}^{1+\nu}}{1+\nu} > \underline{u}$ when $\delta = 0$.

[4] Scaling of the attention cost creates a "bifurcation point" at $\delta = 0$. See Judd (1998) for a discussion on bifurcation in the context of approximating a portfolio choice problem.

Equilibrium objects are approximated by Taylor expansion with respect to δ:

$$\log z(\delta) \approx \log \bar{z} + \hat{z}\delta, \quad \log n_i(\delta) \approx \bar{n} + \hat{n}_i\delta + \frac{\hat{\hat{n}}_i\delta^2}{2}, \quad \log N(\delta) \approx \bar{N} + \hat{N}\delta + \frac{\hat{\hat{N}}\delta^2}{2},$$

where $\bar{N}$, $\hat{N}$, and $\hat{\hat{N}}$ denote the zeroth-, first-, and second-order expansion of $\log N(\delta)$ with respect to δ, and similarly for other variables. As an example, the first-order expansion of $\log N$ is given by:

$$\hat{N} = \frac{d}{d\delta}\boldsymbol{N}(x(\delta), \theta(\delta), \delta)\big|_{\delta=0} = \boldsymbol{N}_x x + \boldsymbol{N}_\theta \theta + \boldsymbol{N}_\delta,$$

where $\boldsymbol{N}_x, \boldsymbol{N}_\theta, \boldsymbol{N}_\delta$ are derivatives of function $\boldsymbol{N}(\cdot)$ at $\delta \to 0$.

The expansions are solved by differentiating equilibrium conditions in Lemma 1 with respect to δ to appropriate orders and evaluating at $\delta = 0$. Details of the method are discussed in Appendix A. In this paper, I solve the equilibrium up to second-order approximation.

3 Fluctuations in Attention and Uncertainty

I characterize the equilibrium and show that a worsening of initial condition $\bar{\theta}$ can generate an increase in attention, which leads to higher aggregate volatility, forecast dispersion, and the uncertainty each agent faces about aggregate output. I then derive its implications on the measures of information rigidity.

3.1 Aggregate Input

Given agents' attention choice, aggregate input is determined by the optimality of input (Equation 2) and the aggregation condition (Equation 1). First-order approximations of the equilibrium conditions are:

$$\hat{n}_i = \mathbb{E}[r(\bar{\theta} + \theta) + s\hat{N}|\bar{\mathcal{F}}_i], \quad \hat{N} = \int \hat{n}_i \, di.$$

These conditions describe how $\hat{n}_i$ responds to aggregate variables and how $\hat{N}$ depends on individual input. Information sets $\bar{\mathcal{F}}_i$ is evaluated at the average attention level $\bar{z}$, containing signal $\hat{x}_i = \theta + \epsilon_i/\sqrt{\bar{z}}$. Parameters

$$r := \frac{1}{\eta + \nu} > 0, \quad s := \frac{\eta}{\eta + \nu} \in [0, 1]$$

depend on preference and technology: r describes the direct response to aggregate productivity $\bar{\theta} + \theta$; s captures the response to aggregate input $\hat{N}$, representing the level of strategic complementarity. Solving the system gives:

$$\hat{N} = \boldsymbol{N}_x x + \boldsymbol{N}_\theta \theta.$$

Coefficient $\boldsymbol{N}_x$ describes how much aggregate input responds to initial condition $\bar{\theta}$ through signal x, which does not depend on attention choice; coefficient $\boldsymbol{N}_\theta$ captures the response to changes in productivity, θ. It depends on the average attention level $\bar{z}$ through Kalman gain λ:

$$\boldsymbol{N}_\theta = \frac{r\lambda}{1 - s\lambda}, \quad \lambda := \frac{\sigma_\theta^2}{\sigma_\theta^2 + 1/\bar{z}} \in [0, 1].$$

When $\bar{z}$ is higher, aggregate input is responsive to θ as agents update expectations with more precise signals. Moreover, there is a feedback loop due to strategic complementarity: When aggregate input varies more with θ, each agent's input responds more. These two forces are represented by λ in the numerator and denominator of $\boldsymbol{N}_\theta$.

While first-order approximations describe how the level of attention $\bar{z}$ affects aggregate input, they fail to capture how the economy behaves as agents adjust their attention. The second-order expansions capture the effect of attention response, $\hat{z}$:

$$\hat{\hat{n}}_i = \underbrace{2 \times \frac{d}{d\delta}\mathbb{E}[r\theta + s\ \hat{N}|\mathcal{F}_i(\delta)]\Big|_{\delta=0}}_{\text{(i)}} + \underbrace{s\mathbb{E}[\hat{\hat{N}}|\bar{\mathcal{F}}_i]}_{\text{(ii)}} + \upsilon_0, \quad \hat{\hat{N}} = \int \hat{\hat{n}}_i\ di + \upsilon_1. \tag{4}$$

Individual input $\hat{\hat{n}}_i$ responds to (i) changes in beliefs about aggregate conditions due to attention response and (ii) the effect of these changes on aggregate input, $\hat{\hat{N}}$. Constants υ_0 and υ_1 represent how agents respond to the average level of uncertainty and input dispersion, both of which are given by the first-order approximations and the average level of attention $\bar{z}$.[5] How agents' beliefs change with attention response in (i) is given by:

$$\frac{d}{d\delta}\mathbb{E}[r\theta + s\hat{N}|\mathcal{F}_i(\delta)]\Big|_{\delta=0} = \boldsymbol{N}_\theta\Big((1-\lambda)\hat{z}\hat{x}_i - \frac{\hat{z}}{2\sqrt{\bar{z}}}\epsilon_i\Big).$$

The first term on the right-hand side says that when agents pay more attention, $\hat{z} > 0$, they rely more on $\hat{x}_i$ to update beliefs. The second term shows that when agents pay more attention, it reduces idiosyncrasy in their signals. Solving $\hat{\hat{N}}$ from Equation 4 and combining it with the solution for $\hat{N}$ gives the following lemma:

[5] Changes in uncertainty affect input in third-order approximations. I abstract from these effects as numerous works have studied how uncertainty affects output.

Lemma 2 *Up to second-order approximation,*

$$\log N \approx \boldsymbol{N}_x x + \boldsymbol{N}_\theta \Big(1 + \frac{1-\lambda}{1-s\lambda} \times \hat{z}\Big)\theta + const.$$

Proof. See Appendix A.2. □

In comparison to the first-order approximation, attention response $\hat{z}$ affects how aggregate input reacts to changes in productivity θ: When agents pay more attention, $\hat{z} > 0$, input response to θ is stronger. Other things equal, this effect is more pronounced when the economy features stronger strategic complementarity (higher s), because each agent's response triggers stronger feedback from other agents when the coordination motive is strong.

3.2 Attention Response

Given how input choices depend on attention, I solve the equilibrium level of attention, $\bar{z}$, and attention response $\hat{z}$, by expanding the optimality condition of attention (Equation 3) to corresponding orders. Lemma 3 summarizes a key implication on how equilibrium attention responds to initial condition $\bar{\theta}$:

Lemma 3 *Parameter $\tilde{\gamma}$ determines the direction of attention response:*

$$\tilde{\gamma} \gtreqless 1 \iff \frac{\partial \hat{z}}{\partial \bar{\theta}} \lesseqgtr 0.$$

Proof. See Appendix A.2. □

Lemma 3 reveals two competing forces that shape attention response to $\bar{\theta}$: an income effect and a substitution effect of expected productivity on attention. To understand these effects, consider a decrease in expected productivity (due to a decrease in $\bar{\theta}$). The substitution effect comes from a decrease in the marginal rate of transformation between attention and agents' payoff: When agents expect low productivity, they lower the level of labor input; consequently, a 1% mistake in input decision is less costly, and agents have less incentive to pay attention. On the other hand, the income effect comes from an increase in the marginal rate of substitution between attention and the consumption-labor composite. Agents expect lower income when productivity is low. With lower income, the expected marginal utility from consumption and labor is high, and agents have more incentive to pay attention and avoid making mistakes. When $\tilde{\gamma} > 1$, the income effect on attention dominates, and agents pay more attention when the initial condition worsens. The assumption of GHH preference over consumption and labor separates the income effect on attention from that on labor.[6] This separation allows for procyclical fluctuations in labor input while permitting attention to be countercyclical, consistent with empirical evidence discussed in Section 4.

3.3 Uncertainty, Volatility, and Forecast Dispersion

As agents pay attention, their input decisions respond strongly to aggregate conditions. Due to dispersed information, agents' reactions not only generate large movements in output but also affect the uncertainty each agent faces about other agents' responses. This mechanism has implications for several key features of the business cycle, which are captured by the following measures of uncertainty, volatility, and forecast dispersion:

Definition *Let* $\widetilde{Y} \coloneqq \log Y$.

1. *Aggregate volatility* : $SD(\widetilde{Y}|\bar{\theta}) \coloneqq \left(\mathbb{E}\left[\left(\widetilde{Y} - \mathbb{E}[\widetilde{Y}|\bar{\theta}]\right)^2 \big| \bar{\theta}\right]\right)^{\frac{1}{2}}$.
2. *Forecast dispersion* : $Disp(\mathbb{E}_i[\widetilde{Y}]) \coloneqq \left(\int \left(\mathbb{E}[\widetilde{Y}|\mathcal{F}_i] - \int \mathbb{E}[\widetilde{Y}|\mathcal{F}_i]di\right)^2\right)^{\frac{1}{2}}$.
3. *Subjective uncertainty* : $SD(\widetilde{Y}|\mathcal{F}_i) \coloneqq \left(\mathbb{E}\left[\left(\widetilde{Y} - \mathbb{E}[\widetilde{Y}|\mathcal{F}_i]\right)^2 \big| \mathcal{F}_i\right]\right)^{\frac{1}{2}}$.

Aggregate volatility is the conditional output volatility given $\bar{\theta}$. It reflects how much aggregate out reacts to changes in the aggregate state, θ, given the initial condition $\bar{\theta}$. Forecast dispersion captures the different views among agents about aggregate output. Subjective uncertainty represents agents' uncertainty about output based on their individual information set. These measures have close empirical counterparts, and they are well-known to exhibit countercyclical fluctuations over the business cycle.

To isolate the effects of attention on these measures, I consider a *fixed-attention economy* in which attention is fixed exogenously at its average level $\bar{z}$:

Definition *An equilibrium of a fixed-attention economy solves Equations 2 and 3, with attention given by* $z_i = \bar{z}, \forall i$.

The fixed-attention economy provides a relevant benchmark. Without attention response, initial condition $\bar{\theta}$ does not affect how agents respond to changes in aggregate condition θ. As a result, $\bar{\theta}$ has no effects on the measures of uncertainty, volatility, forecast dispersion:

Lemma 4 $SD(\widetilde{Y}|\bar{\theta})$*,* $Disp(\mathbb{E}_i[\widetilde{Y}])$*, and* $SD(\widetilde{Y}|\mathcal{F}_i)$ *are constant in* $\bar{\theta}$ *up to second-order*

[6] Parameter $\tilde{\gamma}$ plays a dual role in governing the income effect on attention and relative risk aversion. I provide one way to isolate these roles with two parameters in Section 5.

approximation in a fixed-attention economy.

Proof. See Appendix A.2. ☐

In contrast to this benchmark, when attention is endogenous, all fluctuations in these measures can be attributed to agents' attention responses. Theorem 1 provides conditions under which countercyclical attention generates countercyclical fluctuations in uncertainty, volatility, and forecast dispersion:

Theorem 1 *Suppose that $\tilde{\gamma} > 1$ (attention is countercyclical). Up to second-order approximation,*

$$\text{(1)} \quad \frac{\partial}{\partial \bar{\theta}} SD(\widetilde{Y}|\bar{\theta}) < 0.$$

Moreover, there exists a threshold $\zeta \in \mathbb{R}$ such that if $\bar{z} < \zeta$, then

$$\text{(2)} \quad \frac{\partial}{\partial \bar{\theta}} Disp(\mathbb{E}_i[\widetilde{Y}]) < 0, \quad \text{and} \quad \text{(3)} \quad \frac{\partial}{\partial \bar{\theta}} SD(\widetilde{Y}|\mathcal{F}_i) < 0.$$

The threshold $\zeta > 0$ if and only if $r > \frac{1}{2}$, and $\zeta \to \infty$ as $s \to 1$. In addition, (2) is true only if (3) is true.

Proof. See Appendix A.3. ☐

In response to a decrease in $\bar{\theta}$, agents pay more attention as the income effect on attention dominates: $\tilde{\gamma} > 1$. As agents pay attention, their collective response generates large movements in aggregate output, increasing aggregate volatility. This connection between agents' attention and aggregate volatility follows directly from Lemma 2.

Whether increased attention generates higher uncertainty and forecast dispersion depends on two competing channels. On the one hand, as agents pay attention, they reduce idiosyncratic noises in their signals and receive more accurate information about the aggregate state. This reduces dispersion in their forecasts and lowers their uncertainty about output. On the other hand, information has stronger effects on agents' beliefs as they pay attention. Discrepancies in signals lead to a larger dispersion in forecasts as a result. Moreover, each agent faces uncertainty about how other agents will respond due to dispersed information. This uncertainty is aggravated when all agents pay attention and respond strongly to aggregate conditions. Therefore, each agent can face higher uncertainty about aggregate output despite having learned more about the exogenous changes in θ. An increase in uncertainty of this kind is always accompanied by increased forecast dispersion among agents.

The strength of the two channels depends on the initial level of attention $\bar{z}$. Threshold ζ represents the point below which the second channel dominates and is determined by the preference and technology of the economy. If $r > \frac{1}{2}$, aggregate input has a strong enough effect on output relative to exogenous productivity, and the threshold ζ is positive. The second channel is stronger when the level of strategic complementarity s is higher. As shown in Lemma 2, high strategic complementarity implies that each agent's reaction triggers strong feedback from other agents, and attention response $\hat{z}$ has a large effect on aggregate input. In fact, $\zeta \to \infty$ as $s \to 1$ means that agents' attention always leads to greater uncertainty if the economy features strong enough strategic complementarity.

Dispersed information plays a crucial role in Theorem 1. Without dispersed information, an increase in attention leads to higher aggregate volatility but always reduces agents' uncertainty about output. This is because, if all agents receive the same signal, they face only *fundamental uncertainty* about the exogenous states, which is reduced when the signal becomes more precise. By contrast, when information is dispersed, agents also face *strategic uncertainty* about other agents' actions, which can increase as a result of agents' attempts to acquire more information.

Generality of the Result

At the core of Theorem 1 is a key observation that when information is dispersed, more information can generate increased volatility, forecast dispersion, and uncertainty about aggregate outcomes. This observation is general: it does not rely on the specific reason that motivates agents to pay more attention, and it applies to a large class of models with different assumptions on preferences, technologies, and the sources of shocks.

I demonstrate the generality of this result with a companion theorem (Theorem 1A) in Appendix B for a class of generalized "beauty-contest games" of the following form:

$$\mathbb{E}\big[G(a + \eta\tilde{N}, \tilde{n}_i)\big|\mathcal{F}_i\big] = 0, \quad H(\tilde{N}) = \int H(\tilde{n}_i)di.$$

As shown by Angeletos and Lian (2016), this class of games encompasses some of the most common setups in macroeconomic models. I demonstrate that, with certain restrictions on the function G, an increase in information (precision z) can generate uncertainty, volatility, and forecast dispersion under conditions similar to those stated in Theorem 1. I provide examples in Appendix B.1 that fit into this class of models, featuring alternative assumptions on preferences (allowing for an income effect on labor) and technology (allowing for interactions among agents through roundabout production). Finally, to illustrate that the observation is not specific to the shock considered in Section 2, I show in Appendix B.2 that the same observation applies to an economy featuring extrinsic "sentiment" shocks as in Angeletos and La'O (2013).

3.4 Implications on Expectation Updates

The mechanism in Theorem 1 is closely connected to how agents update their expectations over the business cycle. The connection is useful as it distinguishes this attention-based mechanism from the common mechanism of generating uncertainty fluctuations with exogenous volatility shocks. To illustrate this distinction, consider the following generalization of the model:

Generalization (volatility shocks): Suppose that aggregate productivity contains a cross term $\alpha\bar{\theta}\theta$, with the production function is given by:

$$Y = e^{\bar{\theta}+\theta+\alpha\bar{\theta}\theta} \left(\int y_i^{1-\eta}\, di\right)^{\frac{1}{1-\eta}}.$$

With $\alpha = 0$, the production function reduces to that in Section 2. When $\alpha < 0$, aggregate productivity features countercyclical volatility: $SD(\bar{\theta} + \theta + \alpha\bar{\theta}\theta|\bar{\theta})$ decreases with $\bar{\theta}$. This provides an alternative mechanism for the phenomena in Theorem 1: A decrease in $\bar{\theta}$ increases the volatility of productivity exogenously, and leads to higher uncertainty, volatility, and forecast dispersion of aggregate output without endogenous attention response.

However, the two mechanisms have distinct empirical implications on agents' expectation updates. This distinction is captured by the measure of *information rigidity* from Coibion and Gorodnichenko (2015):

$$\beta_{CG}(\bar{\theta}) \coloneqq \frac{Cov(\tilde{Y} - \bar{\mathbb{E}}[\tilde{Y}], \bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})}{Var(\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})}, \tag{5}$$

where $\bar{\mathbb{E}}[\cdot] = \int \mathbb{E}[\cdot|\mathcal{F}_i]di$.

Measure β_{CG} represents how much agent updates their expectations relative to the size of movements in aggregate output. The measure goes to zero if agents incorporate all available information, $\tilde{Y} - \bar{\mathbb{E}}[\tilde{Y}] \to 0$. It goes to infinity if signals x_i are perfectly uninformative, and agents do not update their expectations: $\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x] \to 0$. The measure varies with the initial state $\bar{\theta}$ as it affects agents' attention choices and the volatility of aggregate productivity. Its empirical counterpart is the regression coefficient of average forecast errors, $\tilde{Y} - \bar{\mathbb{E}}[\tilde{Y}]$, on average forecast revision, $\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]$, conditional on $\bar{\theta}$. Lemma 5 shows how endogenous attention and exogenous volatility shocks affect the measure of information rigidity β_{CG} as a function of $\bar{\theta}$:

Lemma 5 *Up to second-order approximation,*

$$\beta_{CG}(\bar{\theta}) \approx \frac{1-\lambda}{\lambda}\Big(1 + \big[-\lambda_x\phi_z + (1-\lambda_x)\phi_\alpha\big] \times \bar{\theta}\Big),$$

where $\lambda_x := \frac{\sigma^2_{\bar{\theta}}}{\sigma^2_{\bar{\theta}}+\sigma^2_\epsilon} \in [0,1]$, $\phi_z := \frac{\partial}{\partial\bar{\theta}}\hat{z}$, *and* $\phi_\alpha \lesseqgtr 0$ *if* $\alpha \lesseqgtr 0$.

Proof. See Appendix A.4. □

Without endogenous attention and volatility shocks, $\phi_z = \phi_\alpha = 0$, the measure $\beta_{CG} = \frac{1-\lambda}{\lambda}$ represents a constant level of information rigidity: A constant fraction λ of the variation in θ is captured by agents' expectation updates, and the other $1-\lambda$ fraction remains in the forecast error. By contrast, if attention is countercyclical, $\phi_z < 0$, a decline in initial condition $\bar{\theta}$ induces agents to pay more attention. Measure β_{CG} decreases as agents incorporate a larger fraction of the variation in θ into their expectation updates, and the size of expectation updates increases relative to the movements in aggregate output. In comparison, $\phi_\alpha < 0$ if aggregate productivity features countercyclical volatility. A decrease in $\bar{\theta}$ makes productivity more volatile and generates larger movements in output. This exogenous increase in volatility generally leads to a higher information rigidity because, rationally, agents do not fully incorporate the increase in volatility into their expectation updates, except for the knife-edge case in which agents have perfect information about $\bar{\theta}$, namely, $\lambda_x = 1$. When movements in output increase more than agents' expectation updates, the measure of information rigidity increases. The following corollary summarizes the distinct implications of the two mechanisms:

Corollary 1 *If* $\tilde{\gamma} > 1$ *and* $\alpha = 0$, *then* $\beta_{CG}(\bar{\theta})$ *is increasing in* $\bar{\theta}$. *By contrast, in a fixed-attention economy with* $\alpha < 0$, $\beta_{CG}(\bar{\theta})$ *is decreasing in* $\bar{\theta}$.

These distinct implications on information rigidity are useful because they allow me to separate the two mechanisms and assess their quantitative importance in generating macroeconomic uncertainty over the business cycle.

4 Empirics

I present empirical evidence in preparation for a quantitative assessment of the mechanism in Theorem 1. I show the pattern of expectation updates that indicate countercyclical attention, and I construct measures of uncertainty, volatility, and forecast dispersion, which exhibit countercyclical fluctuations. These facts are prominent business cycle phenomena studied extensively. I follow the literature in constructing these measures, with minor modifications so that the evidence can be linked directly to the dynamic model in Section 5 and the quantitative assessment in Section 6. Details are provided in Appendix E.

4.1 Attention and Expectation Updates

I construct measures of information rigidity using the forecasts of GDP growth in the Survey of Professional Forecasters (SPF). Similar to Equation 5, the measure is given by the regression coefficient of average forecast errors on average forecast revisions. In the regression, I allow for an interaction term with an indicator of aggregate condition to capture how the measure varies over time:

$$\overline{FE}_{t,h} = \alpha_{CG}^{\intercal} \begin{pmatrix} 1 \\ \mathbf{1}_t^R \end{pmatrix} + \begin{pmatrix} \beta_{CG} & \Delta\beta_{CG} \end{pmatrix} \begin{pmatrix} \overline{FR}_{t,h} \\ \mathbf{1}_t^R \times \overline{FR}_{t,h} \end{pmatrix} + residual_{t,h}. \tag{6}$$

Average forecast error at period t for output growth between period t and $t+h$ is represented by $\overline{FE}_{t,h} \coloneqq \Delta\widetilde{Y}_{t,h} - \overline{\mathbb{E}}_t[\Delta\widetilde{Y}_{t,h}]$; average forecast revision from the previous period is given by $\overline{FR}_{t,h} \coloneqq \overline{\mathbb{E}}_t[\Delta\widetilde{Y}_{t,h}] - \overline{\mathbb{E}}_{t-1}[\Delta\widetilde{Y}_{t,h}]$. I consider two specifications for the interaction term. The first specification uses the NBER recession periods for the indicator $\mathbf{1}_t^R$ and the second specification uses whether output in the previous period is below trend as the indicator, where output is band-pass filtered with a frequency corresponding to 6-32 quarters. In both specifications, the indicators represent periods of worsening aggregate conditions.

Table 1: Measure of Information Rigidity

	Indicator ($\mathbf{1}_t^R$)	
	NBER recession	below trend
β_{CG}	0.56	0.73
	(0.17)	(0.20)
$\Delta\beta_{CG}$	-0.57	-0.24
	(0.32)	(0.26)

Sample: 1968Q3 to 2019Q4; forecasts horizons: 0 to 3 quarters ahead; robust standard errors in parentheses. Table 6 in Appendix E shows a similar pattern with alternative sample periods and specifications of low output periods.

Table 1 shows the measure of information rigidity and how it changes with the indicator of worsening aggregate conditions. With the indicator of NBER recession, the estimate drops from $\beta_{CG} = .56$ in normal periods to zero during recession periods, a difference of $\Delta\beta_{CG} = -.57$; with an indicator of below-trend output, the estimate reduces by $\Delta\beta_{CG} = -.24$ between above-trend and below-trend periods. This pattern is consistent with the finding of Coibion and Gorodnichenko (2015). As shown in Lemma 5, reduced information rigidity is consistent with agents actively paying attention and incorporating new information to update their expectations. In Table 7 in Appendix E, I extend the analysis to an alternative measure of information rigidity proposed by Goldstein (2023). The alternative measure demonstrates a similar pattern of expectation updates at the individual forecaster level: Forecasters who expect lower aggregate output exhibit reduced information rigidity. In Section 6.1, I use these empirical patterns of expectation updates to quantify the extent to which agents' attention varies over the business cycle.[7]

4.2 Uncertainty, Volatility, and Forecast Dispersion

I construct empirical measures of uncertainty, volatility, and forecast dispersion corresponding to the measures studied in Section 3.3. A short description is provided for each measure, and key moments of their business cycle fluctuations are presented below.

Aggregate volatility (σ_t^Y): I estimate the conditional heteroskedasticity of quarterly real GDP growth with an EGARCH(1,1)-ARMA(1,1) model, which captures how shocks to aggregate output feed into its volatility. The conditional volatility is related to the measures in Jurado, Ludvigson, and Ng (2015), Ilut, Kehrig, and Schneider (2018), and Adrian, Boyarchenko, and Giannone (2019), among others. I adopt a simple univariate GARCH specification because it does not rely on variables absent in the model and allows a direct comparison between data and model in Section 6. I follow the same procedure to estimate the conditional volatility of aggregate productivity (σ_t^{TFP}) for the exogenous volatility shocks, as in Bloom et al. (2018).

Forecast dispersion (d_t^Y): I calculate forecast dispersion as the cross-sectional standard deviation of one-quarter-ahead estimates of GDP growth from the SPF each quarter. Countercyclical forecast dispersion has been documented by previous works such as Bachmann, Elstner, and Sims (2013), Bloom (2014), and Kozeniauskas, Orlik, and Veldkamp (2018).

Subjective uncertainty (υ_t^Y): I measure subjective uncertainty about aggregate output with

[7] Although estimates in Table 7 are more precise, they require a dynamic structure absent in the theoretical analysis. For consistency, I use moments in Table 1 for the calibration in Section 6.1 and compare alternative measures (Goldstein, 2023) between model and data in Table 7.

the probability-range data from the SPF. The survey asks each forecaster to assign probability weights to different ranges of possible GDP growth. This data has been used to document countercyclical uncertainty, for example, by Bloom (2014) and Fajgelbaum, Schaal, and Taschereau-Dumouchel (2017). Following Engelberg, Manski, and Williams (2009), I fit a parametric distribution to the discrete probability weights submitted by each forecaster. I calculate the standard deviation of the distribution and average across forecasters.[8]

Table 2: Uncertainty, Volatility, and Forecast Dispersion

	σ_t^Y	d_t^Y	υ_t^Y	σ_t^{TFP}
$cor(\cdot, \tilde{Y}_t)$	$-.39$	$-.40$	$-.33$	$-.34$
sd/avg	.47	.37	.36	.18

Sample: 1968Q3 to 2019Q4, detrended with a band-pass filter at 6-32 quarters frequency; υ_t^Y available since 1981Q2.

The first three columns in Table 2 show how measures of uncertainty, volatility, and forecast dispersion correlate with aggregate output, along with the magnitude of fluctuations in these measures. The three measures are all negatively correlated with output, and the magnitudes of fluctuations are large: the standard deviation of these measures over the business cycle frequency ranges from around 40% to 50% relative to the long-run average of each respective measure. The last column shows the conditional volatility of aggregate productivity. As a common mechanism of generating uncertainty fluctuations, the volatility of aggregate productivity is also countercyclical. Yet, the magnitude of fluctuations, relative to its long-run average, is only half that of the other three measures.[9]

To what extent can endogenous attention and volatility shocks account for the fluctuations in uncertainty, volatility, and forecast dispersion of aggregate output is a quantitative question. I now extend the model to a dynamic framework, allowing for both endogenous attention and exogenous volatility shocks, and quantify their effects in Section 6.

5 Dynamic Model

The economy consists of the same agents and final good producer as in Section 2. Time lasts from $t = 0, \cdots, \infty$. Each period splits into three stages in which agents pay attention,

[8]For the probability-range data, the SPF asks for fixed-event forecasts of year-over-year GDP growth. Appendix E describes how the series is adjusted to be comparable to the other two measures.

[9]If TFP is adjusted for utilization, the correlation with output is $-.10$, and the standard deviation relative to the long-run average is .08.

supply labor, and consume. Aggregate productivity is persistent and features countercyclical volatility. Information is dispersed among agents. Besides observing idiosyncratic signals, agents observe the prices of their own products subject to unobservable idiosyncratic demand shocks. These shocks make it difficult for agents to perfectly infer the aggregate state, thereby perpetuating the dispersion of information over time.

Preference and Technology

Agents pay attention, supply labor, and consume with preference given by:

$$\mathbb{E}_{i,0} \sum_{t=0}^{\infty} \beta^t U(c_{i,t}, n_{i,t}, z_{i,t}).$$

Agents produce intermediate goods using labor in each period with linear technology $q_{i,t} = n_{i,t}$, and face period-by-period budget constraints $c_{i,t} \leq p_{i,t} q_{i,t}$. I abstract away from capital accumulation and saving to focus on the dynamics generated by agents' attention choice over the business cycle.

The final good producer maximizes profit $Y_t - \int p_{i,t} y_{i,t}$ by combining intermediate goods to produce the final good with technology:

$$Y_t = e^{\theta_t + \vartheta_t} \left(\int \left(\, e^{\omega_{i,t}} y_{i,t} \, \right)^{1-\eta} di \right)^{\frac{1}{1-\eta}},$$

where

$$\theta_t = \rho \, \theta_{t-1} + \omega_t, \quad \vartheta_t = \rho \, \vartheta_{t-1} + \Sigma(\theta_{t-1})\omega_t, \quad \omega_t \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\omega^2),$$

and $\{\omega_{i,t}\}$ are idiosyncratic demand shocks, i.i.d. normal over time and goods with a common variance $\sigma_{\omega_i}^2$. Aggregate state θ_t affects the level of productivity, $\theta_t + \vartheta_t$, and introduces variation in volatility through its effect on ϑ_t. Together, productivity $\theta_t + \vartheta_t$ features persistence ρ and stochastic volatility $1 + \Sigma(\theta_{t-1})$. If $\Sigma \equiv 0$, aggregate productivity is simply given by θ_t; if $\Sigma' < 0$, its volatility decreases with the past aggregate state θ_{t-1}.

Timing and Information

Each period consists of three stages. In stage 1, each agent chooses attention $z_{i,t}$ given information set $\mathcal{F}_{i,t-1}$. In stage 2, according to their attention choice, agents receive idiosyncratic signals about the aggregate state θ_t:

$$x_{i,t} = \theta_t + \frac{\epsilon_{i,t}}{\sqrt{z_{i,t}}},, \text{ where } \epsilon_i \overset{iid}{\sim} \mathcal{N}(0,1).$$

Agents decide labor input $n_{i,t}$ after observing $x_{i,t}$, and the economy proceeds to stage 3. In stage 3, equilibrium prices $\{p_{i,t}\}$ realize. Each agent observes the price of their product $p_{i,t}$ and consumes $c_{i,t}$ subject to budget constraints.[10] The final good producer observes $\{\theta_t + \vartheta_t, \omega_{i,t}, p_{i,t}\}$ and chooses intermediate input $\{y_{i,t}\}$ to maximize profit.

To summarize, each agent i chooses stochastic processes $z_{i,t}, n_{i,t}, c_{i,t}$ under information constraints:

$$z_{i,t} \in \mathcal{F}'_{i,t-1} \coloneqq \boldsymbol{\sigma}(x_i^{t-1},\ p_i^{t-1}), \quad n_{i,t} \in \mathcal{F}_{i,t} \coloneqq \boldsymbol{\sigma}(x_i^t,\ p_i^{t-1}),$$

and $c_{i,t} \in \mathcal{F}'_{i,t}$, where x_i^t, p_i^t denote the histories up to time t. Agents have a common prior $\theta_0 \sim \mathcal{N}(0, \sigma_0^2)$, which is inconsequential for the stationary properties of the economy. The assumption that agents observe signals about the aggregate state θ_t instead of productivity $\theta_t + \vartheta_t$ implies that, given the same attention choices, volatility shocks will not induce exogenous variation in the informativeness of signals.

Definition of Equilibrium

An equilibrium consists of stochastic processes $\{z_{i,t}, n_{i,t}, c_{i,t}, q_{i,t}, y_{i,t}, Y_t, p_{i,t}\}$ such that (i) $z_{i,t}, n_{i,t}$, and $c_{i,t}$ optimize the expected utility for each agent, subject to budget constraints and information constraints; (ii) given prices $\{p_{i,t}\}$, the final good producer chooses $\{y_{i,t}\}$ to optimize profit; (iii) productions of $q_{i,t}, Y_t$ are determined by the respective technologies; and (iv) markets clear for all goods $q_{i,t} = y_{i,t}, \forall i, t$ and $Y_t = \int c_{i,t}\ di, \forall t$.

5.1 Equilibrium Characterization

The prices of intermediate goods $\{p_{i,t}\}$ can be solved from the final good producer's maximization problem and market clearing, given the distribution of labor input:

$$p_{i,t} = e^{\theta_t + \vartheta_t + \omega_{i,t}} N_t^{\eta} n_{i,t}^{-\eta}, \quad N_t = \left(\int \left(e^{\omega_{i,t}} n_{i,t} \right)^{1-\eta} di \right)^{\frac{1}{1-\eta}}.$$

Prices of intermediate goods depend not only on the aggregate variables but also on the idiosyncratic demand shocks $\omega_{i,t}$. Shock $\omega_{i,t}$ shifts the final good producer's demand for good i, but it is not observed by the agent. As a result, the model features persistent dispersed information because agents cannot make perfect inferences about the past aggregate state based on observations of prices. Maintaining dispersed information over time is crucial for the model to generate empirically plausible forecast patterns.

[10] For agents' consumption choices to be internally consistent, they must observe $p_{i,t}$ because $c_{i,t} \le p_{i,t} n_{i,t}$. However, as discussed in Section 6.1, $\{p_{i,t}\}$ plays a minor role in agents' information sets in the calibrated model.

The equilibrium can be characterized by a system of equations involving attention $z_{i,t}$, labor input $n_{i,t}$, and aggregate labor N_t:

Lemma 6 *An equilibrium solves the following system:*

$$\sum_{\tau=t}^{\infty} \beta^{\tau-t}\mathbb{E}\Big[\; U_{i,\tau} \times \frac{1-\epsilon_{i,t}^2}{2z_{i,t}} + \frac{\partial}{\partial z_{i,t}} U_{i,t} \;\Big|\mathcal{F}'_{i,t-1}\Big] = 0,\ \forall i, t,$$

$$\mathbb{E}\Big[\; \frac{\partial}{\partial n_{i,t}} U_{i,t} \;\Big|\mathcal{F}_{i,t}\Big] = 0,\ \forall i, t,$$

$$N_t = \Big(\int \Big(e^{\omega_{i,t}} n_{i,t}\Big)^{1-\eta} di\Big)^{\frac{1}{1-\eta}},\ \forall t,$$

where $U_{i,t} \coloneqq U\big(e^{\theta_t+\vartheta_t+\omega_{i,t}} N_t^{\eta} n_{i,t}^{1-\eta}, n_{i,t}, z_{i,t}\big)$ *denotes agent i's period utility.*

Proof. See Appendix C.1. □

The first two conditions represent the optimality of attention choice and labor input, and the third condition describes how aggregate input depends on individual input and idiosyncratic demand shocks $\omega_{i,t}$. These conditions are natural generalizations of the equilibrium conditions in Section 2. Unlike the static model, however, agents enter each period with different information sets $\mathcal{F}_{i,t-1}$ that depend on their past signal realizations and attention choices.

Infinite Regresss Problem

The dynamic structure of agents' information sets is crucial for generating the empirical pattern observed in forecast surveys.[11] However, an advancement in the existing solution method is necessary to study how attention choice varies with aggregate conditions in dispersed information models.

The persistence of dispersed information poses a challenge because each agent's equilibrium strategies are generally nonlinear functions of an infinite-dimension history of signals and prices. For example, using the approximation described in Section 2, the first- and second-order expansions of input $\hat{n}_{i,t}, \hat{\hat{n}}_{i,t}$ are multilinear functions of the signal history, $s_i^t \coloneqq (x_i^t,\ p_i^{t-1})$:

$$\hat{n}_{i,t} = \boldsymbol{n}_s \hat{s}_i^t + \boldsymbol{n}_\delta, \quad \hat{\hat{n}}_{i,t} = \hat{s}_i^{t\intercal} \boldsymbol{n}_{ss} \hat{s}_i^t + \boldsymbol{n}_{\delta\delta} + \cdots .$$

This challenge in dispersed information models is often referred to as the *infinite regress problem*, under which signal histories cannot be succinctly summarized by a finite-dimensional

[11] For example, the measure of information rigidity presented in Table 7 (Goldstein, 2023) builds on a persistent deviation of individual forecasts from average forecasts.

state. Methods have been proposed to address the problem within linear rational expectation models.[12] However, existing methods are constrained to first-order approximations. As a result, they cannot capture the dynamics of attention and uncertainty, as these dynamics are higher-order features of the model. Below, I introduce a computational procedure that extends the perturbation method developed in Section 2 to compute higher-order dynamics of dispersed information models.

5.2 Higher-Order Dynamics with Infinite Regress

To compute the expansions numerically, I look for a finite-dimension approximation for the signal history. An analogy of the spectrum theorem motivates the search for factors $\{f_{i,t}^{(1)}, f_{i,t}^{(2)}, \ldots\}$ and multilinear functions $\{\Phi_f^n, \Phi_{ff}^n, \ldots\}$ such that:

$$\boldsymbol{n}_s \hat{s}_i^t \approx \Phi_f^n f_{i,t}^{(1)}, \quad \hat{s}_i^{t\intercal} \boldsymbol{n}_{ss} \hat{s}_i^t \approx f_{i,t}^{(2)\intercal} \Phi_{ff}^n f_{i,t}^{(2)}, \ \ldots.$$

I consider factor structures of the following form:

$$f_{i,t+1}^{(m)} = A^{(m)} \ f_{i,t}^{(m)} + C^{(m)} \hat{s}_{i,t}, \ \forall m = 1, 2, \ldots,$$

and I solve for the optimal factor structure for each order of approximation with the corresponding multilinear functions. To solve for the m^{th} order approximation given the first $(m-1)^{th}$ orders, I consider the following two-step procedure: (i) Given $A^{(m)}, \ C^{(m)}$, stimulate the economy and solve the corresponding multilinear functions Φ's that minimize the sum of squared residuals of the expanded equilibrium conditions; (ii) optimize over $A^{(m)}, \ C^{(m)}$ to look for the optimal factor structure. Computationally, the first step generally involves only a linear-quadratic problem, which can be solved efficiently; the second step is a non-linear problem but can be easily parallelized.

The procedure is reminiscent of the methods for solving heterogeneous-agent models, e.g., Krusell and Smith (1998), but differs in subtle ways. The infinite-dimensional-state problem here originates from both the time and cross-sectional dimensions. For the time dimension, agents need to tack the infinite history of signals due to infinite regress; for the cross-section, heterogeneity arises due to agents having dispersed signals and beliefs. The procedure described above searches for finite-dimension factors that summarize signal histories along the time dimension, contrasting methods for solving models where summarizing cross-sectional

[12] Woodford (2001) shows how to represent such equilibrium as a finite system of difference equations despite expectations of arbitrarily high-order matter. Lorenzoni (2009) truncates the history of signals with a fixed time window. Nimark (2017) truncates belief hierarchy above a certain order. Huo and Pedroni (2020) and Huo and Takayama (2023) show a finite-state representation exists for certain cases, and they provide numerical methods for cases where infinite regress is unavoidable.

heterogeneity is key. The complication of cross-sectional heterogeneity in this paper is alleviated by the perturbation approach. Due to the multilinear structure of the Taylor expansions and Gaussian shocks, expansions of aggregate variables can be easily linked to individual variables, given $A^{(m)}$, $C^{(m)}$ and the multilinear functions Φ's. This aggregation result is common in linear dispersed information models, which extends naturally to higher-order approximations.

Appendix D provides a detailed discussion of the procedure. Besides the computational procedure, I provide analytic results necessary for the computation, including the first- and second-order expansion of the equilibrium conditions in Lemma 6, as well as an expansion of the expectation operator to address the non-linear filtering problem resulting from higher-order dynamics.

6 Quantitative Implications

I now quantify the importance of endogenous attention in generating macroeconomic uncertainty over the business cycle and contrast it with the effect of exogenous volatility shocks. I calibrate the model such that (i) agents' expectation updates generate variation in information rigidity as in forecast surveys and (ii) the size of volatility shocks corresponds to its empirical counterpart. Both empirical features are described in Section 4. Using the calibrated model, I assess how much variation in aggregate volatility, forecast dispersion, and subjective uncertainty can be attributed to the two mechanisms.

6.1 Calibration

Each period in the model corresponds to a quarter. Stages 1 and 2 of each period occur at the beginning of a quarter, at which point agents pay attention, receive information, make forecasts, and make input decisions. Stage 3 occurs at the end of the quarter when production takes place, and output is recorded. I assume that agents' forecasts are represented by forecasts in the SPF. The equilibrium is approximated to the second order.

Discount rate β is set at .995, corresponding to the quarterly frequency. The elasticity of substitution between intermediate goods $1/\eta$ is set at 4 so that the average markup over the marginal cost of labor is 33% in the steady state. The standard deviation of idiosyncratic demand shocks σ_{ω_i} is set at 2.5%, generating a dispersion of quarterly price change around 4% for the intermediate goods.

The utility from consumption, labor, and attention takes the following form:

$$U(c_{i,t}, n_{i,t}, z_{i,t}) = \frac{1}{1-\gamma}\left\{ \max\left\{ c_i - \frac{n_i^{1+\nu}}{1+\nu},\ \underline{u} \right\}^{1-\tilde{\gamma}} - (1-\tilde{\gamma})\kappa z_{i,t} \right\}^{\frac{1-\gamma}{1-\tilde{\gamma}}}. \tag{7}$$

Absent attention cost, the flow utility reduces to the standard GHH preference.[13] Parameter κ governs the average level of attention. Parameter $\tilde{\gamma}$ determines the strength of the income effect on attention and effectively controls how much attention varies over the business cycle. I calibrate κ and $\tilde{\gamma}$ jointly with the rest of the parameters to generate expectation updates consistent with the forecast survey.

Quantifying Attention Variation and Volatility Shocks

I calibrate three sets of parameters internally. First, ρ, σ_ω and ν govern the persistence of the aggregate productivity process, its average volatility, and the convexity of the labor cost function. As these parameters are directly linked to the input and output of production, I target (i) the persistence of aggregate output, (ii) the (unconditional) volatility of output, and (iii) the relative volatility of hours to output.

Second, κ and $\tilde{\gamma}$ control the average level of attention $\bar{z}$ and the extent to which attention response $\hat{z}_t$ varies over the business cycle. I calibrate them to match measures of information rigidity, β_{CG} and $\Delta\beta_{CG}$, from the forecast regression in Table 1, where $\Delta\beta_{CG}$ corresponds to the interaction term from the indicator of below-trend output.

Finally, the magnitude of exogenous fluctuations in volatility is governed by the slope of function Σ at the steady state. I denote the slope by $\bar{\alpha} := \bar{\Sigma}'(0)$, corresponding to parameter α in Section 3 that introduces countercyclical volatility. I calibrate $\bar{\alpha}$ to match the standard deviation of σ_t^{TFP} relative to its long-run average, as presented in Table 2.

Table 3 shows the calibrated parameters with targeted moments from the data and the model. By matching measures of information rigidity β_{CG} and $\Delta\beta_{CG}$, attention choice in the model implies that the average size of noise $1/\sqrt{\bar{z}}$ is such that agents update their expectations about the aggregate state θ_t with an average Kalman gain of .54 from signals $x_{i,t}$.[14] For an average agent, attention varies over time such that the size of noise $1/\sqrt{z_{i,t}}$ at its 20 percentile is 33% of that at its 80 percentile. Table 7 in Appendix E shows that the model also generates patterns of expectation updates similar to those from the individual-

[13] The value of γ is inconsequential for the fluctuations of aggregate variables because risk aversion affects only their levels up to second-order approximation.

[14] By contrast, the Kalman gain from price $p_{i,t-1}$ is only .001, indicating that most learning about the aggregate condition comes from active attention choice instead of passive observation of prices.

level regression (Goldstein, 2023) discussed in Section 4, measuring how information rigidity varies with individual forecasters' expectations about aggregate conditions.[15]

Finally, the level of strategic complementarity, $s = \frac{\eta}{\eta+\nu}$, is pinned down by the convexity of labor cost ν, given the elasticity of substitution η. The calibrated model features a high level of strategic complementarity, $s = .78$. This results from a large elasticity of labor input $1/\nu$ due to hours being almost as volatile as output in the data.

Table 3: Calibration

Parameter	ρ	σ_ω	ν	$1/\sqrt{\bar{z}(\kappa)}$	$\tilde{\gamma}$	$\bar{\alpha}$
	.83	.0011	.07	.0012	63.3	$-.116$
Moment	$\rho_1\ Y_t$	$sd\ Y_t$	$\frac{sd\ N_t}{sd\ Y_t}$	β_{CG}	$\Delta\beta_{CG}$	$\frac{sd}{avg}\ \sigma_t^{TFP}$
Data	.93	1.94	.91	.73	$-.24$	.18
Model	.93	1.96	.91	.74	$-.24$	.18

Sample: 1968Q3 to 2019Q4, detrended with a band-pass filter at 6-32 quarters frequency. Model moments are averages of 1000 simulations of 200 quarters with 50 forecasters.

6.2 Decomposition: Attention vs. Volatility Shocks

To isolate the effects of attention choice on uncertainty, volatility, and forecast dispersion, I consider three alternative specifications of the model: (i) the *full model* in which attention endogenously responds to aggregate conditions and productivity features countercyclical volatility, (ii) a model with only *volatility shocks* in which attention is fixed exogenously at its average level, $z_{i,t} = \bar{z}$, and (iii) a model in which *attention* responds endogenously without volatility shocks, $\bar{\alpha} = 0$.

Cyclicality of Uncertainty and Expectation Updates

Table 4 compares the cyclicality of uncertainty, volatility, and forecast dispersion generated by the three specifications to their empirical counterparts, along with implications on information rigidity.

The first row shows the same empirical moments from Section 4: all three measures correlate negatively with output, and $\Delta\beta_{CG}$ is negative, indicating a decrease in information rigidity during low-output periods. The next three rows represent moments generated by the three alternative specifications. Uncertainty, volatility, and forecast dispersion are countercyclical

[15] Alternatively, one can calibrate the model to match these moments instead.

in all three specifications. Yet, only the two specifications with endogenous attention features reduced information rigidity during low-output periods. In fact, the measure of information rigidity increases slightly if volatility shocks are the only driving force, consistent with the analytical result in Corollary 1.

Table 4: Cyclicality of Uncertainty and Expectation Updates

$cor(\cdot, \tilde{Y}_t)$	σ_t^Y	d_t^Y	υ_t^Y	$\Delta\beta_{CG}$
data	$-.39$	$-.40$	$-.33$	$-.24$
full model	$-.90$	$-.88$	$-.87$	$-.24$
vol. shocks	$-.76$	$-.78$	$-.95$	$.01$
attention	$-.93$	$-.87$	$-.93$	$-.20$

Model moments are averages of 1000 simulations of 200 quarters with 50 forecasters.

With only an increase in the volatility of aggregate productivity, agents' expectations updates do not fully account for the size of movements in aggregate output, as they do not have perfect information about how much the volatility $\Sigma(\theta_{t-1})$ has increased. As a result, information rigidity increases, as if agents' expectation updates feature further underreaction. By contrast, with endogenous attention absent any volatility shocks, aggregate volatility, forecast dispersion, and subjective uncertainty are countercyclical due to agents paying attention and responding to aggregate conditions under dispersed information. As agents pay attention, the measure of information rigidity decreases.

The fact that macroeconomic uncertainty arises endogenously with attention is not an a priori assumption of the model. Similar to Theorem 1, endogenous attention leads to countercyclical uncertainty in the dynamic model only because the average level of attention is below a certainty threshold, given the level of strategic complementarity in the economy. Therefore, the fact that uncertainty about output and forecast dispersion increases with attention is a quantitative result implied by the average level of information rigidity in the data.

Magnitude of Fluctuations in Uncertainty

Table 5 quantifies how much variations in uncertainty, volatility, and forecast dispersion can be explained by countercyclical attention, exogenous volatility shocks, and the interaction between the two mechanisms.

Table 5: Magnitude of Fluctuations in Uncertainty

sd/avg	σ_t^Y	d_t^Y	υ_t^Y	σ_t^{TFP}
data	.47	.37	.36	.18
full model	.57	.52	.39	.18
vol. shocks	.15	.18	.15	.18
attention	.37	.32	.14	–

Model moments are averages of 1000 simulations of 200 quarters with 50 forecasters.

In the data, the three measures fluctuate with standard deviations ranging from 40% to 50% relative to their long-run averages. In the full model, the two mechanisms fully account for the fluctuations: Aggregate volatility and forecast dispersion fluctuate with standard deviations that are around 50% relative to their long-run averages, and subjective uncertainty varies by around 40%.

The two alternative specifications isolate the effect of each mechanism. With only volatility shocks, the three measures fluctuate with standard deviations around 20% relative to the long-run averages, which are similar to how much the volatility of aggregate productivity fluctuates. This reflects a lack of internal mechanisms to amplify the exogenous volatility shocks in a fixed-attention economy.

By contrast, endogenous attention generates significant fluctuations in uncertainty, volatility, and forecast dispersion without exogenous volatility shocks, accounting for 40% to 80% of the observed variation. The size of fluctuations in these measures depends on how much attention varies over the business cycle. This feature is quantified by how much information rigidity varies over the business cycle, $\Delta\beta_{CG}$, as targeted by the strength of income effect on attention $\tilde{\gamma}$ in the calibration model.

The difference between the full model and the two mechanisms shows that endogenous attention amplifies exogenous volatility shocks: agents pay attention not only due to the income effect on attention but also as a response to the exogenous increase in volatility. Table 5 shows that the magnitudes of fluctuations in the three measures are larger in the full model than the sum of the two isolated mechanisms. The interaction between the two mechanisms accounts for between 4% to 25% of the fluctuations in these measures.

Finally, the fact that countercyclical attention can generate large fluctuations in uncertainty, volatility, and forecast dispersion does not necessarily depend on the exact mechanism driving attention choices. It may be that countercyclical attention is driven by an alternative mechanism other than the income effect. For example, agents can pay more attention during a recession because attention cost κ decreases due to a lower opportunity cost of their time. However, as long as this alternative mechanism generates a countercyclical attention response consistent with cyclical information rigidity in the data, this response will likely generate the same countercyclical fluctuations in uncertainty, volatility, and forecast dispersion. In this sense, the result in Table 5 represents the extent to which macroeconomic uncertainty can be accounted for by countercyclical attention, independent of the exact mechanism driving agents' attention choice.

7 Conclusion

I show that increases in macroeconomic uncertainty during recessions can arise from economic agents' attempts to acquire more information. This mechanism results from agents interacting under dispersed information, and it is consistent with evidence of reduced information rigidity in those periods, in contrast to the prediction of exogenous volatility shocks. When I quantify the variation in attention with measures of information rigidity, countercyclical attention accounts for half of the observed fluctuations in aggregate volatility, forecast dispersion, and uncertainty about aggregate output.

Exploring the normative implications of the mechanism will be valuable for understanding how macroeconomic policies can mitigate uncertainty during economic crises. Many macroeconomic policies work through their effects on people's expectations about aggregate outcomes. Therefore, understanding how people process information in response to aggregate events is crucial for designing policies that can effectively coordinate and anchor beliefs. Because the responses of information choices to aggregate events are higher-order features in models with information frictions, the perturbation technique developed in this paper is particularly suitable for answering these questions. I leave these topics for future research.

References

[1] Tobias Adrian, Nina Boyarchenko, and Domenico Giannone. "Vulnerable Growth". In: *American Economic Review* 109.4 (2019), pp. 1263–89.

[2] George-Marios Angeletos and Zhen Huo. “Myopia and Anchoring”. In: *American Economic Review* 111.4 (2021), pp. 1166–1200.

[3] George-Marios Angeletos and Jennifer La’O. “Noisy Business Cycles”. In: *NBER Macroeconomics Annual* 24.1 (2010), pp. 319–378.

[4] George-Marios Angeletos and Jennifer La’O. “Sentiments”. In: *Econometrica* 81.2 (2013), pp. 739–779.

[5] George-Marios Angeletos and Chen Lian. “Forward Guidance without Common Knowledge”. In: *American Economic Review* 108.9 (2018), pp. 2477–2512.

[6] George-Marios Angeletos and Chen Lian. “Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination”. In: *Handbook of macroeconomics*. Vol. 2. Elsevier, 2016, pp. 1065–1240.

[7] Rüdiger Bachmann, Steffen Elstner, and Eric R Sims. “Uncertainty and Economic Activity: Evidence from Business Survey Data”. In: *American Economic Journal: Macroeconomics* 5.2 (2013), pp. 217–49.

[8] Jess Benhabib, Xuewen Liu, and Pengfei Wang. “Endogenous Information Acquisition and Countercyclical Uncertainty”. In: *Journal of Economic Theory* 165 (2016), pp. 601–642.

[9] Anmol Bhandari et al. “Inequality, Business Cycles, and Monetary-Fiscal Policy”. In: *Econometrica* 89.6 (2021), pp. 2559–2599.

[10] Nicholas Bloom. “Fluctuations in Uncertainty”. In: *Journal of Economic Perspectives* 28.2 (2014), pp. 153–76.

[11] Nicholas Bloom. “The Impact of Uncertainty Shocks”. In: *Econometrica* 77.3 (2009), pp. 623–685.

[12] Nicholas Bloom et al. “Really Uncertain Business Cycles”. In: *Econometrica* 86.3 (2018), pp. 1031–1065.

[13] Jaroslav Borovicka and Lars Peter Hansen. “Robust Preference Expansions”. In: (2013).

[14] Ryan Chahrour and Kyle Jurado. “Revisiting the Forecasts of Others”. In: (2023).

[15] Olivier Coibion and Yuriy Gorodnichenko. “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts”. In: *American Economic Review* 105.8 (2015), pp. 2644–78.

[16] Joseph Engelberg, Charles F Manski, and Jared Williams. “Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters”. In: *Journal of Business & Economic Statistics* 27.1 (2009), pp. 30–41.

[17] Pablo D Fajgelbaum, Edouard Schaal, and Mathieu Taschereau-Dumouchel. “Uncertainty Traps”. In: *The Quarterly Journal of Economics* 132.4 (2017), pp. 1641–1692.

[18] Joel P Flynn and Karthik Sastry. "Attention Cycles". In: *Available at SSRN 3592107* (2023).

[19] Nathan Goldstein. "Tracking inattention". In: *Journal of the European Economic Association* 21.6 (2023), pp. 2682–2725.

[20] Zhen Huo and Marcelo Pedroni. "A Single-Judge Solution to Beauty Contests". In: *American Economic Review* 110.2 (2020), pp. 526–68.

[21] Zhen Huo and Naoki Takayama. "Rational Expectations Models with Higher-Order Beliefs". In: *Available at SSRN 3873663* (2023).

[22] Cosmin Ilut, Matthias Kehrig, and Martin Schneider. "Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News". In: *Journal of Political Economy* 126.5 (2018), pp. 2011–2071.

[23] Kenneth L Judd. *Numerical Methods in Economics*. MIT press, 1998.

[24] Kyle Jurado, Sydney C Ludvigson, and Serena Ng. "Measuring Uncertainty". In: *American Economic Review* 105.3 (2015), pp. 1177–1216.

[25] Kenneth Kasa. "Forecasting the Forecasts of Others in the Frequency Domain". In: *Review of Economic Dynamics* 3.4 (2000), pp. 726–756.

[26] Nicholas Kozeniauskas, Anna Orlik, and Laura Veldkamp. "What are Uncertainty Shocks?" In: *Journal of Monetary Economics* 100 (2018), pp. 1–15.

[27] Per Krusell and Anthony A Smith Jr. "Income and Wealth Heterogeneity in the Macroeconomy". In: *Journal of Political Economy* 106.5 (1998), pp. 867–896.

[28] Giovanni Lombardo and Harald Uhlig. "A Theory of Pruning". In: *International Economic Review* 59.4 (2018), pp. 1825–1836.

[29] Guido Lorenzoni. "A Theory of Demand Shocks". In: *American Economic Review* 99.5 (2009), pp. 2050–84.

[30] Guido Lorenzoni. "Optimal Monetary Policy with Uncertain Fundamentals and Dispersed Information". In: *The Review of Economic Studies* 77.1 (2010), pp. 305–338.

[31] Robert E Lucas. "Expectations and the Neutrality of Money". In: *Journal of Economic Theory* 4.2 (1972), pp. 103–124.

[32] Bartosz Maćkowiak and Mirko Wiederholt. "Business Cycle Dynamics under Rational Inattention". In: *The Review of Economic Studies* 82.4 (2015), pp. 1502–1532.

[33] Bartosz Maćkowiak and Mirko Wiederholt. "Optimal Sticky Prices under Rational Inattention". In: *American Economic Review* 99.3 (2009), pp. 769–803.

[34] Kristoffer Nimark. "Dynamic Higher Order Expectations". In: (2017).

[35] Kristoffer P Nimark. "Man-Bites-Dog Business Cycles". In: *American Economic Review* 104.8 (2014), pp. 2320–67.

[36] Phelps. *Microeconomic Foundations of Employment and Inflation Theory.* WW Norton New York, 1970.
[37] Christopher A Sims. "Implications of Rational Inattention". In: *Journal of Monetary Economics* 50.3 (2003), pp. 665–690.
[38] Wenting Song and Samuel Stern. "Firm Inattention and the Efficacy of Monetary Policy: A Text-Based Approach". In: *Available at SSRN 3900641* (2020).
[39] Ludwig Straub and Robert Ulbricht. "Endogenous Uncertainty and Credit Crunches". In: *Available at SSRN 2668078* (2023).
[40] Robert M Townsend. "Forecasting the Forecasts of Others". In: *Journal of Political Economy* 91.4 (1983), pp. 546–588.
[41] Stijn Van Nieuwerburgh and Laura Veldkamp. "Learning Asymmetries in Real Business Cycles". In: *Journal of Monetary Economics* 53.4 (2006), pp. 753–772.
[42] Michael Woodford. *Imperfect Common Knowledge and the Effects of Monetary Policy.* Tech. rep. National Bureau of Economic Research, 2001.

A Derivations and Proofs

I derive all results in Lemma 1 to 5 and Theorem 1 with aggregate productivity of the form $\bar{\theta} + (1 + \alpha\bar{\theta})\theta$, allowing for the possibility of exogenous volatility as in Section 3.4. Furthermore, I adopt the following notations:

$$v_i = v(c_i, n_i) := \max\{c_i - \frac{n_i^{1+\nu}}{1+\nu}, \underline{u}\}, \quad \kappa(z_i) := \kappa z_i,$$

so that agents' preferences can be written as $\frac{1}{1-\tilde{\gamma}} v(c_i, n_i)_i^{1-\tilde{\gamma}} - \kappa(z_i)$.

A.1 Proof of Lemma 1

The aggregation condition in Equation 1 results from standard derivation.

To derive the optimality condition in equation 2, substitute $c_i = e^{\bar{\theta}+(1+\alpha\bar{\theta})\theta} N^\eta n_i^{1-\eta}$ and take the first-order condition with respect to n_i. This gives

$$\mathbb{E}_i \left[\left((1-\eta) e^{\bar{\theta}+(1+\alpha\bar{\theta})\theta} N^\eta n_i^{-\eta} - n_i^\nu \right) v_i^{-\tilde{\gamma}} \mathbf{1}_{\{v_i > \underline{u}\}} \right] = 0.$$

Because $u(c_i, n_i) = \frac{1}{1-\tilde{\gamma}} v(c_i, n_i)_i^{1-\tilde{\gamma}}$, we have

$$u_c(c_i, n_i) = v_i^{-\tilde{\gamma}} \mathbf{1}_{\{v_i > \underline{u}\}}.$$

Dividing both sides by $\mathbb{E}_i[u_c(c_i, n_i)]$. And since $n_i \in \mathcal{F}_i$, we can move n_i^ν to the right-hand side and we have:

$$\mathbb{E}_i\Big[\Big((1-\eta)e^{\bar{\theta}+(1+\alpha\bar{\theta})\theta}N^\eta n_i^{-\eta}\Big)\frac{u_c(c_i, n_i)}{\mathbb{E}_i[u_c(c_i, n_i)]}\Big] = n_i^\nu,\ \forall x_i.$$

For the optimality condition of attention in Equation 3, let $V(z_i)$ denote agent i's value function given attention z_i:

$$V(z_i) := \max_{n(\cdot)} \mathbb{E}\Big[\frac{1}{1-\tilde{\gamma}}\Big(\max\{e^{\bar{\theta}+(1+\alpha\bar{\theta})\theta}N(\theta,x)^\eta n(x_i,x)^{1-\eta} - \frac{n(x_i,x)^{1+\nu}}{1+\nu}, \underline{u}\}\Big)^{1-\tilde{\gamma}}\Big| x, z_i\Big] - \kappa(z_i),$$

where aggregate input $N(\theta, x)$ is taken as given.

Attention optimality requires $V'(z_i) = 0$. The envelope theorem implies

$$\int \frac{1}{1-\tilde{\gamma}}\Big(\max\{e^{\bar{\theta}+(1+\alpha\bar{\theta})\theta}N(\theta,x)^\eta n(x_i,x)^{1-\eta} - \frac{n(x_i,x)^{1+\nu}}{1+\nu}, \underline{u}\}\Big)^{1-\tilde{\gamma}}\frac{\partial}{\partial z_i}\varphi(\theta, x_i|x, z_i)dx_i d\theta = \kappa'(z_i).$$

A.2 Detailed Derivations

I provide details of the perturbation method, and I solve the zeroth-, first-, and second-order approximation of the equilibrium. Lemma 2, 3, and 4 follow immediately from the derivation.

Formulation of the Perturbation

Consider a sequence of economies parameterized by δ such that

$$\bar{\theta}(\delta) = \bar{\theta}\delta, \quad \theta(\delta) = \theta\delta, \quad \epsilon(\delta) = \epsilon\delta, \quad \epsilon_i(\delta) = \epsilon_i\delta, \quad \kappa(z,\delta) = \delta^2\kappa(z).$$

For the economy indexed by δ, the equilibrium is described by Lemma 1:

$$\mathbb{E}\Big[\frac{1}{1-\tilde{\gamma}}v_i(\delta)^{1-\tilde{\gamma}} \times \frac{1-\epsilon_i^2}{2z(\delta)}\big|x(\delta), z(\delta)\Big] - \delta^2\kappa'(z(\delta)) = 0$$

$$\mathbb{E}\Big[v_i(\delta)^{-\tilde{\gamma}}\Big((1-\eta)e^{\bar{\theta}(\delta)+(1+\alpha\bar{\theta}(\delta))\theta(\delta)}N^\eta n_i^{-\eta} - n_i(\delta)^\nu\Big)\mathbf{1}_{\{v_i(\delta)>\underline{u}\}}\big|x(\delta), x_i(\delta), z(\delta)\Big] = 0,$$

$$\log N(\delta) = \frac{1}{1-\eta}\log\Big(\int \exp((1-\eta)\log n_i(\delta))\Big),$$

where $v_i(\delta) = \max\{c_i(\delta) - \frac{n_i(\delta)^{1+\nu}}{1+\nu}, \underline{u}\}$ and the term $\frac{1-\epsilon_i^2}{2z(\delta)}$ comes from:

$$\frac{\partial}{\partial z_i}\varphi(\bar{\theta}, \theta, x_i|x, z_i) = \frac{1-\epsilon_i^2}{2z_i}\varphi(\bar{\theta}, \theta, x_i|x, z_i).$$

Assume that the equilibrium can be approximated by Taylor expansions:

$$\log n_i(\delta) \approx \bar{n} + \hat{n}_i\delta + \frac{1}{2}\hat{\hat{n}}_i\delta^2, \quad \log N(\delta) \approx \bar{N} + \hat{N}\delta + \frac{1}{2}\hat{\hat{N}}\delta^2,$$
$$\log z(\delta) \approx \log \bar{z} + \hat{z}\delta, \quad \log v_i(\delta) \approx \bar{v} + \hat{v}_i\delta + \frac{1}{2}\hat{\hat{v}}_i\delta^2.$$

Zeroth-Order Expansion

Evaluating the equilibrium conditions at $\delta \to 0$, we have

$$\mathbb{E}_0\Big[\frac{e^{(1-\tilde{\gamma})\bar{v}}}{1-\tilde{\gamma}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big] = 0, \quad e^{\eta\bar{N}-\eta\bar{n}}(1-\eta) = e^{\nu\bar{n}}, \quad \bar{N} = \bar{n},$$

where I adopt the following shorthand for the expectation operator:

$$\mathbb{E}_0[\cdot] \coloneqq \mathbb{E}[\cdot|\hat{x}, \bar{z}],$$

and $\hat{x} = \bar{\theta} + \epsilon$ and $\bar{v} = \log(e^{\eta\bar{N}+(1-\eta)\bar{n}} - \frac{1}{1+\nu}e^{(1+\nu)\bar{n}})$.

Conditions above determine $\bar{n}, \bar{N}, \bar{v}$, but not $\bar{z}$. This is because both the marginal benefit of attention and the marginal cost $\kappa'(z,\delta)$ are zero at $\delta \to 0$:

$$\frac{e^{(1-\tilde{\gamma})\bar{v}}}{1-\tilde{\gamma}}\mathbb{E}_0\Big[\frac{1-\epsilon_i^2}{2\bar{z}}\Big] \equiv 0, \quad \forall \bar{z}.$$

As I show below, the m^{th}-order expansion of attention $z(\delta)$ will be determined jointly with the $(m+1)^{th}$-order expansion of $n_i(\delta)$ and $N(\delta)$ from the attention optimality condition expanded to the $(m+2)^{th}$ order. This property is similar to a perturbation for portfolio choice problems as discussed in Judd (1998).

First-Order Expansion

Differentiating the equilibrium conditions with respect to δ at $\delta \to 0$,

$$\mathbb{E}_0\Big[e^{(1-\tilde{\gamma})\bar{v}}\hat{v}_i\frac{1-\epsilon_i^2}{2\bar{z}} - \frac{e^{(1-\tilde{\gamma})\bar{v}}}{1-\tilde{\gamma}}\frac{1-\epsilon_i^2}{2\bar{z}}\hat{z}\Big] = 0, \tag{8}$$

$$\hat{n}_i = \mathbb{E}[\ r(\bar{\theta}+\theta) + s\hat{N}\ |\bar{\mathcal{F}}_i], \tag{9}$$

$$\hat{N} = \int \hat{n}_i, \tag{10}$$

where $r \coloneqq \frac{1}{\nu+\eta}$, $s \coloneqq \frac{\eta}{\nu+\eta}$ are as defined in Section 2.

Expand $\log n_i(\delta) = \boldsymbol{n}(x_i(\delta), x(\delta), \delta)$ and $\log N(\delta) = \boldsymbol{N}(\theta(\delta), x(\delta), \delta)$ with respect to δ at

$\delta \to 0$, we have

$$\hat{n}_i = \boldsymbol{n}_{x_i}\hat{x}_i + \boldsymbol{n}_x\hat{x} + \boldsymbol{n}_\delta, \quad \hat{N} = \boldsymbol{N}_\theta\theta + \boldsymbol{N}_x\hat{x} + \boldsymbol{N}_\delta,$$

where $\hat{x} = \bar{\theta} + \epsilon$ and $\hat{x}_i = \theta + \frac{\epsilon_i}{\sqrt{\bar{z}}}$. Similarly, expansion of $v_i(\delta)$ gives

$$\hat{v}_i = \boldsymbol{v}_\theta(\bar{\theta} + \theta) + \boldsymbol{v}_N\hat{N} + \boldsymbol{v}_n\hat{n}_i,$$

where I use $\boldsymbol{v}_\theta = \boldsymbol{v}_{\bar{\theta}}$ at $\delta \to 0$.

Coefficients $\boldsymbol{n}_{x_i}, \boldsymbol{N}_\theta, \boldsymbol{n}_x, \boldsymbol{N}_x$ can be solved from the expansion of the equilibrium conditions in equations (9) to (10). Matching coefficients gives the standard "beauty-contest" result:

$$\boldsymbol{n}_{x_i} = \boldsymbol{N}_\theta = \frac{r\lambda(\bar{z})}{1 - s\lambda(\bar{z})}, \quad \boldsymbol{n}_x = \boldsymbol{N}_x = \frac{r\lambda_x}{1-s}, \tag{11}$$

where $\lambda(\bar{z}) \coloneqq \frac{\sigma_\theta^2}{\sigma_\theta^2 + 1/\bar{z}}$, $\lambda_x \coloneqq \frac{\sigma_{\bar{\theta}}^2}{\sigma_{\bar{\theta}}^2 + \sigma_\epsilon^2}$, and $\boldsymbol{n}_\delta = \boldsymbol{N}_\delta = 0$.

The first-order expansion of attention optimality does not determine $\bar{z}$ because

$$\mathbb{E}_0\Big[e^{(1-\tilde{\gamma})\bar{v}}\hat{v}_i\frac{1-\epsilon_i^2}{2\bar{z}} - \frac{e^{(1-\tilde{\gamma})\bar{v}}}{1-\tilde{\gamma}}\frac{1-\epsilon_i^2}{2\bar{z}}\hat{z}\Big] \equiv 0, \quad \forall \bar{z}.$$

To solve for $\bar{z}$, expand the attention optimality condition to the second order at $\delta \to 0$:

$$\mathbb{E}_0\Big[\,\big((1-\tilde{\gamma})\hat{v}_i^2 + \hat{\hat{v}}_i\big)\, e^{(1-\tilde{\gamma})\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big] - 2\kappa'(\bar{z}) = 0, \tag{12}$$

where

$$\hat{\hat{v}}_i = \boldsymbol{v}_{nn}\hat{n}_i^2 + \boldsymbol{v}_n\hat{\hat{n}}_i + \boldsymbol{v}_{\theta\theta}(\bar{\theta}+\theta)^2 + \boldsymbol{v}_{NN}\hat{N}^2 + 2\boldsymbol{v}_{\theta n}(\bar{\theta}+\theta)\hat{n}_i + 2\boldsymbol{v}_{\theta N}(\bar{\theta}+\theta)\hat{N} + 2\boldsymbol{v}_{nN}\hat{n}_i\hat{N} + \boldsymbol{v}_N\hat{\hat{N}}.$$

From direct calculation, $\boldsymbol{v}_{nn} = -(1-\eta)(1+\nu)$ and $\boldsymbol{v}_n = 0$. Moreover, because ϵ_i is independent of the aggregate variables, we have

Lemma 7 $\mathbb{E}_0[(1-\epsilon_i^2)\ \epsilon_i^m\ \bar{\theta}^h\theta^k\epsilon^l] = 0,\ \forall m \in \{2p+1|p \in \mathbb{N}\}$ *or* $m = 0$.

Substituting the expression for $\hat{\hat{v}}_i$ back into Equation 12 and using Lemma 7, the optimality condition for attention reduces to:

$$e^{(1-\tilde{\gamma})\bar{v}}\ |\boldsymbol{v}_{nn}|\ \Big(\frac{\boldsymbol{n}_{x_i}}{\bar{z}}\Big)^2 - 2\kappa'(\bar{z}) = 0. \tag{13}$$

With linear attention cost $\kappa(z) = \kappa\ z$, the solution is given by

$$\bar{z} = \frac{1}{1-s}\Big(\Big(\frac{r^2 e^{(1-\tilde{\gamma})\bar{v}}|\boldsymbol{v}_{nn}|}{2\kappa}\Big)^{\frac{1}{2}} - \frac{1}{\sigma_\theta^2}\Big).$$

Second-Order Expansion

From the expansion of $x_i(\delta) \approx \hat{x}_i\delta + \frac{1}{2}\hat{\hat{x}}_i\delta^2$, we have $\hat{x}_i = \theta + \frac{1}{\sqrt{\bar{z}}}\epsilon_i$ and $\hat{\hat{x}}_i = -\frac{\hat{z}}{\sqrt{\bar{z}}}\epsilon_i$.

The second-order expansions of individual and aggregate input and the first-order expansion of attention are:

$$\begin{aligned}
\hat{\hat{n}}_i &= \boldsymbol{n}_{x_ix_i}\hat{x}_i^2 + 2\boldsymbol{n}_{xx_i}\hat{x}\hat{x}_i + 2\boldsymbol{n}_{xx}\hat{x}^2 + \boldsymbol{n}_{\delta\delta} + \boldsymbol{n}_{x_i}\hat{\hat{x}}_i,\\
\hat{\hat{N}} &= \boldsymbol{N}_{\theta\theta}\theta^2 + 2\boldsymbol{N}_{x\theta}\hat{x}\theta + \boldsymbol{N}_{xx}\hat{x}^2 + \boldsymbol{N}_{\delta\delta},\\
\hat{z} &= \boldsymbol{z}_x\hat{x}.
\end{aligned}$$

Cross-derivatives for (x, δ) and (x_i, δ) are omitted for ease of exposition. It is easy to show that they are all zeros.

From the optimality condition of input, the second-order expansion gives

$$\hat{\hat{n}}_i = \mathbb{E}[2r\alpha\bar{\theta}\theta + s\hat{\hat{N}}|\bar{\mathcal{F}}_i] + 2\frac{d}{d\delta}\mathbb{E}[r(\bar{\theta}+\theta) + s\hat{N}|x(\delta), x_i(\delta)]\Big|_{\delta=0} - 2(\Gamma - \frac{1}{2})\ Var_i((\bar{\theta}+\theta) + \eta\hat{N}),$$

where $\hat{\mathbb{E}}_i[\cdot] := \frac{d}{d\delta}\mathbb{E}[\cdot|\mathcal{F}_i(\delta)]\big|_{\delta=0}$ and $\Gamma := \frac{(1+\nu)(1-\eta)}{\eta+\nu}\tilde{\gamma}$. The derivation of the equation above uses

$$\mathbb{E}[\hat{\hat{n}}_i|\bar{\mathcal{F}}_i] + 2\hat{\mathbb{E}}_i[\hat{n}_{i,t}] = \hat{\hat{n}}_{i,t},$$

which results from differentiating $\mathbb{E}[n_i(\delta)|\mathcal{F}_i(\delta)] = n_i(\delta)$ twice at $\delta \to 0$.

The first term on the right-hand-side can be expressed as

$$\mathbb{E}[2r\alpha\bar{\theta}\theta + s\hat{\hat{N}}|\bar{\mathcal{F}}_i] = s\boldsymbol{N}_{\theta\theta}\big(\big(\lambda\hat{x}_i\big)^2 + Var_i(\theta)\big) + 2\lambda\hat{x}_i(r\alpha\lambda_x + s\boldsymbol{N}_{x\theta})\hat{x} + s\boldsymbol{N}_{xx}\hat{x}^2 + s\boldsymbol{N}_{\delta\delta};$$

the second term as

$$\hat{\mathbb{E}}_i[r(\bar{\theta}+\theta) + s\hat{N}] = \boldsymbol{N}_\theta\big((1-\lambda)\hat{z}\hat{x}_i + \frac{1}{2}\hat{\hat{x}}_i\big),$$

where I use $\boldsymbol{N}_\theta = (r + s\boldsymbol{N}_\theta)\lambda$ and the expansion of $\lambda(\delta) = \frac{\sigma_\theta^2}{\sigma_\theta^2 + 1/z(\delta)}$:

$$\hat{\lambda} = \frac{\sigma_\theta^2}{(\sigma_\theta^2 + 1/\bar{z})^2}\frac{\hat{z}}{\bar{z}} = \lambda(1-\lambda)\hat{z}.$$

From the optimality of input:

$$\begin{aligned}
&\boldsymbol{n}_{x_ix_i}\hat{x}_i^2 + 2\boldsymbol{n}_{xx_i}\hat{x}\hat{x}_i + 2\boldsymbol{n}_{xx}\hat{x}^2 + \boldsymbol{n}_{\delta\delta} + \boldsymbol{n}_{x_i}\hat{\hat{x}}_i\\
&= s\boldsymbol{N}_{\theta\theta}\big(\big(\lambda\hat{x}_i\big)^2 + Var_i(\theta)\big) + 2\lambda\hat{x}_i(r\alpha\lambda_x + s\boldsymbol{N}_{x\theta})\hat{x} + s\boldsymbol{N}_{xx}\hat{x}^2 + s\boldsymbol{N}_{\delta\delta}\\
&\quad + 2\boldsymbol{N}_\theta\big((1-\lambda)\hat{x}_i\hat{z} + \frac{1}{2}\hat{\hat{x}}_i\big) - 2(\Gamma - \frac{1}{2})\ Var_i(\theta + \eta\hat{N}).
\end{aligned}$$

From the aggregation condition:

$$\boldsymbol{N}_{\theta\theta}\theta^2 + 2\boldsymbol{N}_{x\theta}\hat{x}\theta + \boldsymbol{N}_{xx}\hat{x}^2 + \boldsymbol{N}_{\delta\delta} = \boldsymbol{n}_{x_i x_i}\big(\theta^2 + \frac{1}{\sqrt{\bar{z}}}\big) + 2\boldsymbol{n}_{xx_i}\hat{x}\theta + 2\boldsymbol{n}_{xx}\hat{x}^2 + \boldsymbol{n}_{\delta\delta} + (1-\eta)\frac{\boldsymbol{n}_{x_i}^2}{\sqrt{\bar{z}}}.$$

Matching coefficients in the aggregation condition, we have $\boldsymbol{N}_{\theta\theta} = \boldsymbol{n}_{x_i x_i}$, $\boldsymbol{N}_{xx} = \boldsymbol{n}_{xx}$, and $\boldsymbol{N}_{x\theta} = \boldsymbol{n}_{xx_i}$.

From input optimality, matching coefficients implies $\boldsymbol{N}_{\theta\theta} = \boldsymbol{N}_{xx} = 0$. Moreover, because $\boldsymbol{N}_{x\theta} = \boldsymbol{n}_{xx_i}$, we have, from the terms involving (x, x_i),

$$\boldsymbol{N}_{x\theta} = \boldsymbol{n}_{xx_i} = \boldsymbol{N}_{\theta}\big(\alpha\lambda_x + \frac{1-\lambda}{1-s\lambda} \times \boldsymbol{z}_x\big). \tag{14}$$

From attention optimality, the third-order expansion gives

$$\mathbb{E}_0\Big[\,\Big((1-\tilde{\gamma})^2\hat{v}_i^3 + 3(1-\tilde{\gamma})\hat{v}_i\hat{\hat{v}}_i + \hat{\hat{\hat{v}}}_i - 3\hat{z}\left((1-\tilde{\gamma})\hat{v}_i^2 + \hat{v}_i\right)\Big)\, e^{(1-\tilde{\gamma})\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\,\Big] = 0,$$

where

$$\hat{\hat{\hat{v}}}_i = 3\hat{\hat{n}}_i(\boldsymbol{v}_{nn}\hat{n}_i + \boldsymbol{v}_{n\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nN}\hat{N}) + \boldsymbol{v}_{nnn}\hat{n}_i^3 + 3\boldsymbol{v}_{nn\theta}(\bar{\theta}+\theta)\hat{n}_i^2 + 3\boldsymbol{v}_{nnN}\hat{N}\hat{n}_i^2 + \cdots,$$

and the omitted terms are irrelevant to the calculation.

Using $e^{(1-\tilde{\gamma})\bar{v}}\ |\boldsymbol{v}_{nn}|\ \left(\frac{\boldsymbol{n}_{x_i}}{\bar{z}}\right)^2 = 2\kappa$, the condition reduces to

$$\mathbb{E}_0\Big[\Big(3(1-\tilde{\gamma})\hat{v}_i\hat{\hat{v}}_i + \hat{\hat{\hat{v}}}_i\Big)e^{(1-\tilde{\gamma})\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big] - 6\kappa\hat{z} = 0.$$

Substitute the expression for $\hat{\hat{\hat{v}}}_i$,

$$\begin{aligned}\mathbb{E}_0\Big[3\Big(\big((1-\tilde{\gamma})(\boldsymbol{v}_\theta(\bar{\theta}+\theta) + \boldsymbol{v}_N\hat{N})\boldsymbol{v}_{nn} + \boldsymbol{v}_{nnn}\hat{n}_i + \boldsymbol{v}_{nn\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nnN}\hat{N}\big)\boldsymbol{n}_{x_i}^2\frac{\epsilon_i^2}{\bar{z}} \\ + \hat{\hat{n}}_i(\boldsymbol{v}_{nn}\hat{n}_i + \boldsymbol{v}_{n\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nN}\hat{N})\Big)e^{(1-\tilde{\gamma})\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big] = 6\kappa\hat{z},\end{aligned} \tag{15}$$

where $\boldsymbol{v}_\theta = 1,\ \boldsymbol{v}_N = \eta$, and

$$\begin{aligned}\boldsymbol{v}_{nn} &= -(1-\eta)(1+\nu), \quad \boldsymbol{v}_{n\theta} = r(1-\eta)(1+\nu), \quad \boldsymbol{v}_{nN} = s(1-\eta)(1+\nu), \\ \boldsymbol{v}_{nnn} &= (2+\nu-\eta)\boldsymbol{v}_{nn}, \quad \boldsymbol{v}_{nn\theta} = (2+\nu-\eta)\boldsymbol{v}_{n\theta}, \quad \boldsymbol{v}_{nnN} = (2+\nu-\eta)\boldsymbol{v}_{nN}.\end{aligned}$$

Because $\epsilon_i \perp \bar{\theta}, \theta$ and $\mathbb{E}[\epsilon_i(\epsilon_i^2 - \epsilon_i^4)|\hat{x}, \bar{z}] = 0$, we have

$$\mathbb{E}_0[(\boldsymbol{v}_{nnn}\hat{n}_i + \boldsymbol{v}_{nn\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nnN}\hat{N})(\epsilon_i^2 - \epsilon_i^4)] = 0,$$

where I use

$$\mathbb{E}_0[(-\hat{n}_i + r(\bar{\theta}+\theta) + s\hat{N})(\epsilon_i^2 - \epsilon_i^4)] = \mathbb{E}_0[-\big(\boldsymbol{n}_{x_i}\big(\theta + \frac{\epsilon_i}{\sqrt{\bar{z}}}\big) + \boldsymbol{n}_x\hat{x}\big) + r(\bar{\theta}+\theta) + s\hat{N}]\mathbb{E}_0[\epsilon_i^2 - \epsilon_i^4]$$
$$= \mathbb{E}_0[\mathbb{E}_i[-\hat{n}_i + r(\bar{\theta}+\theta) + s\hat{N}]]\mathbb{E}_0[\epsilon_i^2 - \epsilon_i^4] = 0.$$

Terms in the second line of Equation 15 can be simplified as

$$\mathbb{E}_0\Big[\boldsymbol{n}_{x_ix_i}\hat{x}_i^2(\boldsymbol{v}_{nn}(\boldsymbol{n}_{x_i}\hat{x}_i + \boldsymbol{n}_x\hat{x}) + \boldsymbol{v}_{n\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nN}\hat{N})\frac{1-\epsilon_i^2}{2\bar{z}}\Big] = \mathbb{E}_0\Big[2\boldsymbol{n}_{x_ix_i}\theta \times \boldsymbol{v}_{nn}\frac{\boldsymbol{n}_{x_i}}{\bar{z}}\frac{\epsilon_i^2-\epsilon_i^4}{2\bar{z}}\Big],$$
$$\mathbb{E}_0\Big[2\boldsymbol{n}_{xx_i}\hat{x}_i(\boldsymbol{v}_{nn}\boldsymbol{n}_{x_i}\hat{x}_i + \boldsymbol{v}_{n\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nN}\hat{N})\frac{1-\epsilon_i^2}{2\bar{z}}\Big] = \mathbb{E}_0\Big[2\boldsymbol{n}_{xx_i} \times \boldsymbol{v}_{nn}\frac{\boldsymbol{n}_{x_i}}{\bar{z}}\frac{\epsilon_i^2-\epsilon_i^4}{2\bar{z}}\Big],$$
$$\mathbb{E}_0\Big[\boldsymbol{n}_{x_i}\hat{\hat{x}}_i(\boldsymbol{v}_{nn}\boldsymbol{n}_{x_i}\hat{x}_i + \boldsymbol{v}_{n\theta}(\bar{\theta}+\theta) + \boldsymbol{v}_{nN}\hat{N})\frac{1-\epsilon_i^2}{2\bar{z}}\Big] = \mathbb{E}_0\Big[-\boldsymbol{n}_{x_i}\hat{z} \times \boldsymbol{v}_{nn}\frac{\boldsymbol{n}_{x_i}}{\bar{z}}\frac{\epsilon_i^2-\epsilon_i^4}{2\bar{z}}\Big].$$

As a result, the attention optimality condition reduces

$$\mathbb{E}_0\Big[\Big(3(1-\tilde{\gamma})(\boldsymbol{v}_\theta(\bar{\theta}+\theta) + \boldsymbol{v}_N\hat{N}) + 3\Big(\frac{2\boldsymbol{n}_{x_ix_i}\theta + 2\boldsymbol{n}_{xx_i} - \boldsymbol{n}_{x_i}\hat{z}}{\boldsymbol{n}_{x_i}}\Big)\Big)e^{(1-\tilde{\gamma})\bar{v}}\boldsymbol{v}_{nn}\frac{\boldsymbol{n}_{x_i}^2}{\bar{z}}\ \frac{\epsilon_i^2-\epsilon_i^4}{2\bar{z}}\Big] = 6\kappa\hat{z}.$$

Using the solution of $\boldsymbol{n}_{xx_i}$ in Equation 14 and that $e^{(1-\tilde{\gamma})\bar{v}}\ |\boldsymbol{v}_{nn}|\ \left(\frac{\boldsymbol{n}_{x_i}}{\bar{z}}\right)^2 = 2\kappa$ and $\boldsymbol{n}_{x_ix_i} = 0$, we have

$$\big(3(1-\tilde{\gamma})\mathbb{E}_0[\boldsymbol{v}_\theta(\bar{\theta}+\theta) + \boldsymbol{v}_N\hat{N}] + 6\big(\alpha\lambda_x\hat{x} + \frac{1-\lambda}{1-s\lambda}\hat{z}\big) - 3\hat{z}\big) \times 2\kappa = 6\kappa\hat{z},$$

Rearranging gives

$$\big(3(1-\tilde{\gamma})(\boldsymbol{v}_\theta + \boldsymbol{v}_N\boldsymbol{N}_x) + 6\alpha\big) \times \lambda_x\hat{x} = 6\frac{(1-s)\lambda}{1-s\lambda} \times \hat{z}. \tag{16}$$

Volatility, Forecast Dispersion, and Uncertainty

Consider the measures of aggregate volatility, forecast dispersion, and subjective uncertainty in the sequence of economies indexed by δ :

$$SD(\widetilde{Y}(\delta)|\bar{\theta}(\delta)) = \Big(\mathbb{E}\Big[\big(\widetilde{Y}(\delta) - \mathbb{E}[\widetilde{Y}(\delta)|\bar{\theta}(\delta)]\big)^2\Big|\bar{\theta}(\delta)\Big]\Big)^{\frac{1}{2}},$$
$$Disp(\mathbb{E}_i[\widetilde{Y}(\delta)]) = \Big(\int\big(\mathbb{E}_i[\widetilde{Y}(\delta)] - \int\mathbb{E}_i[\widetilde{Y}(\delta)]\ di\big)^2 di\Big)^{\frac{1}{2}},$$
$$SD_i(\widetilde{Y}(\delta)) = \Big(\mathbb{E}_i\Big[\big(\widetilde{Y}(\delta) - \mathbb{E}_i[\widetilde{Y}(\delta)]\big)^2\Big]\Big)^{\frac{1}{2}}.$$

It is easy to show the zeroth-order expansions of the three measures are zeros.

To approximate the measures, note that the first-order expansions of equilibrium output are given by

$$\hat{Y} = \boldsymbol{Y}_{\bar{\theta}}\bar{\theta} + \boldsymbol{Y}_{\theta}\theta + \boldsymbol{Y}_{x}\hat{x} + \boldsymbol{Y}_{\delta},$$

where $\boldsymbol{Y}_{\bar{\theta}} = 1$, $\boldsymbol{Y}_x = \boldsymbol{N}_x$, and $\boldsymbol{Y}_\theta = 1 + \boldsymbol{N}_\theta$.

As a result, the first-order expansions of these measures are

$$\widehat{SD}(\widetilde{Y}|\bar{\theta}) = \Big(\mathbb{E}[(\hat{Y} - \mathbb{E}[\hat{Y}|\bar{\theta}])^2|\bar{\theta}]\Big)^{1/2} = \big(\boldsymbol{Y}_\theta^2\ \sigma_\theta^2 + \boldsymbol{Y}_x^2\ \sigma_\epsilon^2\big)^{1/2}, \tag{17}$$

$$\widehat{Disp}(\mathbb{E}_i[\widetilde{Y}]) = \Big(\int \Big(\mathbb{E}_i[\hat{Y}] - \int \mathbb{E}_i[\hat{Y}]di\Big)^2 di\Big)^{1/2} = \boldsymbol{Y}_\theta \lambda(\bar{z})\frac{1}{\sqrt{\bar{z}}}, \tag{18}$$

$$\widehat{SD}_i(\widetilde{Y}) = \Big(\mathbb{E}_i[(\hat{Y} - \mathbb{E}_i[\hat{Y}])^2]\Big)^{1/2} = \big(\boldsymbol{Y}_{\bar{\theta}}(1-\lambda_x)\sigma_{\bar{\theta}}^2 + \boldsymbol{Y}_\theta(1-\lambda(\bar{z}))\sigma_\theta^2\big)^{1/2}\ , \forall i. \tag{19}$$

Note that the first-order expansions do not depend on $\bar{\theta}$. To capture the state dependency, second-order expansion is necessary.

For second-order expansion of uncertainty, volatility, and forecast dispersion, we write the expansion of output as

$$\hat{\hat{Y}} = 2\boldsymbol{Y}_{\bar{\theta}\theta}\bar{\theta}\theta + 2\boldsymbol{Y}_{x\theta}\hat{x}\theta + \boldsymbol{Y}_{\delta\delta},$$

where $\boldsymbol{Y}_{\bar{\theta}\theta} = \alpha$ and $\boldsymbol{Y}_{x\theta} = \boldsymbol{N}_{x\theta}$. The other terms are omitted as they are zero from the second-order expansions of n_i and N.

Aggregate Volatility

Direct calculation gives

$$\widehat{\widehat{SD}}(\widetilde{Y}|\bar{\theta}) = \widehat{SD}(\widetilde{Y}|\bar{\theta})^{-1}\Big(\mathbb{E}[(\hat{Y} - \mathbb{E}[\hat{Y}|\bar{\theta}])(\hat{\hat{Y}} - \mathbb{E}[\hat{\hat{Y}}|\bar{\theta}])|\bar{\theta}]\Big).$$

From $\hat{Y} - \mathbb{E}[\hat{Y}|\bar{\theta}] = \boldsymbol{Y}_\theta\theta + \boldsymbol{Y}_x\epsilon$ and $\hat{\hat{Y}} - \mathbb{E}[\hat{\hat{Y}}|\bar{\theta}] = 2(\boldsymbol{Y}_{\bar{\theta}\theta}\bar{\theta} + \boldsymbol{Y}_{x\theta}\hat{x})\theta$,

$$\widehat{\widehat{SD}}(\widetilde{Y}|\bar{\theta}) = \widehat{SD}(\widetilde{Y}|\bar{\theta})^{-1} \times 2\boldsymbol{Y}_\theta(\boldsymbol{Y}_{\bar{\theta}\theta} + \boldsymbol{Y}_{x\theta})\bar{\theta}\sigma_\theta^2. \tag{20}$$

Note that when $\alpha = 0$ (no exogenous volatility), we have $\boldsymbol{Y}_{\bar{\theta}\theta} = 0$ and $\boldsymbol{Y}_{x\theta}$ reflects how agents' endogenous attention affects aggregate volatility through $\boldsymbol{z}_x$ in Equation 14.

Forecast Dispersion

Direct calculation gives

$$\widehat{\widehat{Disp}}(\mathbb{E}_i[\tilde{Y}]) = \widehat{Disp}(\mathbb{E}_i[\tilde{Y}])^{-1} \int \Big(\mathbb{E}[\hat{Y}|\bar{\mathcal{F}}_i] - \int \mathbb{E}[\hat{Y}|\bar{\mathcal{F}}_i]di\Big) \times\Big(\mathbb{E}[\hat{\hat{Y}}|\bar{\mathcal{F}}_i] + 2\hat{\mathbb{E}}_i[\hat{Y}] - \int \mathbb{E}[\hat{\hat{Y}}|\bar{\mathcal{F}}_i] + 2\hat{\mathbb{E}}_i[\hat{Y}]di\Big)di.$$

From

$$\mathbb{E}[\hat{Y}|\bar{\mathcal{F}}_i] - \int \mathbb{E}[\hat{Y}|\bar{\mathcal{F}}_i]di = \boldsymbol{Y}_\theta \lambda \frac{1}{\sqrt{\bar{z}}}\epsilon_i,$$

$$\mathbb{E}[\hat{\hat{Y}}|\bar{\mathcal{F}}_i] + 2\hat{\mathbb{E}}_i[\hat{Y}] - \int \mathbb{E}[\hat{\hat{Y}}|\bar{\mathcal{F}}_i] + 2\hat{\mathbb{E}}_i[\hat{Y}]di = 2(\boldsymbol{Y}_{\bar{\theta}\theta}\lambda_x + \boldsymbol{Y}_{x\theta})\hat{x}\frac{\lambda}{\sqrt{\bar{z}}}\epsilon_i + 2\boldsymbol{Y}_\theta\lambda(\frac{1}{2} - \lambda)\frac{\boldsymbol{z}_x\hat{x}}{\sqrt{\bar{z}}}\epsilon_i,$$

$$\widehat{\widehat{Disp}}(\mathbb{E}_i[\tilde{Y}]) = \widehat{Disp}(\mathbb{E}_i[\tilde{Y}])^{-1} \times 2\boldsymbol{Y}_\theta\Big((\boldsymbol{Y}_{\bar{\theta}\theta}\lambda_x + \boldsymbol{Y}_{x\theta})\hat{x} + \boldsymbol{Y}_\theta\big(\frac{1}{2} - \lambda\big)\boldsymbol{z}_x\hat{x}\Big)\frac{\lambda^2}{\bar{z}}. \tag{21}$$

When $\alpha = 0$, we have $\boldsymbol{Y}_{\bar{\theta}\theta} = 0$ and Equation 21 captures two forces: (1) $\boldsymbol{Y}_{x\theta} = \boldsymbol{N}_{x\theta}$ captures how forecast dispersion increases with the volatility of aggregate input, and (2) the direct effect of attention $\boldsymbol{z}_x$ can either increase or decrease forecast dispersion, depending on $\lambda \gtrless \frac{1}{2}$.

Subjective Uncertainty

Direct calculation gives

$$\widehat{\widehat{SD}}_i(\widetilde{Y}) = \widehat{SD}_i(\widetilde{Y})^{-1}\mathbb{E}\Big[(\hat{Y} - \mathbb{E}[\hat{Y}|\bar{\mathcal{F}}_i])(\hat{\hat{Y}} - \mathbb{E}[\hat{\hat{Y}}|\bar{\mathcal{F}}_i] - 2\hat{\mathbb{E}}_i[\hat{Y}])\Big|\bar{\mathcal{F}}_i\Big].$$

Because

$$\hat{\hat{Y}} - \mathbb{E}[\hat{\hat{Y}}|\bar{\mathcal{F}}_i] - 2\hat{\mathbb{E}}_i[\hat{Y}] = 2\boldsymbol{Y}_{\bar{\theta}\theta}(\bar{\theta} - \lambda_x\hat{x})\theta + 2(\boldsymbol{Y}_{\bar{\theta}\theta}\lambda_x\hat{x} + \boldsymbol{Y}_{x\theta})\hat{x}(\theta - \lambda\hat{x}_i) - 2\boldsymbol{Y}_\theta\lambda\Big((1-\lambda)\hat{x}_i - \frac{\epsilon_i}{2\sqrt{\bar{z}}}\Big)\boldsymbol{z}_x\hat{x},$$

we have

$$\widehat{\widehat{SD}}_i(\widetilde{Y}) = \widehat{SD}_i(\widetilde{Y})^{-1} \times 2\Big(\boldsymbol{Y}_{\bar{\theta}}\boldsymbol{Y}_{\bar{\theta}\theta}\lambda\hat{x}_i(1-\lambda_x)\sigma_{\bar{\theta}}^2 + \boldsymbol{Y}_\theta\big(\boldsymbol{Y}_{\bar{\theta}\theta}\lambda_x + \boldsymbol{Y}_{x\theta} - \frac{1}{2}\boldsymbol{Y}_\theta\lambda\boldsymbol{z}_x\big)\hat{x}(1-\lambda)\sigma_\theta^2\Big), \tag{22}$$

where the derivation uses

$$\mathbb{E}[(\theta - \lambda\hat{x}_i)\frac{\epsilon_i}{\sqrt{\bar{z}}}|\bar{\mathcal{F}}_i] = \mathbb{E}[(\theta - \lambda\hat{x}_i)(\hat{x}_i - \theta)|\bar{\mathcal{F}}_i] = -\mathbb{E}[(\theta - \lambda\hat{x}_i)\theta|\bar{\mathcal{F}}_i] = -(1-\lambda)\sigma_\theta^2.$$

When $\alpha = 0$, we have $\boldsymbol{Y}_{\bar{\theta}\theta} = 0$, and Equation 22 again reflects two forces: (1) $\boldsymbol{Y}_{x\theta}$ captures how the volatility of aggregate input affects agent's subjective uncertainty about aggregate output, and (2) the direct effect of attention $\boldsymbol{z}_x$ decreases uncertainty as agents learn about the aggregate state.

Proofs of Lemma 2, 3, and 4

Lemma 2: The second-order approximation is given by $\boldsymbol{N}_x, \boldsymbol{N}_\theta$ from Equation 11 and $\boldsymbol{N}_{x\theta}$ from Equation 14 with $\alpha = 0$.

Lemma 3: Because $\boldsymbol{v}_\theta, \boldsymbol{v}_N, \boldsymbol{N}_x > 0$, and $\alpha = 0$, Equation 16 implies $\tilde{\gamma} > 1 \iff \boldsymbol{z}_x < 0$. Lemma 3 follows from $\frac{\partial \hat{z}}{\partial \bar{\theta}} = \boldsymbol{z}_x$.

Lemma 4: In a fixed-attention economy without volatility shocks, $\boldsymbol{z}_x = 0$ and $\alpha = 0$ in Equation 14. As a result, Equation 20, 21, and 22 all equal to zero.

A.3 Proof of Theorem 1

Without exogenous volatility ($\alpha = 0$), Equation 20 implies, up to second-order, aggregate volatility decreases with $\bar{\theta}$

$$\frac{\partial}{\partial \bar{\theta}} SD(\widetilde{Y}|\bar{\theta}) < 0 \iff \boldsymbol{Y}_{x\theta} = \boldsymbol{N}_\theta \Big(\frac{1-\lambda}{1-s\lambda}\Big) \boldsymbol{z}_x < 0,$$

where $\boldsymbol{Y}_{x\theta}$ is given by the express in Equation 14, evaluated at $\alpha = 0$. This is true when $\tilde{\gamma} > 1$ since Lemma 3 implies $\boldsymbol{z}_x < 0$ in this case.

Using the solutions of $\boldsymbol{Y}_\theta$ and $\boldsymbol{Y}_{x\theta}$ for forecast dispersion and subjective uncertainty,

$$\frac{\partial}{\partial \bar{\theta}} Disp(\mathbb{E}_i[\widetilde{Y}]) < 0 \iff \Big(\boldsymbol{N}_\theta \frac{1-\lambda}{1-s\lambda} + (1+\boldsymbol{N}_\theta)\Big(\frac{1}{2} - \lambda\Big)\Big) \boldsymbol{z}_x < 0, \tag{23}$$

$$\frac{\partial}{\partial \bar{\theta}} SD_i(\widetilde{Y}) < 0 \iff \Big(\boldsymbol{N}_\theta \frac{1-\lambda}{1-s\lambda} - \frac{1}{2}(1+\boldsymbol{N}_\theta)\lambda\Big) \boldsymbol{z}_x < 0. \tag{24}$$

Using the solution of $\boldsymbol{N}_\theta$ from Equation 11 for forecast dispersion, we have

$$\frac{\partial}{\partial \bar{\theta}} Disp(\mathbb{E}_i[\widetilde{Y}]) < 0 \iff f_d(\lambda) := s(r-s)\lambda^3 + \frac{(s+4)(s-r)}{2}\lambda^2 + \frac{3r-2s-2}{2}\lambda + \frac{1}{2} > 0$$

Since

$$f_d(0) = 1/2 > 0, \quad f_d(1) = -(1-s)(1+r-s)/2 < 0,$$

and $r > s$ implies $f_d(-\infty) = -\infty$ and $f_d(\infty) = \infty$, we know that polynomial $f_d(\lambda)$ has exactly one root in $(0, 1)$. Moreover, since $f_d(1) = -(1-s)(1+r-s)/2 \to 0$ as $s \to 1$, the root goes to 1 as $s \to 1$.

Using the solution of $\boldsymbol{N}_\theta$ from Equation 11 again for subjective uncertainty, we have

$$\frac{\partial}{\partial\bar{\theta}} SD_i(\widetilde{Y}) < 0 \iff f_u(\lambda) := \frac{s(r-s)}{2}\lambda^3 + (s - \frac{3}{2}r)\lambda^2 + (r - \frac{1}{2})\lambda > 0$$

Since

$$f_u(0) = 0, \quad f_u(1) = -\frac{1}{2}(1-s)(1+r-s) < 0, \quad f_u'(0) = r - \frac{1}{2},$$

and $r > s$ implies $f_u(-\infty) = -\infty$ and $f_u(\infty) = \infty$, we know that if $r > 1/2$, then polynomial $f_u(\lambda)$ has exactly one root in $(0, 1)$. Moreover, since $f_u(1) = -\frac{1}{2}(1-s)(1+r-s) \to 0$ as $s \to 1$, the root goes to 1 as $s \to 1$.

Moreover, by comparing the expressions, we have

$$f_d(\lambda) - f_u(\lambda) = (1-\lambda)(1-s\lambda)(1+r\lambda-s\lambda) > 0.$$

This implies $f_u(\lambda) > 0 \implies f_d(\lambda) > 0$, and

$$\frac{\partial}{\partial\bar{\theta}} SD_i(\widetilde{Y}) < 0 \implies \frac{\partial}{\partial\bar{\theta}} Disp(\mathbb{E}_i[\widetilde{Y}]) < 0.$$

In other words, whenever attention increases uncertainty, it is necessarily accompanied by an increase in forecast dispersion.

A.4 Proof of Lemma 5

Rewrite $\beta_{CG}(\bar{\theta})$ as

$$\beta_{CG}(\bar{\theta}) = \frac{Cov(\tilde{Y} - \mathbb{E}[\tilde{Y}|x],\ \bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})}{Var(\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})} - 1.$$

First-order

The first-order expansions of the equilibrium are captured by the zeroth-order expansion of $\beta_{CG}(\bar{\theta})$. Because both the denominator and numerator of $\beta_{CG}(\bar{\theta})$ as well as their derivatives with respect to δ are all zeros at $\delta \to 0$, the limit is given by applying L'Hopital's rule twice:

$$\bar{\beta}_{CG}(\bar{\theta}) = \frac{Cov(\hat{Y} - \mathbb{E}[\hat{Y}|x], \bar{\mathbb{E}}[\hat{Y}] - \mathbb{E}[\hat{Y}|x]|\bar{\theta})}{Var(\bar{\mathbb{E}}[\hat{Y}] - \mathbb{E}[\hat{Y}|x]|\bar{\theta})} - 1.$$

From

$$\hat{Y} - \mathbb{E}[\hat{Y}|\hat{x}] = (\bar{\theta} - \lambda_x\hat{x}) + \boldsymbol{Y}_\theta\theta, \quad \bar{\mathbb{E}}[\hat{Y}] - \mathbb{E}[\hat{Y}|\hat{x}] = \boldsymbol{Y}_\theta\lambda\theta,$$

we have

$$\bar{\beta}_{CG}(\bar{\theta}) = \frac{1}{\lambda} - 1.$$

Second-order

The second-order expansions of the equilibrium are captured by the first-order expansion of the measure. By using $\bar{\beta}_{CG}(\bar{\theta})$, applying L'Hopital's rule, and rearranging the expression, we have

$$\hat{\beta}_{CG}(\bar{\theta}) = \frac{\frac{d^3}{d\delta^3}Cov(\tilde{Y} - \mathbb{E}[\tilde{Y}|x],\ \bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta}) - \frac{1}{\lambda}\frac{d^3}{d\delta^3}Var(\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})}{3\frac{d^2}{d\delta^2}Var(\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})}\Big|_{\delta=0}.$$

From the second-order expansion,

$$\hat{\hat{Y}} - \mathbb{E}[\hat{\hat{Y}}|\hat{x}] = 2\alpha\bar{\theta}\theta + 2\boldsymbol{Y}_{x\theta}\hat{x}\theta,$$
$$\bar{\mathbb{E}}[\hat{\hat{Y}}] + 2\frac{d}{d\delta}\bar{\mathbb{E}}[\hat{Y}|\delta]\Big|_{\delta=0} - \mathbb{E}[\hat{\hat{Y}}|\hat{x}] = 2\alpha\lambda_x\hat{x}\lambda\theta + 2\boldsymbol{Y}_{x\theta}\hat{x}\lambda\theta + 2\boldsymbol{Y}_{\theta}\lambda(1-\lambda)\hat{z}\theta.$$

Direct calculation gives

$$\begin{aligned}
&\frac{d^3}{d\delta^3}Cov(\tilde{Y} - \mathbb{E}[\tilde{Y}|x],\ \bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta}) - \frac{1}{\lambda}\frac{d^3}{d\delta^3}Var(\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})\Big|_{\delta=0} \\
&= 3Cov(\hat{\hat{Y}} - \mathbb{E}[\hat{\hat{Y}}|\hat{x}] - \frac{1}{\lambda}\big(\bar{\mathbb{E}}[\hat{\hat{Y}}] + 2\frac{d}{d\delta}\bar{\mathbb{E}}[\hat{Y}|\delta]\Big|_{\delta=0} - \mathbb{E}[\hat{\hat{Y}}|\hat{x}]\big),\ \bar{\mathbb{E}}[\hat{Y}] - \mathbb{E}[\hat{Y}|\hat{x}]|\bar{\theta}) \\
&= 6\alpha(1-\lambda_x)\bar{\theta}\ \boldsymbol{Y}_{\theta}\lambda\sigma_{\theta}^2 - 6\boldsymbol{Y}_{\theta}(1-\lambda)z_x\bar{\theta}\boldsymbol{Y}_{\theta}\lambda\sigma_{\theta}^2,
\end{aligned}$$

and

$$\frac{d^2}{d\delta^2}Var(\bar{\mathbb{E}}[\tilde{Y}] - \mathbb{E}[\tilde{Y}|x]|\bar{\theta})\Big|_{\delta=0} = 2Var(\bar{\mathbb{E}}[\hat{Y}] - \mathbb{E}[\hat{Y}|\hat{x}]|\bar{\theta}) = 2\boldsymbol{Y}_{\theta}^2\lambda^2\sigma_{\theta}^2.$$

As a result,

$$\hat{\beta}_{CG}(\bar{\theta}) = \big(\frac{1-\lambda_x}{\boldsymbol{Y}_{\theta}\lambda}\alpha - \frac{1-\lambda}{\lambda}z_x\big)\bar{\theta},$$

and Lemma 5 follows.

Online Appendix

B Generalization and Examples

The result that increased attention (the precision of private signal) can lead to higher volatility, uncertainty, and forecast dispersion is general and not tied to the specific assumptions on preference and production technology in Section 2.

Consider an economy with aggregate state $a(\bar{\theta}, \theta)$ and the information structure described in Section 2. Instead of endogenous attention, consider an exogenous increase in the precision of private signals as given by $z(\delta) = \bar{z} + \boldsymbol{z}_\delta \delta$, for some $\boldsymbol{z}_\delta > 0$ that parameterizes comparative statics with respect to the signal precision.

Given the information structure, an equilibrium of the economy is described by optimality and aggregation conditions of the following form:

$$\mathbb{E}\big[G(a(\delta) + \eta \tilde{N}(\delta), \tilde{n}_i(\delta))\big|\mathcal{F}_i(\delta)\big] = 0, \quad H(\tilde{N}(\delta)) = \int H(\tilde{n}_i(\delta))di,$$

where, at the steady state, function G satisfies

$$\frac{G_{11}}{G_1^2} + \frac{G_{22}}{G_2^2} - 2\frac{G_{12}}{G_1 G_2} = 0.$$

This condition allows the function G to be nonlinear in either of its argument, but requires the curve $y(x)$ described by $G(x, y) = 0$ to have curvature zero at the steady state. Under this condition, the best response under certainty features constant elasticities up to the second order. Consequently, the nonlinearity in individual responses results only from changes in the information structure. The condition holds for the setup in Sectoin 2 and a few extensions in Appendix B.1.

For this class of models, the second-order expansion of the equilibrium is described by

$$\hat{n}_t = \mathbb{E}[r\hat{\hat{a}} + s\hat{\hat{N}}|\bar{\mathcal{F}}_i] + 2\frac{d}{d\delta}\mathbb{E}[r\hat{a} + s\hat{N}|\mathcal{F}_i(\delta)] + const, \quad \hat{\hat{N}} = \int \hat{\hat{n}}_i di + const.$$

where $r = \frac{G_1}{-G_2}$ and $s = \eta\frac{G_1}{-G_2}$. Since the expansion is described by the same system as that in Section 3 up to a reparameterization of r and s, the same second-order approximation of the aggregate action $\tilde{N}$ is described by the same formula is in Lemma 2.

To study how a change in the precision of private signal affects the volatility, uncertainty, and forecast dispersion of aggregate outcomes, consider some aggregate variable $\widetilde{Y}$ of the form

$$\widetilde{Y} = \varsigma_a a + \varsigma_N \tilde{N}$$

for some $\varsigma_a, \varsigma_N > 0$. If $\varsigma_N = 0$, the aggregate variable captures the exogenous state a; if $\varsigma_a = 0$, it reflects only the endogenous aggregate response $\tilde{N}$. Following the same steps as the proof for Theorem 1 gives the following result:

Theorem 1A *Suppose that $z(\delta) = \bar{z} + \boldsymbol{z}_\delta \delta$, with $\boldsymbol{z}_\delta > 0$ and $\frac{G_{11}}{G_1^2} + \frac{G_{22}}{G_2^2} - 2\frac{G_{12}}{G_1 G_2} = 0$, then up to second-order approximation,*

$$(1) \quad \frac{\partial}{\partial \delta} SD(\widetilde{Y}|\bar{\theta}) < 0.$$

Moreover, there exists a threshold $\zeta \in \mathbb{R}$ such that if $\bar{z} < \zeta$, then

$$(2) \quad \frac{\partial}{\partial \delta} Disp(\mathbb{E}_i[\widetilde{Y}]) < 0, \quad and \quad (3) \quad \frac{\partial}{\partial \delta} SD(\widetilde{Y}|\mathcal{F}_i) < 0.$$

The threshold $\zeta > 0$ if and only if $r > \frac{\varsigma_a}{2\varsigma_N}$, and $\zeta \to \infty$ as $s \to 1$. In addition, (2) is true only if (3) is true.

If $\varsigma_a = 0$, that is, when aggregate variable $\tilde{Y}$ only reflects the endogenous response $\tilde{N}$, we have $r > 0$ and therefore $\zeta > 0$. In this case, an increase in z always increases subjective uncertainty when $\bar{z}$ is low enough (and therefore also increases forecast dispersion and volatility). On the other hand, if $\varsigma_N = 0$, aggregate variable $\tilde{Y}$ only concerns the exogenous state. In this case, agents' uncertainty about $\tilde{Y}$ always decreases in z.

B.1 Examples

Alternative Preference

Consider the following preference that generalizes the specification in Section 2:

$$v(c_i, n_i) = \left(c_i^{1-\sigma} - (1-\sigma)\frac{n_i^{1+\nu}}{1+\nu}\right)^{\frac{1}{1-\sigma}}$$

Parameter σ governs the income effect on labor; $\sigma = 0$ corresponds to the case in Section 2 and labor is procyclical if $\sigma \in [0, 1)$.

$$\mathbb{E}_i\left[\left((1-\sigma)(1-\eta)e^{(1-\sigma)a(\bar{\theta},\theta)} N^{(1-\sigma)\eta} n_i^{(1-\sigma)(1-\eta)-1}\right) \frac{u_c(c_i, n_i)}{\mathbb{E}_i[u_c(c_i, n_i)]}\right] = n_i^\nu,$$

where $u_c(c_i, n_i) = v_i^{\sigma-\tilde{\gamma}} \mathbf{1}_{\{v_i > \underline{u}\}}$.

This is equivalent to having

$$r = \frac{1-\sigma}{(1-\sigma)\eta + \sigma + \nu}, \quad s = \frac{(1-\sigma)\eta}{(1-\sigma)\eta + \sigma + \nu}.$$

Alternative Technology

Consider a simple "trade model" where agents can combine labor input n_i with purchases of final goods m_i to produce intermediate goods q_i with a roundabout production function:

$$q_i = \left(\frac{n_i}{\varpi}\right)^{\varpi} \left(\frac{m_i}{1-\varpi}\right)^{1-\varpi}, \quad c_i = p_i q_i - m_i,$$

where $n_i, m_i \in \mathcal{F}_i$ and the timing of decision is identical to that in Section 2.

The derivation below shows that the technology is equivalent to a reparameterization of ν in Section 2, demonstrating that agents' production inputs in Section 2 do not have to be taken literally as their labor.

For a given level of q_i, consumption net of labor cost is

$$c_i - \frac{n_i^*(q_i)^{1+\nu}}{1+\nu} = e^{a(\bar{\theta},\theta)} q_i^{1-\eta} Q^{\eta} - m_i^*(q_i) - \frac{n_i^*(q_i)^{1+\nu}}{1+\nu},$$

where $n_i^*(q_i)$ and $m_i^*(q_i)$ solves the following cost minimization problem given q_i:

$$\min_{m_i, n_i} \quad m_i + \frac{n_i^{1+\nu}}{1+\nu} \quad \text{s.t.} \quad q_i = \left(\frac{n_i}{\varpi}\right)^{\varpi} \left(\frac{m_i}{1-\varpi}\right)^{1-\varpi}.$$

Optimality requires

$$m_i = (1-\varpi)\mu q_i, \quad n_i^{1+\nu} = \varpi \mu q_i,$$

and solving for the multiplier gives $\mu = (\varpi q_i)^{\tilde{\nu}}$. Substituting back, the cost of producing q_i is:

$$m_i^*(q_i) + \frac{n_i^*(q_i)^{1+\nu}}{1+\nu} = \frac{\varpi^{\tilde{\nu}}}{1+\tilde{\nu}} q_i^{1+\tilde{\nu}}, \quad \text{where} \quad \tilde{\nu} := \frac{\varpi \nu}{1 + (1-\varpi)\nu}.$$

This is equivalent to having

$$r = \frac{1}{\eta + \tilde{\nu}}, \quad s = \frac{\eta}{\eta + \tilde{\nu}}.$$

B.2 Alternative Shocks: Sentiment

To demonstrate that the same reasoning behind Theorem 1*A* does not rely on the specific type of shocks studied in Section 2, I show below that the same connection between information and uncertainty, volatility, and forecast dispersion exists for extrinsic "sentiment shocks" in a simplified version of Angeletos and La'O (2013).

Consider the following environment: Agents (i, j) match pairwise. Agent i's consumption c_i is a composite of l_i units of "home good" i and l_i^* units of "foreign goods" j:

$$v(c_i, n_i) = c_i - \frac{n_i^{1+\nu}}{1+\nu}, \quad c_i = \left(\frac{\ell_i}{\varpi}\right)^{\varpi} \left(\frac{\ell_i^*}{1-\varpi}\right)^{1-\varpi}.$$

Each agent i produces good i with linear technology $A_i n_i$. Productivity $A_i \sim N(0, 1)$ is idiosyncratic and unknown to the agent. The budget constraint is given by

$$p_i^* \ell_i^* \leq p_i(A_i n_i - \ell_i),$$

where prices of the two goods p_i and p_i^* are normalized such that the multiplier on the budget constraint is one.

Agent i's information set consists of signals s_i and x_i:

$$s_i = A_i + e_i, \quad x_i = s_j + \theta + \frac{\epsilon_i}{\sqrt{z}},$$

where $e_i \sim N(0, \sigma_e^2)$ and $\epsilon_i \sim N(0, 1)$ are idiosyncratic noises and $\theta \sim N(0, \sigma_\theta^2)$ is an aggregate noise that represents agents' "sentiment". Signal s_i provides agent i with information about their own productivity, and signal x_i represents agent i's information about agent j's signal s_j.

From Angeletos and La'O (2013) (p. 771 and 772), aggregate output is given by

$$\tilde{Y} = \bar{Y} + \boldsymbol{Y}_\theta \theta,$$

where

$$\boldsymbol{Y}_\theta = \frac{(1+\nu)(1-\varrho)\varrho^2\sigma_e^2}{(1-\varrho^2)(\sigma_e^4 + z^{-1} + (1-\varrho)\sigma_\theta^2) + \sigma_e^2(1-\varrho^2 + z^{-1} + (1-\varrho)\sigma_\theta^2)}$$

and $\varrho := \frac{1-\varpi}{1-\varpi+\nu}$.

While their model is not isomorphic to the one in Section 2, the result from Theorem $1A$ extends to their setup:

- Aggregate volatility $SD(\tilde{Y}|z) = \boldsymbol{Y}_\theta \sigma_\theta^2$ is increasing in z, since $\boldsymbol{Y}_\theta$ is increasing z.
- There exists a threshold ζ such that forecast dispersion $Disp(\mathbb{E}[\tilde{Y}|s_i, x_i, z])$ and subjective uncertainty $SD(\tilde{Y}|s_i, x_i, z)$ are increasing in z for $z < \zeta$:

Because

$$\frac{\partial}{\partial z} SD(\tilde{Y}|s_i, x_i, z) > 0 \iff const. + z^{-1}(\cdots) + z^{-2}\left(\sigma_e^2 + \frac{2(1-\varpi)\nu + \nu^2}{(1-\varpi+\nu)^2}\right) > 0$$
$$\frac{\partial}{\partial z} Disp(\mathbb{E}[\tilde{Y}|s_i, x_i, z]) > 0 \iff const. + z^{-1}(\cdots) + z^{-2}\Big(\sigma_e^2 + \frac{2(1-\varpi)\nu + \nu^2}{(1-\varpi+\nu)^2}\Big) > 0,$$

these conditions hold as $z \to 0$.

- When $\nu \to 0$, we have $\boldsymbol{Y}_\theta \to \frac{z}{1-\varpi}$, and

$$\frac{\partial}{\partial z} SD(\tilde{Y}|s_i, x_i, z) = \frac{4(\sigma_e^2 + z^{-1})^2\sigma_\theta^2 + 2(2\sigma_e^2 + z^{-1})\sigma_\theta^4}{(1-\varpi)^2 z^{-1}(\sigma_e^2 + z^{-1} + \sigma_\theta^2)^2} > 0,$$
$$\frac{\partial}{\partial z} Disp(\mathbb{E}[\tilde{Y}|s_i, x_i, z]) = \frac{(3z^{-1} + \sigma_e^2 + \sigma_\theta^2)\sigma_\theta^4}{(1-\varpi)^2(\sigma_e^2 + z^{-1} + \sigma_\theta^2)^3} > 0.$$

Therefore, in this case, increasing signal precision always increases agents' forecast dispersion and their uncertainty about aggregate output.

C Dynamic Model: Derivations and Proofs

In this Appendix, I derive equilibrium conditions for the dynamic economy in Lemma 6, and then I derive the expansions of these conditions that characterize the first- and second-order expansions of the equilibrium objects.

C.1 Proof of Lemma 6

Let S_i^t be the collection of possible histories of signals and prices agent i receive before taking actions in period t, and denote a typical element of S_i^t by s_i^t :

$$s_i^t := \{x_i^t, \tilde{p}_i^{t-1}\}, \ \forall t \geq 0,$$

where $\tilde{p}_{i,t} = \theta_t + \vartheta_t + \eta \log N_t + \omega_{i,t}$ is a transformation of $p_{i,t}$ that contains the same information. Similarly, let S'^{t-1}_i be a collection of histories, s'^{t-1}_i, up to the start of period t:

$$s'^{t-1}_i := \{x_i^{t-1}, \tilde{p}_i^{t-1}\}, \forall t \geq 0.$$

A strategy is a sequence of mappings $\{z_t, n_t\}_{t=0}^\infty$ such that

$$z_t : S'^{t-1}_i \to \mathbb{R}_+, \quad n_t : S_i^t \to \mathbb{R}_+.$$

Write agents' period payoff as $V(\theta+\vartheta, N, n, \omega, z) := U(c(\theta+\vartheta, N, n, \omega), n, z)$. Denote the distribution of $\omega^\tau, \omega_i^\tau, s_i^\tau$ conditional on ${s'}_i^{t-1}, z_i^t$ and s_i^t, z_i^t as

$$\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | {s'}_i^{t-1}, z_i^t), \ \Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | s_i^t, z_i^t).$$

Proof. A strategy $\{n_t, z_t\}_{t=0}^\infty$ is optimal for agent i only if, $\forall \tilde{n}, \tilde{z} \in \mathbb{R}_+$ and history s_i^t,

$$\begin{aligned}
&\sum_{\tau=t}^{\infty} \beta^{\tau-t} \int V(\theta_\tau+\vartheta_\tau, N_\tau, \boldsymbol{n}_\tau(s_i^\tau), \omega_{i,\tau}, \boldsymbol{z}_\tau({s'}_i^{\tau-1})) d\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | {s'}_i^{t-1}, \boldsymbol{z}^t({s'}_i^{t-1})) \\
&\geq \sum_{\tau=t+1}^{\infty} \beta^{\tau-t} \int V(\theta_\tau+\vartheta_\tau, N_\tau, \boldsymbol{n}_\tau(s_i^\tau), \omega_{i,\tau}, \boldsymbol{z}_\tau({s'}_i^{\tau-1})) d\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | {s'}_i^{t-1}, \boldsymbol{z}^{t-1}({s'}_i^{t-2}), \tilde{z}) \\
&\quad + \int V(\theta_t+\vartheta_t, N_t, \boldsymbol{n}_t(s_i^t), \omega_{i,t}, \tilde{z}) d\Phi(\omega^t, \omega_i^t, s_i^t | {s'}_i^{t-1}, \boldsymbol{z}^{t-1}({s'}_i^{t-2}), \tilde{z}),
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{\tau=t}^{\infty} \beta^{\tau-t} \int V(\theta_\tau+\vartheta_\tau, N_\tau, \boldsymbol{n}_\tau(s_i^\tau), \omega_{i,\tau}, \boldsymbol{z}_\tau({s'}_i^{\tau-1}))\ d\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | s_i^t, \boldsymbol{z}^t({s'}_i^{t-1})) \\
&\geq \sum_{\tau=t+1}^{\infty} \beta^{\tau-t} \int V(\theta_\tau+\vartheta_\tau, N_\tau, \boldsymbol{n}_\tau(s_i^\tau), \omega_{i,\tau}, \boldsymbol{z}_\tau({s'}_i^{\tau-1}))\ d\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | s_i^t, \boldsymbol{z}^t({s'}_i^{t-1})) \\
&\quad + \int V(\theta_t+\vartheta_t, N_t, \tilde{n}, \omega_{i,t}, z_t({s'}_i^{t-1}))\ d\Phi(\omega^t, \omega_i^t, s_i^t | s_i^t, \boldsymbol{z}^t({s'}_i^{t-1})),
\end{aligned}$$

where

$$N_t = \left(\int \left(e^{\omega_{i,t}}\ \boldsymbol{n}_t(s_i^t) \right)^{1-\eta} d\Phi(\omega_{i,t}, s_i^t | \omega^t, \boldsymbol{z}^t({s'}_i^{t-1})) \right)^{\frac{1}{1-\eta}} \in \boldsymbol{\sigma}(\omega^t).$$

The following two first-order conditions follow

$$\begin{aligned}
&\frac{\partial}{\partial \tilde{z}} \Bigg(\sum_{\tau=t+1}^{\infty} \beta^{\tau-t} \int V(\theta_\tau+\vartheta_\tau, N_\tau, n_{i,\tau}, \omega_{i,\tau}, z_{i,\tau}) d\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | {s'}_i^{t-1}, z_i^{t-1}, \tilde{z}) + \\
&\quad + \int V(\theta_\tau+\vartheta_\tau, N_\tau, n_{i,\tau}, \omega_{i,\tau}, z_{i,\tau}) d\Phi(\omega^\tau, \omega_i^\tau, s_i^\tau | {s'}_i^{t-1}, z_i^{t-1}, \tilde{z}) \Bigg) \Bigg|_{\tilde{z}=z_{i,t}} = 0, \\
&\frac{\partial}{\partial \tilde{n}} \left(\int V(\theta_t+\vartheta_t, N_t, \tilde{n}, \omega_{i,t}, z_{i,t})\ d\Phi(\omega^t, \omega_i^t | s_i^t, z_i^t) \right) \Bigg|_{\tilde{n}=n_{i,t}} = 0.
\end{aligned}$$

The F.O.C. with respect to $\tilde{n}$ gives the second condition in Lemma 6.

For the F.O.C. with respect to $\tilde{z}$, let $\varphi(\cdot|\cdot)$ denote the density of $\Phi(\cdot|\cdot)$; then

$$\begin{aligned}&\frac{\partial}{\partial \tilde{z}}\varphi(\omega^\tau, \omega_i^\tau, s_i^\tau | s_i'^{t-1}, z_i^{t-1}, \tilde{z})\\ =&\frac{\partial}{\partial \tilde{z}}\varphi(x_{i,t}|\tilde{z}, \omega^t)\varphi(\tilde{p}_{i,t}|\omega_{i,t}, \omega^t) \prod_{l=t+1}^{\tau} \varphi(x_{i,l}|z_{i,l}, \omega^l)\varphi(\tilde{p}_{i,l}|\omega_{i,l}, \omega^l)\ \varphi(\omega^\tau, \omega_i^\tau | s_i'^{t-1}, z^{t-1})\end{aligned}$$

Differentiating with respect to $\tilde{z}$ and evaluating at $\tilde{z} = z_{i,t}$ gives

$$\frac{\partial}{\partial \tilde{z}}\varphi(x_{i,t}|\tilde{z}, \omega^t)\Big|_{\tilde{z}=z_{i,t}} = \frac{\partial}{\partial \tilde{z}}\phi\left(\frac{x_{i,t} - \theta_t(\omega^t)}{1/\sqrt{\tilde{z}}}\right)\Big|_{\tilde{z}=z_{i,t}} = \frac{1-\epsilon_{i,t}^2}{2z_{i,t}}\varphi(x_{i,t}|z_{i,t}, \omega^t),$$

where $\phi(\cdot)$ denotes the density of standard normal distribution. As a result,

$$\frac{\partial}{\partial \tilde{z}}\varphi(\omega^\tau, \omega_i^\tau, s_i^\tau | s_i'^{t-1}, z_i^{t-1}, \tilde{z})\Big|_{\tilde{z}=z_{i,t}} = \frac{1-\epsilon_{i,t}^2}{2z_{i,t}}\varphi(\omega^\tau, \omega_i^\tau, s_i^\tau | s_i'^{t-1}, z_i^t),$$

and the F.O.C. for $\tilde{z}$ is given by the first equation in Lemma 6. □

C.2 Expansions of Equilibrium Conditions

Taking derivatives of the conditions in Lemma 6 with respect to δ to corresponding orders and at $\delta = 0$ gives the systems that characterize the expansions of equilibrium objects.

Write the utility function in Equation 7 as

$$V(\theta_t + \vartheta_t, N_t, n_{i,t}, \omega_{i,t}, z_{i,t}) = f\big(v(\theta_t + \vartheta_t, N_t, n_{i,t}, \omega_{i,t}) - \kappa(z_{i,t})\big),$$

where $v(\theta_t + \vartheta_t, N_t, n_{i,t}, \omega_{i,t}) := u(c(\theta_t + \vartheta_t, N_t, n_{i,t}, \omega_{i,t}), n_{i,t})$ represents the payoff from consumption and labor.

The input and attention optimality conditions are, respectively,

$$\mathbb{E}\Big[f'\Big(v_{i,t}(\delta) - \delta^2\kappa(z_{i,t}(\delta))\Big) \times \frac{\partial}{\partial n}v_{i,t}(\delta)\Big|\mathcal{F}_{i,t}(\delta)\Big] = 0,$$

$$\sum_{\tau=t}^{\infty}\beta^{\tau-t}\mathbb{E}\Big[f(\cdot)\frac{1-\epsilon_{i,t}^2}{2z_{i,t}(\delta)}\Big|\mathcal{F}'_{i,t-1}(\delta)\Big] - \mathbb{E}[f'(\cdot)\Big|\mathcal{F}'_{i,t-1}(\delta)]\delta^2\kappa'(z_{i,t}(\delta)) = 0.$$

The following derivation shows that function $f(\cdot)$ affects the equilibrium only up to a constant for the first- and second-order approximation:

Input Optimality

The first- and second-order expansions of the input optimality conditions are given by

$$\mathbb{E}\Big[\bar{f}'\frac{d}{d\delta}\frac{\partial}{\partial n}v_{i,t}(\delta)\Big|\bar{\mathcal{F}}_{i,t}\Big] = 0,$$

and

$$\mathbb{E}\Big[\bar{f}'\frac{d^2}{d\delta^2}\frac{\partial}{\partial n}v_{i,t}(\delta) + 2\bar{f}''\frac{d}{d\delta}v_{i,t}(\delta)\frac{d}{d\delta}\frac{\partial}{\partial n}v_{i,t}(\delta)\Big|\bar{\mathcal{F}}_{i,t}\Big] + 2\hat{\mathbb{E}}_{i,t}\Big[\bar{f}'\frac{d}{d\delta}\frac{\partial}{\partial n}v_{i,t}(\delta)\Big] = 0. \tag{25}$$

Note that in Equation 25, the term multiplying $\bar{f}''$ is

$$\begin{aligned}
&\mathbb{E}\Big[\frac{d}{d\delta}v_{i,t}(\delta)\frac{d}{d\delta}\frac{\partial}{\partial n}v_{i,t}(\delta)\Big|\bar{\mathcal{F}}_{i,t}\Big]\\
&= \mathbb{E}\Big[\big(v_\theta(\hat{\theta}_t+\hat{\vartheta}_t) + v_N\hat{N}_t + v_\omega\hat{\omega}_{i,t}\big)\big(v_{n\theta}(\hat{\theta}_t+\hat{\vartheta}_t) + v_{nN}\hat{N}_t + v_{n\omega}\hat{\omega}_{i,t} + v_{nn}\hat{n}_{i,t}\big)\Big|\bar{\mathcal{F}}_{i,t}\Big]\\
&= Cov\big[v_\theta(\hat{\theta}_t+\hat{\vartheta}_t) + v_N\hat{N}_t + v_\omega\hat{\omega}_{i,t}, v_{n\theta}(\hat{\theta}_t+\hat{\vartheta}_t) + v_{nN}\hat{N}_t + v_{n\omega}\hat{\omega}_{i,t}\big|\bar{\mathcal{F}}_{i,t}\big],
\end{aligned}$$

where the second equation uses the solution of $\hat{n}_{i,t}$. The expression is a constant because the variables are Gaussian. Therefore, Equation 25 reduces to

$$\mathbb{E}\Big[\frac{d^2}{d\delta^2}\frac{\partial}{\partial n}v_{i,t}(\delta)\Big|\bar{\mathcal{F}}_{i,t}\Big] + 2\hat{\mathbb{E}}_{i,t}\Big[\frac{d}{d\delta}\frac{\partial}{\partial n}v_{i,t}(\delta)\Big] + const. = 0.$$

Attention Optimality

The second- and third-order expansions of the attention optimality condition are given by

$$\sum_{\tau=t}^{\infty}\beta^{\tau-t}\mathbb{E}\Big[\Big\{\bar{f}'\Big(\frac{d^2}{d\delta^2}v_{i,\tau}(\delta) - 2\kappa(\bar{z})\Big) + \bar{f}''\Big(\frac{d}{d\delta}v_{i,\tau}(\delta)\Big)^2\Big\}\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big|\bar{\mathcal{F}}'_{i,t-1}\Big] = 2\bar{f}'(\cdot)\kappa'(\bar{z}),$$

and

$$\begin{aligned}
&\sum_{\tau=t}^{\infty}\beta^{\tau-t}\mathbb{E}\Big[\Big\{\bar{f}'\Big(\frac{d^3}{d\delta^3}v_{i,\tau}(\delta) - 2\kappa'(\bar{z})\hat{z}_{i,\tau}\Big) + \bar{f}'''\Big(\frac{d}{d\delta}v_{i,\tau}(\delta)\Big)^3\\
&\qquad + 3\bar{f}''\frac{d}{d\delta}v_{i,\tau}(\delta)\Big(\frac{d^2}{d\delta^2}v_{i,\tau}(\delta) - 2\kappa(\bar{z})\Big)\Big\}\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big|\bar{\mathcal{F}}'_{i,t-1}\Big]\\
&= 2\bar{f}'\kappa''(\bar{z})\hat{z}_{i,t} + 6\bar{f}''\mathbb{E}\Big[\frac{d}{d\delta}v_{i,t}(\delta)\Big|\bar{\mathcal{F}}'_{i,t-1}\Big]\kappa'(\bar{z}).
\end{aligned} \tag{26}$$

Because $\bar{v}_n = 0$, $\frac{d}{d\delta}v_{i,\tau}$ does not contain any term involving $\hat{n}_{i,\tau}$. The second-order expansion reduces to

$$\sum_{\tau=t}^{\infty}\beta^{\tau-t}\mathbb{E}\Big[\Big\{\frac{d^2}{d\delta^2}v_{i,\tau}(\delta)\Big\}\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big|\bar{\mathcal{F}}'_{i,t-1}\Big]\Big] = 2\kappa'(\bar{z}).$$

For the third order, using again that $\bar{v}_n = 0$ and $\frac{d}{d\delta}v_{i,\tau}$ does not contain any term involving $\hat{n}_{i,\tau}$, the term multiplying $\bar{f}''$ on the left-hand side of Equation 26 reduces to

$$\begin{aligned}
&\sum_{\tau=t}^{\infty}\beta^{\tau-t}3\bar{f}''\mathbb{E}\Big[\Big\{\frac{d}{d\delta}v_{i,\tau}(\delta)\Big(\frac{d^2}{d\delta^2}v_{i,\tau}(\delta)-2\kappa(\bar{z})\Big)\Big\}\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big|\bar{\mathcal{F}}'_{i,t-1}\Big] \\
&= \sum_{\tau=t}^{\infty}\beta^{\tau-t}3\bar{f}''\mathbb{E}\Big[\frac{d}{d\delta}v_{i,\tau}(\delta)\Big|\bar{\mathcal{F}}'_{i,t-1}\Big]\mathbb{E}\Big[\frac{d^2}{d\delta^2}v_{i,\tau}(\delta)\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big|\bar{\mathcal{F}}'_{i,t-1}\Big] \\
&= 6\bar{f}''\mathbb{E}\Big[\frac{d}{d\delta}v_{i,t}(\delta)\Big|\bar{\mathcal{F}}'_{i,t-1}\Big]\kappa'(\bar{z}),
\end{aligned}$$

which equals the term multiplying $\bar{f}''$ on the right-hand side of Equation 26.

As a result, Equation 26 simplifies to

$$\sum_{\tau=t}^{\infty}\beta^{\tau-t}\mathbb{E}_{i,t-1}\Big[\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big\{\frac{d^3}{d\delta^3}u_{i,\tau}(\delta)-2\kappa'(\bar{z})\hat{z}_{i,\tau}\Big\}\Big] = 2\kappa''(\bar{z})\hat{z}_{i,t}.$$

Exogenous Processes

For the exogenous processes θ_t and ϑ_t, we have

$$\hat{\theta}_{t+1} = \rho\hat{\theta}_t + \omega_{t+1}, \quad \hat{\vartheta}_{t+1} \equiv 0, \quad \hat{\hat{\theta}}_{t+1} \equiv 0, \quad \hat{\hat{\vartheta}}_{t+1} = \rho\hat{\hat{\vartheta}}_t + 2\bar{\Sigma}'\hat{\theta}_t\omega_{t+1}. \tag{27}$$

Summary of Equilibrium Conditions

The following two lemmas summarize the expansions of equilibrium conditions that characterize the equilibrium up to the second order:

Lemma 8 *The first-order expansions of input $\hat{n}_{i,t}, \hat{N}_t$ and the zeroth-order expansion of attention $\bar{z}$ solve the following system:*

$$\sum_{\tau=t}^{\infty}\beta^{s-t}\mathbb{E}\Big[\hat{v}_{i,\tau}\frac{1-\epsilon_{i,t}^2}{2\bar{z}}\Big|\bar{\mathcal{F}}'_{i,t-1}\Big] = 2\kappa, \quad \mathbb{E}\big[\hat{v}_{n,i,t}\big|\bar{\mathcal{F}}_{i,t}\big] = 0, \quad \hat{N}_t = \int \hat{n}_{i,t},$$

where

$$\hat{v}_{n,i,t} = \nabla\bar{v}_n\begin{pmatrix}\hat{\theta}_t\\ \hat{N}_t\\ \hat{n}_{i,t}\\ \hat{\omega}_{i,t}\end{pmatrix}, \quad \hat{v}_{i,t} = \nabla\bar{v}\begin{pmatrix}\hat{\hat{\vartheta}}_t\\ \hat{\hat{N}}_t\\ \hat{\hat{n}}_{i,t}\\ 0\end{pmatrix} + (\hat{\theta}_t\ \hat{N}_t\ \hat{n}_{i,t}\ \hat{\omega}_{i,t})\bar{\mathcal{H}}_v\begin{pmatrix}\hat{\theta}_t\\ \hat{N}_t\\ \hat{n}_{i,t}\\ \hat{\omega}_{i,t}\end{pmatrix},$$

$\nabla \bar{v}_n$ and $\nabla \bar{v}$ are the gradients of $v_n(\cdot)$ and $v(\cdot)$ at $\delta \to 0$, and $\bar{\mathcal{H}}_v$ represents the Hessian of $v(\cdot)$ at $\delta \to 0$.

Lemma 9 *The second-order expansion of input $\hat{\hat{n}}_{i,t}$, $\hat{\hat{N}}_t$, and the first-order expansion of attention, $\hat{z}_{i,t}$ solve the following system:*

$$\sum_{\tau=t}^{\infty} \beta^{\tau-t} \mathbb{E}\Big[v_{i,\tau}^{(3)} \times \frac{1-\epsilon_{i,t}^2}{2\bar{z}} \Big| \bar{\mathcal{F}}'_{i,t-1}\Big] = 6\kappa \hat{z}_{i,t},$$

$$\mathbb{E}\big[\hat{\hat{v}}_{n,i,t} \big| \bar{\mathcal{F}}_{i,t}\big] + 2\frac{d}{d\delta}\mathbb{E}\big[\hat{v}_{n,i,t} \big| s_i^t(\delta)\big]\Big|_{\delta=0} = 0,$$

$$\hat{\hat{N}}_t = (1-\eta)\int \Big(\hat{n}_{i,t} - \int \hat{n}_{i,t}\Big)^2 + \int \hat{\hat{n}}_{i,t},$$

where

$$\hat{\hat{v}}_{n,i,t} = \nabla \bar{v}_n \begin{pmatrix} \hat{\hat{\vartheta}}_t \\ \hat{\hat{N}}_t \\ \hat{\hat{n}}_{i,t} \\ 0 \end{pmatrix} + (\hat{\theta}_t \;\; \hat{N}_t \;\; \hat{n}_{i,t} \;\; \hat{\omega}_{i,t}) \bar{\mathcal{H}}_{v_n} \begin{pmatrix} \hat{\theta}_t \\ \hat{N}_t \\ \hat{n}_{i,t} \\ \hat{\omega}_{i,t} \end{pmatrix}, \text{ and}$$

$$\begin{aligned} &\mathbb{E}\Big[v_{i,\tau}^{(3)} \frac{1-\epsilon_{i,t}^2}{2\bar{z}} \Big| \bar{\mathcal{F}}'_{i,t-1}\Big] \\ &= \mathbb{E}\Big[\big(3\hat{v}_{n,i,t}\hat{\hat{n}}_{i,\tau} + \hat{n}_{i,\tau}^2\big(3\bar{v}_{nn\theta}\hat{\theta}_\tau + 3\bar{v}_{nnN}\hat{N}_\tau + \bar{v}_{nnn}\hat{n}_{i,\tau}\big)\big)\frac{1-\epsilon_{i,t}^2}{2\bar{z}} \Big| \bar{\mathcal{F}}'_{i,t-1}\Big]. \end{aligned}$$

As in the static model, the first-order expansions of input $\hat{n}_{i,t}, \hat{N}_t$ are jointly determined with the zeroth-order expansion of attention $\bar{z}$, and similarly, the second-order expansions of input $\hat{\hat{n}}_{i,t}, \hat{\hat{N}}_t$ are jointly determined with the first-order expansion of attention $\hat{z}$.

Expansion of the Expectation Operator

In Lemma 9, the system involves an expansion with respect to the expectation operator $\frac{d}{d\delta}\mathbb{E}[\cdot|s_i^t(\delta)]\big|_{\delta=0}$. The following lemma shows how it can be calculated:

Lemma 10 *Given a generic random variable $\xi_t \in \sigma(\omega^t)$,*

$$\frac{d}{d\delta}\mathbb{E}[\xi_t|s_i^t(\delta)]\big|_{\delta=0} = \sum_{\tau=0}^{t} Cov\Big(\xi_t, \frac{\frac{d}{d\delta}\phi(s_{i,t-\tau}(\delta)|\omega^{t-\tau},\delta)}{\phi(s_{i,t-\tau}(\delta)|\omega^{t-\tau},\delta)}\Big|_{\delta=0} \Big| \bar{\mathcal{F}}_{i,t}\Big),$$

where $\phi(s_{i,t}(\delta)|\omega^t,\delta)$ denote the density of signals $s_{i,t}(\delta)$ conditional on ω^t in the economy indexed by δ.

Proof. Write the expectation as an integral over the probability density,

$$\frac{d}{d\delta}\mathbb{E}[\xi_t|s_i^t(\delta)] = \int \xi(\omega^t)\ \frac{d}{d\delta}\phi(\omega^t|s_i^t(\delta),\delta)d\omega^t. \tag{28}$$

Bayes rule implies

$$\phi(\omega^t|s_i^t(\delta),\delta) = \frac{\phi(s_{i,t}(\delta)|\omega^t,\delta)\phi(\omega^t|s_i^{t-1}(\delta),\delta)}{\int \phi(s_{i,t}(\delta)|\tilde{\omega}^t,\delta)\phi(\tilde{\omega}^t|s_i^{t-1}(\delta),\delta)d\tilde{\omega}^t}.$$

Differentiate both sides with respect to δ and divide by $\phi(\omega^t|s_i^t(\delta),\delta)$,

$$\begin{aligned}\frac{\frac{d}{d\delta}\phi(\omega^t|s_i^t(\delta),\delta)}{\phi(\omega^t|s_i^t(\delta),\delta)} &= \frac{\frac{d}{d\delta}\phi(s_{i,t}(\delta)|\omega^t,\delta)}{\phi(s_{i,t}|\omega^t,\delta)} - \mathbb{E}\Big[\frac{\frac{d}{d\delta}\phi(s_{i,t}(\delta)|\omega^t,\delta)}{\phi(s_{i,t}|\omega^t,\delta)}\Big|s_i^t(\delta)\Big] \\ &+ \frac{\frac{d}{d\delta}\phi(\omega^t|s_i^{t-1}(\delta),\delta)}{\phi(\omega^t|s_i^{t-1}(\delta),\delta)} - \mathbb{E}\Big[\frac{\frac{d}{d\delta}\phi(\omega^t|s_i^{t-1}(\delta),\delta)}{\phi(\omega^t|s_i^{t-1}(\delta),\delta)}\Big|s_i^t(\delta)\Big].\end{aligned} \tag{29}$$

Using Equation 29, Equation 28 can be written as

$$\frac{d}{d\delta}\mathbb{E}[\xi_t|s_i^t(\delta)] = Cov\Big(\xi_t, \frac{\frac{d}{d\delta}\phi(s_{i,t}(\delta)|\omega^t,\delta)}{\phi(s_{i,t}(\delta)|\omega^t,\delta)}\ \Big|s_i^t(\delta)\Big) + Cov\Big(\xi_t, \frac{\frac{d}{d\delta}\phi(\omega^t|s_i^{t-1}(\delta),\delta)}{\phi(\omega^t|s_i^{t-1}(\delta),\delta)}\ \Big|s_i^t(\delta)\Big).$$

Iterating backward,

$$\frac{d}{d\delta}\mathbb{E}[\xi_t|s_i^t(\delta)] = \sum_{\tau=0}^{t} Cov\Big(\xi_t, \frac{\frac{d}{d\delta}\phi(s_{i,t-\tau}(\delta)|\omega^t,\delta)}{\phi(s_{i,t-\tau}(\delta)|\omega^t,\delta)}\ \Big|s_i^t(\delta)\Big).$$

Finally, $\omega_{t-\tau+1}^t \perp s_{i,t-\tau}(\delta)\big|_{\omega^{t-\tau}}$ implies

$$\phi(s_{i,t-\tau}(\delta)|\omega^t,\delta) = \phi(s_{i,t-\tau}(\delta)|\omega^{t-\tau},\delta),$$

and the lemma follows from evaluating the expression at $\delta \to 0$. □

It is useful to clarify the expression in Lemma 10. Note that $\{s_{i,t-\tau}(\delta)\}_{\tau\geq 0}$ are signals in agent i's information set $\boldsymbol{\sigma}(s_i^t(\delta))$ when forming expectations, whereas ω^t is a running variable integrated over the density function. Therefore, given a path of realizations of shocks $\{\omega_\tau^*, \omega_{i,\tau}^*, \epsilon_{i,\tau}^*\}_{\tau=0}^t$,

$$s_{i,t}^*(\delta) = \begin{pmatrix} \theta_t^*(\delta) + \frac{\delta\epsilon_{i,t}^*}{\sqrt{z_{i,t}^*(\delta)}} \\ \theta_{t-1}^*(\delta) + \vartheta_{t-1}^*(\delta) + \eta\log N_t^*(\delta) + \delta\omega_{i,t-1}^* \end{pmatrix},$$

and

$$\frac{\frac{d}{d\delta}\phi(s^*_{i,t}(\delta)|\omega^t,\delta)}{\phi(s^*_{i,t}(\delta)|\omega^t,\delta)}\bigg|_{\delta=0} = \frac{-\bar{z}}{2}(x^*_t - \hat{\theta}_t(\omega^t))(\hat{\theta}^*_t - \hat{\theta}_t(\omega^t))\hat{z}^*_{i,t}$$
$$+ \frac{-1}{2\sigma^2_{\omega_i}}\left(p^*_{t-1} - \left(\hat{\theta}_{t-1}(\omega^{t-1}) + \eta\hat{N}_{t-1}(\omega^{t-1})\right)\right)$$
$$\times \left((\hat{\hat{\vartheta}}^*_{t-1} + \eta\hat{\hat{N}}^*_{t-1}) - (\hat{\hat{\vartheta}}_{t-1}(\omega^{t-1}) + \eta\hat{\hat{N}}_{t-1}(\omega^{t-1}))\right).$$

In this case, the covariance in Lemma 10 is $Cov(\cdot,\cdot|\hat{s}^{*t}_i)$, which is conditional on $\hat{s}^{*t}_i$ and integrating over ω^t.

Write the signal as

$$s_{i,t}(\delta) = H(g_t(\delta),\delta) + \begin{pmatrix} \frac{1}{\sqrt{z_{i,t}(\delta)}} & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \delta\epsilon_{i,t} \\ \delta\omega_{i,t-1} \end{pmatrix}, \quad g_t := \begin{pmatrix} f_t \\ f_{t-1} \end{pmatrix},$$

where $H(g_t(\delta),\delta)$ has expansions

$$\hat{H}_t = \phi_H^\intercal g_t, \quad \hat{\hat{H}}_t = \sum_{k=1}^{|s_{i,t}|} \hat{\hat{H}}^{(k)}_t \times e_k, \quad \hat{\hat{H}}^{(k)}_t = g_t^\intercal \Phi_{H_k} g_t.$$

Let

$$\mu^\tau_{i,t} := \mathbb{E}[g_\tau|\bar{\mathcal{F}}_{i,t}], \quad V^{t,\tau}_t := Cov[g_\tau, g_\tau^\intercal|\bar{\mathcal{F}}_{i,t}],$$
$$\Omega^{t,\tau,\tau}_{k,t} := 2V^{t,\tau}_t\Phi_{H_k}V^{\tau,\tau}_t + V^{t,\tau}_t tr[\Phi_{H_k}V^{\tau,\tau}_t],$$

and

$$\bar{Z} = \begin{pmatrix} \bar{z} & 0 \\ 0 & 1 \end{pmatrix}, \quad \hat{Z}_{i,t} = \begin{pmatrix} \bar{z}\times\hat{z}_{i,t} & 0 \\ 0 & 1 \end{pmatrix}.$$

Define $\hat{\mathcal{E}}_{i,t} := \frac{d}{d\delta}\mathbb{E}[g_t|s^t_i(\delta)]\big|_{\delta=0}$, then Lemma 10 implies

$$\hat{\mathcal{E}}_{i,t} = \frac{-1}{2}\sum_{\tau=0}^{t}\Bigg\{2V^{t,t-\tau}_t\phi_H\left(\hat{Z}_{i,t-\tau}(\phi_H^\intercal\mu^{t-\tau}_{i,t} - \hat{s}_{i,t-\tau}) - \bar{Z}\hat{\hat{s}}_{i,t-\tau}\right)\Bigg) \tag{30}$$
$$+ \sum_{k=1}^{|s_{i,t}|}\left(\Omega^{t,t-\tau}_{k,t} + V^{t,t-\tau}_t\mu^{t-\tau\intercal}_{i,t}\Phi_{H_k}\mu^{t-\tau}_{i,t} + 2V^{t,t-\tau}_t\Phi_{H_k}\mu^{t-\tau}_{i,t}(\phi_H\mu^{t-\tau}_{i,t} - \hat{s}_{i,t-\tau})^\intercal\right)\bar{Z}^{(k)}\Bigg\}.$$

D Computation

In this appendix, I detail the procedure to compute the first- and second-order expansions, which I use for the quantitative results in Section 6. A similar procedure applies to higher-order approximations.

The expansions of $\log N_t, \log n_{i,t}$ and $\log z_{i,t}$ are of the following forms:

$$\hat{N}_t = \boldsymbol{N}_\omega \omega^t, \quad \hat{n}_{i,t} = \boldsymbol{n}_s \hat{s}_i^t, \quad \bar{z}_{i,t} = \bar{\boldsymbol{z}},$$
$$\hat{\hat{N}}_t = \omega^{t\top} \boldsymbol{N}_{\omega\omega} \omega^t + \boldsymbol{N}_{\delta\delta}, \quad \hat{\hat{n}}_{i,t} = \hat{s}_i^{t\top} \boldsymbol{n}_{ss} \hat{s}_i^t + \boldsymbol{n}_{\delta\delta} + \boldsymbol{n}_s \hat{\hat{s}}_{i,t}, \quad \hat{z}_{i,t} = \boldsymbol{z}_s \hat{s}'^{t-1}_i,$$

where $\boldsymbol{N}_\delta, \boldsymbol{n}_\delta, \boldsymbol{N}_{\omega\delta}, \boldsymbol{n}_{s\delta}, \boldsymbol{z}_\delta$ are zeros and omitted for ease of exposition.

To compute the expansions, I use the following finite-dimensional approximation:

$$\tilde{N}_t^{(1)} = \Phi_f^N f_t^{(1)}, \quad \tilde{n}_{i,t}^{(1)} = \Phi_f^n f_{i,t}^{(1)}, \quad \tilde{z}^{(0)} = \Phi^z,$$
$$\tilde{N}_t^{(2)} = f_t^{(2)\top} \Phi_{ff}^N f_t^{(2)} + \Phi_{\delta\delta}^N, \quad \tilde{n}_{i,t}^{(2)} = f_{i,t}^{(2)\top} \Phi_{ff}^n f_{i,t}^{(2)} + \Phi_{\delta\delta}^n + \Phi_f^n f_{i,t}^{(1,1)}, \quad \tilde{z}_{i,t}^{(1)} = \Phi_f^z f'^{(2)}_{i,t-1},$$

together with $\tilde{\theta}^{(1)} = \Phi_f^\theta f_t^{(1)}$ and $\tilde{\vartheta}^{(2)} = f_t^{(2)\top} \Phi_{ff}^\vartheta f_t^{(2)}$ for the exogenous state.

Denote $\Phi^{(1)} := \{\Phi_f^\theta, \Phi_f^N, \Phi_f^n, \Phi^z\}$ and $\Phi^{(2)} := \{\Phi_{ff}^\vartheta, \Phi_{ff}^N, \Phi_{\delta\delta}^N, \Phi_{ff}^n, \Phi_{\delta\delta}^n, \Phi_f^z\}$. These are scalars, vectors, and matrices that correspond to the derivatives of the policy functions; $f_t^{(1)}, f_t^{(2)}, f_{i,t}^{(1)}, f_{i,t}^{(2)}, f'^{(2)}_{i,t}, f_{i,t}^{(1,1)}$ are factors that summarize histories of shocks and signals. I impose the aggregate factors with the following structure:

$$f_{t+1}^{(1)} = G^{(1)} f_t^{(1)} + \mathbf{1} \times \omega_t, \quad f_{t+1}^{(2)} = G^{(2)} f_t^{(2)} + \mathbf{1} \times \omega_t$$

for some maxtrices $G^{(1)}, G^{(2)}$ and $\mathbf{1}$ is a vector of ones. And I use the following structure for the individual factors:

$$f_{i,t}^{(1)} := \mathbb{E}[f_t^{(1)} | \tilde{s}_i^{(1),t}], \quad f_{i,t}^{(2)} := \mathbb{E}[f_t^{(2)} | \tilde{s}_i^{(1),t}], \quad f'^{(2)}_{i,t-1} := \mathbb{E}[f_t^{(2)} | \tilde{s}'^{(1),t}_i], \tag{31}$$

where $\tilde{s}_i^{(1),t}, \tilde{s}'^{(1),t}_i$ are the first-order expansion of signals given $\tilde{\theta}_t, \tilde{N}_t$, and $\tilde{z}^{(0)}$. This gives

$$f_{i,t+1}^{(1)} = A^{(1)} f_{i,t}^{(1)} + C^{(1)} \tilde{s}_{i,t}, \quad f_{i,t+1}^{(2)} = A^{(2)} f_{i,t}^{(2)} + C^{(2)} \tilde{s}_{i,t}$$

with matrices $A^{(1)}, C^{(1)}, A^{(2)}, C^{(2)}$ from the corresponding Kalman filter, and a similar structure for $f'^{(2)}_{i,t-1}$ and $f_{i,t+1}^{(1,1)}$. The structure in Equation 31 is simply a convenient form. In principle, one does not need to impose an a priori connection between the aggregate factors and individual factors.

Given the factor structure, conditions in Lemma 8 and 9 correspond to:

$$\mathbf{0} \approx \boldsymbol{\Gamma}^{(1)}(\{\Phi^{(1)}, f_\tau^{(1)}, f_{i,\tau}^{(1)}\}_{\tau \le t}), \quad \forall t = 0, \ldots, \infty,$$
$$\mathbf{0} \approx \boldsymbol{\Gamma}^{(2)}(\{\Phi^{(2)}, f_\tau^{(2)}, f_{i,\tau}^{(2)}, f_{i,\tau}^{(1,1)}, f'^{(2)}_{i,\tau}\}_{\tau \le t}; \{\Phi^{(1)}, f_\tau^{(1)}, f_{i,\tau}^{(1)}\}_{\tau \le t}), \quad \forall t = 0, \ldots, \infty,$$

where $\mathbf{\Gamma}^{(1)}(\cdot)$ and $\mathbf{\Gamma}^{(2)}(\cdot)$ are functions that represents the equilibrium conditions.

As an example, suppose that we have already solved the first-order $\mathbf{\Gamma}^{(1)}(\cdot)$ and would like to obtain the second-order approximations. The equilibrium conditions in Lemma 9 implies a system $\mathbf{\Gamma}^{(2)}(\cdot)$ as follows:

$$\begin{aligned}
residual_{1,t} &= f_{i,t}^{(2)\intercal}\Phi_{ff}^{n} f_{i,t}^{(2)} + \phi_f^{n\intercal} f_{i,t}^{(1,1)} - r f_{i,t}^{(2)\intercal}\Phi_{ff}^{\vartheta} f_{i,t}^{(2)} - s f_{i,t}^{(2)\intercal}\Phi_{ff}^{N} f_{i,t}^{(2)} \\
&\quad - (r\Phi_f^{\theta} + s\Phi_f^{N})^\intercal \hat{\mathcal{E}}_{i,t}^{f} - f_{i,t}^{(1)\intercal}\boldsymbol{\phi}^\intercal \bar{\mathcal{H}}_{v_n}\boldsymbol{\phi}\ f_{i,t}^{(1)} + const_1, \\
residual_{2,t} &= f_t^{(2)\intercal}\Phi_{ff}^{N} f_t^{(2)} - \int f_{i,t}^{(2)\intercal}\Phi_{ff}^{n} f_{i,t}^{(2)} + \Phi_f^{n\intercal} f_{i,t}^{(1,1)} di + const_2, \\
residual_{3,t} &= \sum_{\tau=t}^{\infty}\beta^{\tau-t}\Big(\big((\sigma_{\tau,t}^{n})^2\phi_{\mathcal{H}}^\intercal + 2\sigma_{\tau,t}^{n}\Phi_{ff}^{n}\big)G^{\tau-t} - 2(\sigma_{\tau,t}^{n})^2\Phi_f^{z}\Big) f_{i,t-1}^{\prime(2)},
\end{aligned}$$

where

$$\boldsymbol{\phi} := \begin{pmatrix}\Phi_f^{\theta} & \Phi_f^{N} & \Phi_f^{n}\end{pmatrix}^\intercal,\ \phi_{\mathcal{H}}^\intercal = \begin{pmatrix}\frac{\bar{v}_{nn\theta}}{\bar{v}_{nn}} & \frac{\bar{v}_{nnN}}{\bar{v}_{nn}} & \frac{\bar{v}_{nnn}}{\bar{v}_{nn}}\end{pmatrix}\boldsymbol{\phi},\ \sigma_{\tau,t}^{n} := \Phi_f^{n\intercal} Cov[f_{i,\tau}, \epsilon_{i,t}],$$

and $\hat{\mathcal{E}}_{i,t}^{f} := \begin{pmatrix}I & 0\end{pmatrix}\hat{\mathcal{E}}_{i,t}$ are the first $|f_t|$ elements of $\hat{\mathcal{E}}_{i,t}$, given by Equation 30. Given the solution from the first-order approximation, the computational goal is to solve for $\Phi^{(2)}$ and $G^{(2)}$ that minimize the residuals.

Computational Procedure

Given the first $(m-1)^{th}$ order of approximations, $\{\Phi^{(k)}, f_\tau^{(k)}, f_{i,\tau}^{(k-l,l)}\}_{\tau\leq t, l\leq k}$, $\forall k \in \{1,\ldots,m-1\}$, the m^{th}-order approximation is solved successively with the following procedure:

(1) Fix a dimension d for the factors, and specify matrices $G^{(m)} \in \mathbb{R}^{d\times d}$. Simulate $\{\omega_t, \omega_{i,t}, \epsilon_{i,t}\}_{t\leq T}$ for some large T, and construct $\{f_\tau^{(m)}, f_{i,\tau}^{(m-l,l)}\}_{\tau\leq t, l\leq m}$.

(2) For any given $G^{(m)}$ and the associated factors $\{f_\tau^{(m)}, f_{i,\tau}^{(m-l,l)}\}_{\tau\leq t,l\leq m}$, solve for coefficients $\Phi^{(m)}$ that minimize the sum of the squared residuals:

$$\begin{aligned}
&R_j(G^{(m)}) := \\
&\quad \min_{\Phi^{(m)}}\sum_{t=0}^{T}\Big\{\mathbf{\Gamma}_j^{(m)}(\{\Phi^{(m)}, f_\tau^{(m)}, f_\tau^{(m-l,l)}\}_{\tau\leq t,l\leq m}; \{\Phi^{(k)}, f_\tau^{(k)}, f_\tau^{(k-l,l)}\}_{\tau\leq t,l\leq k})\Big\}^2,
\end{aligned}$$

where $j \in \{1,2,3\}$ corresponds to the residuals from the three equilibrium conditions.

(3) Solve for $G^{(m)}$ that minimize $R(G^{(m)}) = w_n R_1(G^{(m)}) + w_N R_2(G^{(m)}) + w_z R_3(G^{(m)})$, given weights (w_n, w_N, w_z).

Note that for $m > 1$, the minimization in (2) is generally a linear-quadratic problem, as $\Gamma^{(m)}$ is linear in $\Phi^{(m)}$. As a result, the problem can be solved efficiently. Moreover, the optimization over matrix $G^{(m)}$ can be restricted to matrices of Jordan canonical form without loss of generality. This procedure seeks the optimal d-factor representation of the equilibrium given the structure proposed, and one could successively increase the dimension d to reduce the size of residuals until it meets a certain computational criterion.

Implementation and Validation

The quantitative results are based on the following specifications:

- Period of simulation: 1500, discarding the first 50 periods.
- Number of factors: I use two factors for the first-order expansion, one of which is (proportional to) θ_t. I include an additional two factors for the second-order expansion, one of which, together with θ_t, minimizes the residual for $\hat{\hat{\vartheta}}_t$.
- The summation in Equation 30 is truncated at 10 periods.
- $G^{(1)}$, $G^{(2)}$ are restricted to be diagonal.
- Residuals weights $(w_n, w_N, w_z) = (1, 1, 3.5 \times 10^3)$. The weights are chosen so that the three residuals are of similar magnitudes. Residuals from attention optimality result from the third-order expansion. For numerical performance, they are scaled by the weights so that residuals from the three equilibrium conditions have similar magnitudes.
- Error tolerance: For the first- and second-order expansions, the size of residuals (Euclidean norm) per period is at the magnitude of 10^{-5}. The size can be interpreted as the errors in input decisions each period relative to the steady-state input.

Validation: When $\sigma_{\omega_i} \to \infty$, the first-order expansion of the model can be solved with the analytical solution provided by Huo and Pedroni (2020). The first-order expansion from the computation procedure described in this section produces a numerically identical result for this special case.

E Empirical Appendix

Data Description

The empirical evidence provided in Section 4 comes from two data sources.

1. For aggregate data, I use the quarterly series on output, hours, and TFP from Fernald's website.

2. The Survey of Professional Forecasters (SPF) is available on the Federal Reserve Bank of Philadelphia's website. The survey was formerly conducted by the American Statistical Association and the National Bureau of Economic Research, began in 1968:Q4, and was taken over by the Philadelphia Fed in 1990:Q2. The Philadelphia Fed conducts quarterly surveys with around 40 forecasters around the end of the second month in a quarter. It provides forecaster-level data, in which forecasters report forecasts for outcomes in the current and next four quarters, typically about the level of economic variables in each quarter. The outcomes predicted include a range of aggregate variables, including the real GDP forecasts. The Philadelphia Fed also provides the realized values of the forecasted aggregate variables, including all vintages of real GDP. In addition to asking forecasters for point estimates of these variables, the SPF also asks forecasters to report probabilistic forecasts for fixed-event year-over-year percentage changes in GDP growth. The SPF provides intervals of possible GDP growth and asks respondents to report their subjective probabilities that the variable of interest will take a value in each interval.

E.1 Measures of Information Rigidity

For the measure of information rigidity from the regression in Equation 6, the variables are constructed as follows:

1. To construct $\overline{FE}_{t,h} \coloneqq \Delta\widetilde{Y}_{t,h} - \overline{\mathbb{E}}_t[\Delta\widetilde{Y}_{t,h}]$ and $\overline{FR}_{t,h} \coloneqq \overline{\mathbb{E}}_t[\widetilde{Y}_{t,h}] - \overline{\mathbb{E}}_{t-1}[\widetilde{Y}_{t,h}]$:

   - $\Delta\widetilde{Y}_{t,h}$ is calculated as the growth of quarterly real GDP in period $t+h$ relative to that in the period $t-1$, using real GDP series from vintage $t+h$.

   - $\overline{\mathbb{E}}_t[\Delta\widetilde{Y}_{t,h}]$ is calculated as the SPF forecasts of real GDP for period $t+h$ relative to forecasts of GDP for period $t-1$ from surveys reported in period t, averaged across forecasters.

   Finally, I demean $\overline{FE}_{t,h}$ and $\overline{FR}_{t,h}$ by the averages of corresponding horizons to remove horizon fixed effects.

2. For the indicators of low economic activities, $\mathbf{1}_t^R$, I consider the following two specifications:

   (i) NBER recession, available at NBER's website.

(ii) Below trend: I construct a detrended output series $\{Y_t\}$ using a band-pass filter at 6-32 quarters frequency, and define $\mathbf{1}_t^R := \mathbf{1}_{Y_{t-1}<median(\{Y_t\})}$ as an indicator of whether output last period is below trend.

Besides the two specifications shown in Table 1, I consider a few additional specifications where (i) the cutoff for low output periods are, respectively, the 33% and 20% quantile of the detrended output series, $\{Y_t\}$, and (ii) the sample period is extended to include the Covid recession.

Table 6 reports the baseline specifications (columns "recession" and "50%", rows "Pre-Covid") and the alternative specifications. Although the moments vary among the specifications, all specifications show a decrease in the measure of information rigidity in low output periods. I use the pre-Covid sample as the baseline specification to avoid the result being driven by extreme periods at the end of the sample. Additionally, estimates from the pre-Covid sample are also consistent with the evidence from individual-level regression in Table 7 as I discuss below.

Table 6: Alternative Specifications of Coibion and Gorodnichenko (2015)

		Indicator ($\mathbf{1}_t^R$)			
		recession	50%	33%	20%
Pre-Covid	β_{CG}	0.56 (0.17)	0.73 (0.20)	0.60 (0.16)	0.61 (0.16)
	$\Delta\beta_{CG}$	−0.57 (0.32)	−0.24 (0.26)	−0.20 (0.24)	−0.20 (0.25)
Incl. Covid	β_{CG}	0.49 (0.16)	0.72 (0.20)	0.50 (0.16)	0.53 (0.16)
	$\Delta\beta_{CG}$	−0.71 (0.24)	−0.79 (0.24)	−0.63 (0.22)	−0.67 (0.21)
Model	β_{CG}	–	0.74	0.69	0.65
	$\Delta\beta_{CG}$	–	−0.24	−0.23	−0.24

Forecasts horizons: 0 to 3 quarters ahead; robust standard errors in parentheses. Pre-Covid: 1968 Q3 - 2019 Q4; Incl. Covid: 1968 Q3 - 2022 Q4. Model moments are averages of 1000 simulations of 200 quarters with 50 forecasters.

Individual-Level Regression

I consider a variation of the regression in Goldstein (2023) to provide supporting evidence that forecasters pay more attention when they expect worsening aggregate conditions. The regression relies on the persistence in deviation of individual forecast from the average forecast:

$$FDEV_{i,t,h} \coloneqq \mathbb{E}_{i,t}[\tilde{Y}_{t+h}] - \bar{\mathbb{E}}_t[\tilde{Y}_{t+h}].$$

Intuitively, how much an individual's forecast deviation persists across a certain period is closely connected to the amount of information received. This logic lies behind the regression in Goldstein (2023), projecting $FDEV_{i,t,h}$ and the deviation last period, $FDEV_{i,t-1,h+1}$.

To capture the possibility that forecasters may vary the amount of information they acquire depending on what they expect about the aggregate state of the economy, I allow for an interaction term with an indicator $\mathbf{1}^R_{i,t}$

$$FDEV_{i,t,h} = \alpha_G^\intercal \begin{pmatrix} 1 \\ \mathbf{1}^R_{i,t} \end{pmatrix} + \begin{pmatrix} \beta_G & \Delta\beta_G \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{1}^R_{i,t} \end{pmatrix} FDEV_{i,t-1,h+1} + residual_{i,t,h}., \tag{32}$$

where

$$\mathbf{1}^R_{i,t} \coloneqq \mathbf{1}_{\{FDEV_{i,t-1,h+1} < cut_{i,h}\}},$$

and cutoff $cut_{i,h}$ is given by some $x\%$ quantile of the individual's deviations for horizon h for $x \in \{50, 33, 20\}$. A negative coefficient $\Delta\beta_G$ is indicative that forecasters update their expectations more when they expect worsening economic conditions.

Table 7: State-Dependent Attention: Individual Level Regression

		Indicator ($\mathbf{1}^R_{i,t}$)		
		50%	33%	20%
Data	β_G	0.50 (0.02)	0.50 (0.02)	0.49 (0.02)
	$\Delta\beta_G$	−0.14 (0.04)	−0.15 (0.04)	−0.16 (0.04)
Model	β_G	0.61	0.60	0.59
	$\Delta\beta_G$	−0.15	−0.16	−0.15

Forecasts horizons: 0 to 3 quarters ahead; robust standard errors in parentheses. Sample period: 1968 Q3 - 2019 Q4; extending to 2022 Q4 gives identical results up to rounding errors.

E.2 Uncertainty, Volatility, and Forecast Dispersion

Measures reported in Table 2 are constructed as follows.

1. Aggregate volatility, σ_t^Y, is measured as the conditional heteroskedasticity of quarterly real GDP growth with a univariate EGARCH(1,1)-ARMA(1,1) model. I follow the same procedure to construct the volatility of TFP, σ_t^{TFP}, which is similar to the estimation in Bloom et al. (2018).

2. Forecast dispersion about aggregate output, d_t^Y, is calculated from the SPF point estimates. For each period t and forecaster i, I calculate the forecasts of real GDP growth $\mathbb{E}_{i,t}[\Delta\widetilde{Y}_t]$ as the forecasts of real GDP in period t relative to that in period t from the survey reported in period t. The forecast dispersion, d_t^Y, is calculated as the standard deviation of $\mathbb{E}_{i,t}[\Delta\widetilde{Y}_t]$ for each period across forecasters. Finally, to alleviate changes due to survey design when the Philadelphia Fed took over the SPF, I removed a version fixed effect for periods before 1992Q1 (after which the Philadelphia Fed started reporting under the new survey design).

3. Subjective uncertainty about aggregate output, υ_t^Y, is calculated using the SPF probability-range data. I fit a Beta distribution with parameter a, b and support $[l, r]$ to the response of each forecaster in each period. Specifically, let $\{m_k\}_{k=1}^n$ denote the endpoints of intervals specified by the SPF, where $m_1 = -\infty$ and $m_n = \infty$, and $F_{i,t}(m_l)$ denote the empirical CDF provided by forecaster i. I look for parameters $a_{i,t}, b_{i,t}$ and bounds $l_{i,t}, r_{i,t}$ that solve

$$\min_{a_{i,t}>1, b_{i,t}>1, l_{i,t}, r_{i,t}} \sum_{k=1}^{n} \Big(Beta(t_k, a_{i,t}, b_{i,t}, l_{i,t}, r_{i,t}) - F_{i,t}(m_{i,t}) \Big)^2$$

   such that

$$\begin{cases} l_{i,t} = \inf\ Supp(F_{i,t}), & \text{if } \inf\ Supp(F_{i,t}) > -\infty \\ l_{i,t} > l_{min}, & \text{if } \inf\ Supp(F_{i,t}) = -\infty, \\ r_{i,t} = \sup\ Supp(F_{i,t}), & \text{if } \sup\ Supp(F_{i,t}) < \infty \\ r_{i,t} < r_{max}, & \text{if } \sup\ Supp(F_{i,t}) = \infty, \end{cases}$$

   where l_{min} and r_{max} are bounds on the support. In other words, if a forecaster places a positive probability on the unbounded intervals provided by the SPF, I estimate finite bounds $l_{i,t}$ and $r_{i,t}$ with limit l_{min} and r_{max}; otherwise, I take the support provided by the forecaster as the support for the Beta distribution. I set (l_{min}, r_{max}) to be $(-16\%, 16\%)$. Given $a_{i,t}, b_{i,t}, l_{i,t}, r_{i,t}$, I calculate the standard deviation of the fitted distribution for each forecaster i in period t. I subtract the standard deviation of a uniform distribution over the minimal bin size of 1% so that the measure is zero when a forecaster places 100% probability in one bin. I aggregate the measures across forecasters by computing the averages of standard deviations across forecasters for each period.

Finally, the design of the probability range survey introduces a few issues:

(i) The survey asks forecasters to report probability forecasts for year-over-year GDP growth in different quarters throughout the year. That is, for a period t, the forecasters report their beliefs on

$$\sum_{\tau \in yr(t)} Y_\tau / \sum_{\tau \in yr(t)-1} Y_\tau,$$

where $yr(t)$ denote the year in which period t is in. To make the forecasts comparable to the other two series, which are based on quarterly output growth, Y_t/Y_{t-1}. I adjust the series by multiplying $\sum_{\tau \in yr(t)-1} Y_\tau$ and dividing by Y_{t-1} and by the number of quarters remaining in the year. The adjusted series represents the subjective uncertainty about quarterized real GDP growth for the remainder of the year. I remove quarter-of-the-year fixed effects to control for differences due to forecast horizons.

(ii) The upper and lower bounds of the survey occasionally introduce bunching at the top and bottom cells of the survey. To address this issue, I control for an indicator of whether more than 20% of the forecasters put 50% of the probability in the top or bottom cell.

(iii) To control for changes due to survey design in the 1990s, I remove a version fixed effect for periods before 1992Q1, similar to the adjustment for forecast dispersion discussed above.

Figure 1 shows the measures of aggregate volatility, forecast dispersion, and subjective uncertainty in log deviations from their respective long-run averages, where the NBER recession periods are marked by gray areas. All three measures are countercyclical, rising sharply during recessions and declining during booms.

Table 8 shows the same moments as Table 2 using the full sample data, including the Covid recession. In comparison to Table 2, the measures are more negatively correlated to output, and the magnitudes of fluctuations are larger. I use moments from the Pre-Covid sample in Table 2 as the baseline to avoid the quantitative results being driven by extreme periods at the end of the sample.

Table 8: Uncertainty, Volatility, and Forecast Dispersion

	σ_t^Y	d_t^Y	υ_t^Y	σ_t^{TFP}
$cor(\cdot, \tilde{Y}_t)$	$-.44$	$-.47$	$-.38$	$-.36$
sd/avg	.68	.47	.46	.20

Sample: 1968Q3 to 2022Q4, detrended with a band-pass filter at 6-32 quarters frequency; υ_t^Y available only after 1981Q2.

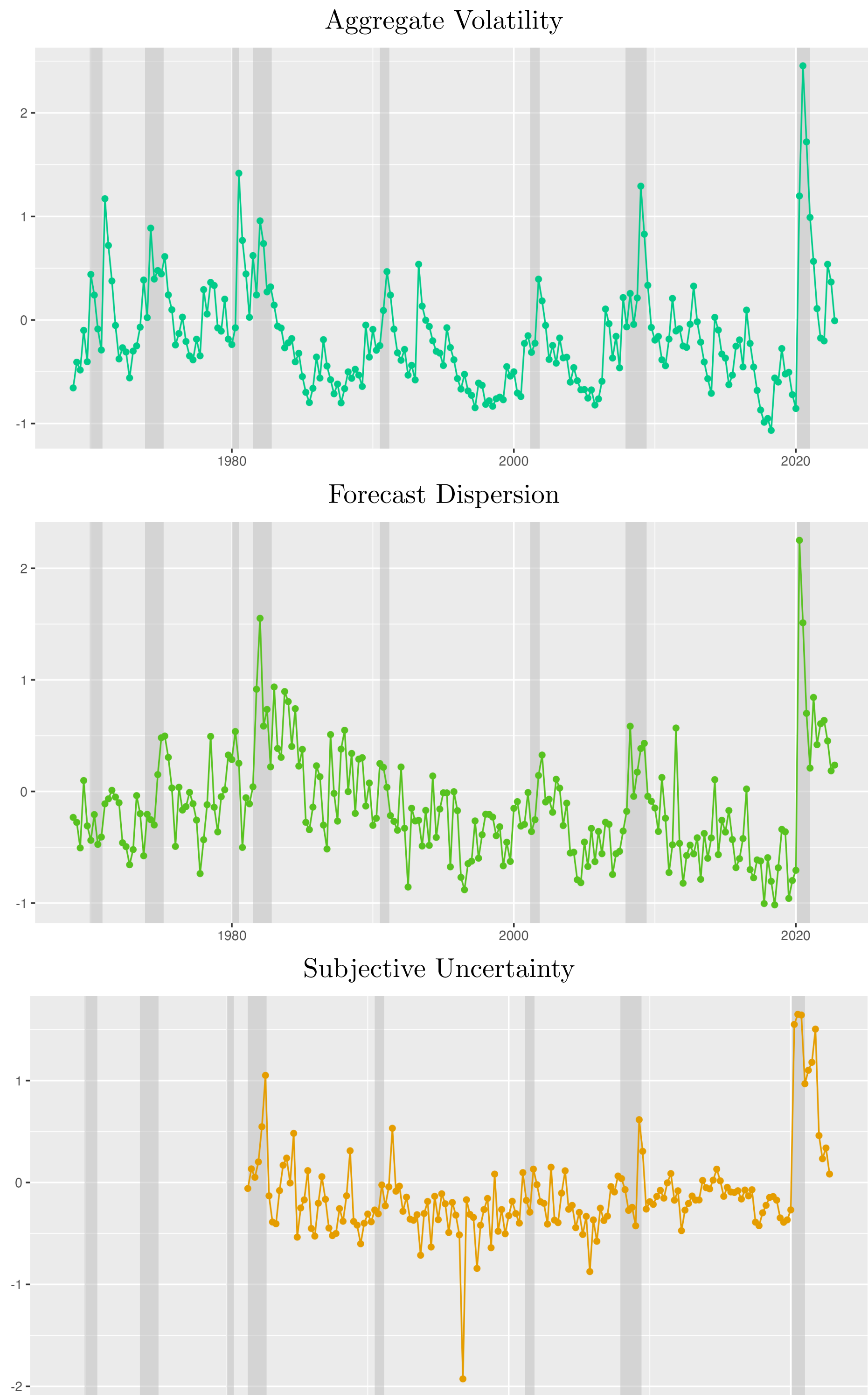

Figure 1: Aggregate volatility, forecast dispersion, and subjective uncertainty about aggregate output over time. X-axis: quarters; Y-axis: log deviation of variables from long-run averages. Gray areas indicate NBER recessions.