ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS
WORKING PAPER SERIES

Attention and Fluctuations in Macroeconomic Uncertainty

Authors	Yu-Ting Chiang
Working Paper Number	2022-004A
Creation Date	March 2022
Citable Link	https://doi.org/10.20955/wp.2022.004
Suggested Citation	Chiang, Y.-T., 2022; Attention and Fluctuations in Macroeconomic 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

The views expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment.

Attention and Fluctuations in Macroeconomic Uncertainty

Yu-Ting Chiang[†]

March 8, 2022

Abstract

This paper studies a dispersed information economy in which agents can exert costly attention to learn about an unknown aggregate state of the economy. Under certain conditions, attention and four measures of uncertainty are countercyclical: Agents pay more attention when they expect the economy to be in a bad state, and their reaction generates higher (i) aggregate output volatility, (ii) cross-sectional output dispersion, (iii) forecast dispersion about aggregate output, and (iv) subjective uncertainty about aggregate output faced by each agent. All these phenomena are prominent features of the U.S. data. When attention cost is calibrated to forecast survey data, the model generates countercyclical fluctuations in attention and uncertainty, consistent with untargted moments from the data. Fluctuations in attention and uncertainty are higher-order properties of the model. A new method is developed to solve higher-order dynamics of the equilibrium under an infinite regress problem.

Keywords: business cycles, macroeconomic uncertainty, dispersed information, rational inattention

JEL code: D8, E1, E3, E7

[†]Federal Reserve Bank of St. Louis, contact email: yu-ting.chiang@stls.frb.org. This paper is a revision of the first chapter of my dissertation. I thank my advisors Mikhail Golosov and Lars Hansen for their guidance and support and my thesis committee members, Fernando Alvarez and Yueran Ma, for their helpful advice. I appreciate valuable discussions with Kai-Hao Yang, Piotr Źoch, Mikayel Sukiasyan, Ezra Karger, and Hyunmin Park. All errors are my own. Views in this paper do not necessarily represent those of the Federal Reserve Bank of St. Louis or the Federal Reserve System.

1. Introduction

This paper studies a dispersed information economy in which economic agents can exert costly attention to learn about the state of the economy over the business cycle. In equilibrium, countercyclical attention generates countercyclical fluctuations in macroeconomic uncertainty: Agents pay more attention to the economy when they expect it to enter a downturn, and their reaction generates (i) an increase in aggregate output volatility, (ii) an increase in cross-sectional output dispersion, (iii) higher forecast dispersion about aggregate output, and (iv) more subjective uncertainty about aggregate output for each agent in the economy. Countercyclical fluctuations in attention and the four measures of uncertainty are prominent business cycle phenomena in the data. The economy studied below exhibits these phenomena without reliance on exogenous volatility shocks: All exogenous shocks that hit the economy are homoskedastic, yet endogenous variables, such as input and output, are heteroskedastic due to agents' endogenous attention responses.

The economy I study is a standard dispersed information economy with one key departure: agents' attention to the economy varies endogenously with economic conditions. In the economy, agents observe unknown aggregate productivity with idiosyncratic noise and make production decisions based on their information. Agents can "pay attention" in the sense that they can exert costly effort to reduce the noise in their observations. When agents expect the economy to be in a bad state, they pay more attention to avoid bad decisions because resources become more valuable when expected income is low. As agents pay attention, they react strongly to the information they receive and, under certain conditions, their *endogenous reaction* increases the four measures of macroeconomic uncertainty: Agents' production decisions respond strongly to realizations of aggregate productivity, and their aggregate response increases aggregate volatility. At the same time, agents' expectations and reactions diverge as they react to different information. Moreover, because information is dispersed, agents are uncertain about others' aggregate responses; as agents pay attention and react, each agent faces more uncertainty about the aggregate outcome due to their uncertainty about other agents' endogenous responses, despite having learned more about the exogenous state of the economy.

I characterize the aforementioned results analytically in a static economy where the mechanism is transparent. I then extend the economy to a dynamic setup to explore the quantitative potential of the mechanism. Key parameters of the model are calibrated to match features of the U.S. economy: The model matches the average size of forecast errors from forecast surveys, as well as salient business cycle moments such as the average volatility and persistence of aggregate input and output. These features of the data jointly pin down two key features of economy: the average level of attention and the level of strategic complementarity in the economy, which effectively determine the strength of the mechanism. With the calibrated model, I simulate and construct untargeted statistics for measures of uncertainty and attention, and compare them to corresponding statistics for the U.S. economy.

In the calibrated model, endogenous attention fluctuations generate countercyclical fluctuations in all four measures of uncertainty with cyclical, magnitudes, and persistence consistent with their respective counterparts from the data. Moreover, agents' forecast patterns in the model are consistent with empirical evidence documented by Coibion and Gorodnichenko (2015) that indicates countercyclical attention.

The analysis carried out in this paper (especially the quantitative results) is based on a new approximation method that may be of independent interest. The method builds on perturbation techniques to solve higher-order approximation of the equilibrium for dispersed information economies. The method departs from standard methods to address a few issues that arise from dispersed information and endogenous attention fluctuations, including an *infinite regress problem* under which the models lack a finite recursive state space. Recent progress by Huo and Takayama (2015) and others have found ways to handle the problem under first-order approximations. Yet, first-order approximations miss important features of the approximated economies: Among other well-known limitations, they cannot capture the fluctuations in attention and measures of uncertainty, because these fluctuations are higher-order properties of the model. These higher-order dynamics are captured by the method developed in this paper.

Literature

This paper is related to several strands of the literature.

The phenomena I seek to explain is the focus of an extensive literature that studies countercyclical fluctuations in macroeconomic uncertainty. Since Bloom (2009), much work has provided empirical evidence on countercyclical uncertainty and studied its effect.¹ Among works that provide explanations for fluctuations in uncertainty, Van Nieuwerburgh and Veldkamp (2006), Fajgelbaum et al. (2017), and Benhabib et al. (2016) feature procyclical information about aggregate or idiosyncratic productivity, Orlik and Veldkamp (2014) show that learning about the skewness of the GDP process can generate countercyclical subjective uncertainty, Ilut et al. (2018) use a concave hiring rule to generate time-varying responsiveness to shocks, which give rise to countercyclical aggregate volatility and cross-sectional dispersion of employment. Kozeniasukas, Orlik, and Veldkamp (2018) generate countercyclical fluctuations in the four measures of uncertainty with countercyclical shocks to the volatility of aggregate TFP. While these papers either focus on a few of the four phenomena or rely on exogenous volatility shocks, I show that all four phenomena related macroeconomic uncertainty are interconnected through agents' endogenous attention choice, which generates time-varying responsiveness along with forecast patterns consistent with evidence on countercyclical information acquisition documented by Coibion and

¹See, for example, Bachmann et al. (2013), Vavra (2014), Jurado et al. (2015), Ilut et al. (2018), Bloom et al. (2018), Fernández-Villaverde et al. (2011), Christiano et al. (2014), Gilchrist et al. (2014), Vavra (2014), and Basu and Bundick (2017).

Gorodnichenko (2015).

The model studied below builds on the dispersed information and rational inattention literature, following Phelps (1970), Lucas (1972), and Sims (2003). The modeling of attention choice in this paper is closet to Myatt and Wallace (2012). Dynamic models in the literature mostly feature static information structure: Woodford (2001), Lorenzoni (2009, 2010), Angeletos and La'O (2010, 2013), Angeletos and Lian (2018), and Angeletos and Huo (2021) feature exogenously static information structure; the information structure in Maćkowiak and Wiederholt (2015, 2009) is endogenous but static as a result of the approximation adopted. Among papers with time varying information structures, Nimark (2014) generates fluctuations in aggregate volatility and forecast dispersion with TFP volatility shocks. The arrivals of volatility shocks are accompanied by the availability of public signals, which capture the role of news media in coordinating agents' responses. Flynn and Sastry (2021) study attention choice with a flexible information structure, and show that attention fluctuations produce fluctuations in aggregate output volatility. In this paper, endogenous attention choice produces a time-varying information structure that turns homoskedastic shocks into heteroskedastic responses. Due to dispersed information, agents' heteroskedastic responses generate a broad set of phenomena in the microdata, including countercyclical dispersion in individual outputs and forecasts, as well as countercyclical subjective uncertainty.

Finally, the method developed in this paper is related to works that use perturbation methods to address economic problems,² and works that address the infinite regress problem in linear dispersed information models since Townsend (1983), and followed by Kasa (2000), Lorenzoni (2009), Nimark (2017), Huo and Takayama (2015), Huo and Pedroni (2020), and Angeletos and Huo (2021) among others. The method departs from standard perturbations on two fronts: (i) a new formulation of the perturbation sequence to address issues that arise from endogenous information choice, and (ii) a computation procedure to solve the perturbation sequence under the complication of the infinite regress problem. The perturbation formulation results in a higher-order approximation that generalizes the linear-quadratic approximation commonly used to solve dispersed information and rational inattention models. The computation procedure builds on insights of methods that address the infinite regress problem in linear models: despite each agents needing to track the whole history of signals to construct a belief hierarchy, the relevant equilibrium condition often depends only on a few state variables that appropriately summarize the history of signals. This paper provides one way to construct these state variables for higher-order approximation with an algorithm that resembles methods that deal with cross-sectional heterogeneity, such as the one provided by Krusell and Smith (1998).

²The method here is more closely related to the small shock expansion in Lombardo and Uhlig (2018), robust preference expansion in Borovicka and Hansen (2013), and perturbation with cross-sectional heterogeneity in Bhandari et al. (2018). See Judd (1998) for a classical exposition.

Outline

The rest of this paper is organized as follows. Section 2 sets up a static dispersed information economy with endogenous information choice. Section 3 shows that a decrease in expected productivity can generate increases in the four measures of uncertainty through an endogenous increase in attention. Section 4 discusses empirical evidence on the four measures of uncertainty and attention. Section 5 extends the economy to a dynamic setup, discusses the challenge posed by the infinite regress problem, and gives an overview of the computational method. Section 6 calibrates the model. Section 7 compares business cycle statistics of the four measures of uncertainty and attention between the model and the data. Section 8 concludes. Proofs and details of the computational method are provided in Appendix A and B. Extensions and empirical details are provided in Supplementary Appendix C and D.

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. All agents and the final good producer are price-takers. Aggregate productivity of the final good producer is unknown to the agents when they make labor input decisions, but agents can pay attention and acquire information about the aggregate productivity. The economy proceeds in three stages. Agents make their attention choices in stage 1. Labor input and intermediate goods production are set in stage 2. Final good production and consumption occur in stage 2.

Preferences and technologies

Preferences of an agent i are given by

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

where $c_i \in \mathbb{R}^+$ is final good consumption, $n_i \in \mathbb{R}^+$ is labor, and $z_i \in \mathbb{R}^+$ is agent i 's attention level. Utility from c_i and n_i takes a modified Greenwood–Hercowitz–Huffman (GHH) form:

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

where a lower bound $\underline{u} > 0$ ensures that preferences are well-defined for all realizations of consumption after any labor input choice. Parameter ν is the inverse Frisch elasticity of labor supply. Parameter γ plays a dual role: Besides its traditional role of parameterizing relative risk aversion over realizations of consumption-labor bundles, it also governs how agents trade off between attention level z_i and real variables c_i, n_i (in a sense to be made

clear below).

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 constraint $c_i \leq p_i q_i$, where p_i is the relative price of intermediate good i using the final good as the numeraire.

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

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

where θ is the aggregate productivity of the final good producer, and $\eta \in [0, 1)$ is the inverse elasticity of substitution. Aggregate productivity is stochastic and normally distributed with mean $\bar{\theta}$ and variance σ^2 :

$$\theta \sim \mathcal{N}(\bar{\theta}, \sigma^2).$$

The key comparative static exercise in this model studies how the economy responds to changes in $\bar{\theta}$, which captures “fluctuations” in the aggregate economic condition. The intuition of this exercise carries through to the dynamic model in Section 5, where the economy responds to changes in expected productivity as a result of productivity shocks that hit the economy in the past.

Timing and information

The economy proceeds in three stages. In stage 1, agents simultaneously choose attention z_i based on common prior $\theta \sim \mathcal{N}(\bar{\theta}, \sigma^2)$. In stage 2, each agent receives an idiosyncratic signal x_i about aggregate productivity with precision equal to their attention level z_i :

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

Based on signal x_i , agent i forms an expectation about the price of their intermediate good p_i , and chooses labor input n_i to produce the intermediate good. In stage 3, aggregate productivity θ and prices $\{p_i\}$ realize. 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.

Equilibrium definition

An (symmetric) equilibrium is a collection of real number z and functions $\{n, q, p, c, y, Y\}$, where $n_i = n(x_i)$, $q_i = q(x_i)$, $p_i = p(x_i, \theta)$, $c_i = c(x_i, \theta)$, $y_i = y(x_i, \theta)$, $Y = Y(\theta)$, such that (i) z_i optimizes expected utility under agents’ priors over x_i , n_i , and p_i ; (ii) n_i optimizes

expected utility under each agent's conditional expectation over p_i given x_i ; (iii) c_i is optimal, subject to budget constraints given p_i ; (iv) the final good producer chooses $\{y_i\}$ to maximize profit given the distribution of prices $\{p_i\}$; (v) intermediate goods production $\{q_i\}$ and final good production Y are given by their respective technology; and (vi) prices $\{p_i\}$ clear all markets: $y_i = q_i, \forall i$ and $Y = \int c_i di$.

Equilibrium characterization

I now characterize the equilibrium. Consider first the final good producer's profit maximization problem in stage 3. The final good producer takes prices $\{p_i\}$ as given and chooses $\{y_i\}$ to maximize profit. Their profit-maximization problem leads to the standard CES demand for intermediate goods: $p_i = e^{(1-\eta)\theta} Y^\eta y_i^{-\eta}$. Given labor input $\{n_i\}$ chosen by the intermediate goods producers in state 2, market clearing and production feasibility imply the equilibrium price of intermediate good i can be solved as a function of θ , n_i , and "aggregate labor" N :

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

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

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

The price of intermediate good p_i (relative to the final good) increases in θ because more of the final good is produced when productivity is high; it increases with N because intermediate goods are complements and good i is more valuable when other agents produce more; it decreases with n_i because the marginal value of an intermediate good is decreasing. Agents' consumption c_i increases with n_i despite a decrease in p_i , because the elasticity of substitution between intermediate goods is greater than 1.

Consider now the labor supply decisions that take place in stage 2. Given signal realization x_i , each agent forms expectations about price p_i based on their expectations about aggregate productivity θ and other agents' aggregate labor decisions N . Labor optimality implies

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

where $\mathbb{E}_i[\cdot] := \mathbb{E}[\cdot | x_i, z_i, \bar{\theta}]$ is a shortcut for the rational expectation conditional on x_i given attention z_i and expected productivity $\bar{\theta}$. The condition, in words, says that agents equalize the expected marginal product of labor (weighted by the normalized marginal utility of consumption) to their marginal disutility of labor.

Finally, consider the attention choice problem agents face in stage 1. Given other agents'

equilibrium attention z , each agent chooses attention z_i to equalize the marginal value of attention to the marginal cost of attention κ :

$$\frac{\partial}{\partial z_i} \mathbb{E}[u(c_i, n_i) \mid z_i, \bar{\theta}] = \kappa,$$

where each agent chooses attention z_i , understanding its effect on signal x_i and their labor input $n_i = n(x_i)$, as well as their consumption $c_i = e^\theta N^\eta n_i^{1-\eta}$, which also depends on other agents' attention z through aggregate labor N . Labor input function $n(\cdot)$ is evaluated at the optimal attention level z_i as a result of the envelope theorem.

Lemma 1 *An equilibrium is characterized by attention z , individual labor input $n(x_i)$, and aggregate labor $N(\theta)$ that solve the following fixed-point problem:*

$$\begin{aligned} \int u(c(x_i, \theta), n(x_i)) \frac{\partial}{\partial z} \varphi(x_i, \theta \mid z, \bar{\theta}) dx_i d\theta &= \kappa, \\ \mathbb{E}\left[(1 - \eta)e^\theta N(\theta)^\eta n(x_i)^{1-\eta} \frac{u_c(c(x_i, \theta), n(x_i))}{\mathbb{E}[u_c(c(x_i, \theta), n(x_i)) \mid x_i, z]} \mid x_i, z\right] &= n(x_i)^\nu, \quad \forall x_i, \\ N(\theta) &= \left(\int n(x_i)^{1-\eta} di\right)^{\frac{1}{1-\eta}}, \end{aligned}$$

where $c(x_i, \theta) = e^\theta N(\theta)^\eta n(x_i)^{1-\eta}$, and $\varphi(x_i, \theta \mid z, \bar{\theta})$ is the Gaussian density function of θ, x_i given z and $\bar{\theta}$.

Equilibrium approximation

The fixed-point problem does not have a closed-form solution in general, and therefore I proceed with an approximation method. The method builds on the standard perturbation method, but it is modified to address certain issues that emerge from endogenous attention choice. The approximation starts from formulating a more general sequence of economies indexed by a perturbation parameter δ that scales the size of the shocks, noises, and the attention cost in these economies:

$$\theta(\delta) = \bar{\theta} + \delta \hat{\theta}, \quad \epsilon_i(\delta) = \delta \epsilon_i, \quad \kappa(\delta) = \delta^2 \kappa, \quad \text{where} \quad \hat{\theta} = \theta - \bar{\theta}.$$

As δ goes to 0, the sequence of economies converges to a deterministic economy with a vanishing attention cost and no shocks, which can be solved easily.³ On the other hand, the economy with $\delta = 1$ corresponds to the economy to be approximated. Equilibrium objects are approximated with Taylor expansion along the sequence of economies with

³I focus on the case where the consumption-labor composite that enters the GHH utility is greater than the lower bound $\underline{u}$ when $\delta = 0$. In this case, $\underline{u}$ is inconsequential for σ small enough.

respect to δ :

$$\log n_i(\delta) = \bar{n} + \hat{n}_i \delta + \dots, \quad \log N(\delta) = \bar{N} + \hat{N} \delta + \dots, \quad \log z(\delta) = \log \bar{z} + \hat{z} \delta + \dots,$$

where $\bar{n}$ and $\hat{n}_i$ denote the zeroth and first-order expansion of $\log n_i(\delta)$ with respect to δ , and similarly for other variables and higher orders. The scaling of shocks $\hat{\theta}$ and noises ϵ_i is standard for the formulation of perturbation sequences, whereas scaling of the attention cost is an essential modification to approximate problems with attention choice. The reason that the marginal cost of attention $\kappa(\delta)$ should be scaled to vanish in order δ^2 is because the marginal value of information is second order: For economies with small shocks, deviation of labor input from its deterministic optimal level has no first-order effect on agents' payoff, therefore the gain from receiving a more precise signal is second-order. Keeping the marginal cost at order δ^2 as the marginal value of information allows the approximating sequence starting from $\delta = 0$ to feature an equilibrium information structure that resembles the one to be approximated at $\delta = 1$.⁴ The Taylor expansion described above provides an accurate approximation when the size of shock σ is small.

Similar to the standard perturbation method, the expansions can be characterized by differentiating equilibrium conditions in Lemma 1 with respect to δ to appropriate orders and evaluating them at $\delta = 0$. Each order of expansion can be solved recursively as a linear system given lower-order expansions. Equilibrium objects can be approximated to arbitrarily high order successively with the corresponding expansions. Further details are discussed in the Appendix A.

The method contains the quadratic approximation commonly used in the literature as a special case, and generalizes to higher-order approximations. The method applies not only to the static economy in this section, but also to the dynamic economy in Section 5, as well as a large class of dispersed information economies with similar structures.

3. Endogenous fluctuations in attention and uncertainty

I now study how changes in expected productivity $\bar{\theta}$ lead to variations in agents' attention level and generate movements in macroeconomic uncertainty, as captured by the four measures of uncertainty below.

Definition *Let $\tilde{Y} := \log Y$ and $\tilde{y}_i := \log y_i$. Consider the following four measures of*

⁴Scaling of the attention cost creates a “bifurcation point” at $\delta = 0$. See Judd (1998) for discussion related to bifurcation in the context of approximating a portfolio choice problem with the perturbation method.

uncertainty:

1. *Aggregate output volatility :*

$$SD(\tilde{Y} | \bar{\theta}) := \left(\mathbb{E} \left[(\tilde{Y} - \mathbb{E}[\tilde{Y} | \bar{\theta}])^2 | \bar{\theta} \right] \right)^{\frac{1}{2}}.$$

2. *Cross-sectional output dispersion :*

$$Disp(\tilde{y}_i | \bar{\theta}) := \left(\int (\tilde{y}_i - \int \tilde{y}_i)^2 di \right)^{\frac{1}{2}}.$$

3. *Forecast dispersion :*

$$Disp(\mathbb{E}_i[\tilde{Y}] | \bar{\theta}) := \left(\int (\mathbb{E}_i[\tilde{Y}] - \int \mathbb{E}_i[\tilde{Y}] di)^2 \right)^{\frac{1}{2}}.$$

4. *Subjective uncertainty :*

$$SD(\tilde{Y} | x_i, \bar{\theta}) := \left(\mathbb{E}_i \left[(\tilde{Y} - \mathbb{E}_i[\tilde{Y}])^2 \right] \right)^{\frac{1}{2}}, \forall i.$$

Each measure corresponds to a different phenomenon studied in the literature, but all are tightly connected to the notion of “fluctuations in macroeconomic uncertainty” over the business cycle. Aggregate output volatility captures how strongly aggregate production responds to aggregate shock $\hat{\theta}$. Cross-sectional output dispersion captures the differences in outcomes among agents in the economy. Forecast dispersion captures the different views held by agents about aggregate output based on their dispersed observations of the aggregate state θ . Subjective uncertainty captures the uncertainty agents face about aggregate output based on their signal x_i given their attention level z_i , which is also the size of the forecast errors agents expect for their own forecasts.

As agents’ attention level varies in response to $\bar{\theta}$, how much attention they pay affects how much information they receive, as well as how they respond to the information — both are reflected in the movements in the four measures of uncertainty. In order to isolate the role of attention in driving fluctuations in uncertainty, I show below that without endogenous attention responses, expected productivity $\bar{\theta}$ has no effect on the four measures of uncertainty (up to second-order approximation). After that, I characterize conditions under which endogenous attention can generate countercyclical movements in measures of uncertainty in response to $\bar{\theta}$.

Benchmark: An economy with exogenous attention

Consider an economy identical to the one described above, except that agents’ attention level is fixed at an exogenous level z . Given the exogenous attention level, agents make labor input choices based on idiosyncratic signals with precision z , and the final good is produced

as before. The equilibrium of this economy is characterized by the same system in Lemma 1, except that the attention optimality condition is dropped and z is fixed exogenously. First-order expansion of the system characterizes the expansion of equilibrium objects as follows:

$$\hat{n}_i = r \mathbb{E}_i[\hat{\theta}] + s \mathbb{E}_i[\hat{N}], \quad \hat{N} = \int \hat{n}_i di,$$

where the first equation is a best-response function of $\hat{n}_i$ resulting from the labor input optimality, and the second equation corresponds to the aggregation condition. Parameters $r := \frac{1}{\eta+\nu} > 0$ and $s := \frac{\eta}{\eta+\nu} \in [0, 1]$ depend on the agents' preferences and technologies, and characterize the best-response function. Parameter r captures the direct response of individual labor input to productivity shock $\hat{\theta}$. Parameter s captures how much individual labor input responds to the endogenous change in aggregate labor $\hat{N}$ (i.e., parameter s captures the level of strategic complementarity in this economy). The solution of the system has the following form:

Lemma 2

$$\hat{n}_i = \Lambda(z) \hat{x}_i, \quad \hat{N} = \Lambda(z) \hat{\theta},$$

where $\Lambda(z) > 0$ is an increasing function of z that does not depend on $\bar{\theta}$.

Function $\Lambda(z)$ captures the responsiveness of the economy to shock $\hat{\theta}$ and signal $\hat{x}_i$ in equilibrium. Responsiveness increases with attention level z because agents respond more to aggregate variables $\hat{\theta}$ and $\hat{N}$ when their signals $\{x_i\}$ are more precise. Furthermore, there is a feedback loop for how strongly agents respond: When agents are more responsive, endogenous aggregate labor $\hat{N}$ becomes more responsive to $\hat{\theta}$ as a result, and each agent becomes even more responsive. The fact that $\Lambda(z)$ does not depend on $\bar{\theta}$ is a result of the constant elasticity preference and technology: how much agents respond to a percentage change in expected productivity does not depend on the level of expected productivity. This property of the economy highlights movements in the measures of uncertainty as a sole consequence of changes in attention, as opposed to the inherent properties of preference and technology.

Lemma 3 *In a fixed-attention economy, $SD(\tilde{Y} | \bar{\theta})$, $Disp(\tilde{y}_i | \bar{\theta})$, $Disp(\mathbb{E}_i[\tilde{Y}] | \bar{\theta})$, and $SD(\tilde{Y} | x_i, \bar{\theta})$ are constant in $\bar{\theta}$ up to second-order approximation.*

With an exogenous level of attention z , responsiveness $\Lambda(z)$ is fixed. Along with homoskedastic shocks and noises, the economy is homoskedastic in expected productivity: Changes in $\bar{\theta}$ have no effect on any of the four measures of uncertainty in a fixed-attention economy.

Endogenous response in attention and uncertainty

The economy behaves quite differently when agents choose how much attention to pay endogenously in response to changes in expected productivity $\bar{\theta}$. Changes in $\bar{\theta}$ induce agents

to adjust their attention levels, which not only affect the precision of the signals they receive, but also the agents' responsiveness to information and shocks. As a result, endogenous variables in the economy are heteroskedastic in $\bar{\theta}$, despite the fact that all exogenous shocks are homoskedastic, and preferences and technologies have constant elasticity. How equilibrium attention changes with expected productivity $\bar{\theta}$ is characterized below:

Lemma 4 *Parameter γ determines the equilibrium response of attention $\bar{z}(\bar{\theta})$ to expected productivity $\bar{\theta}$:*

$$\frac{\partial}{\partial \bar{\theta}} \bar{z}(\bar{\theta}) \leq 0 \iff \gamma \geq 1.$$

There are two competing forces that determine how equilibrium attention responds to expected productivity $\bar{\theta}$: an income effect and a substitution effect of expected productivity on attention. To understand these effects, consider a decrease in expected productivity $\bar{\theta}$. The substitution effect comes from a decrease in the marginal rate of transformation between attention and “income” (by “income” I refer to the consumption-labor composite). When agents expect low productivity, they expect low labor input. When expected labor input is low, avoiding a 1% mistake in labor input translates into fewer units of income. As a result, agents have less incentive to pay attention and learn about unknown aggregate productivity θ . On the other hand, the income effect comes from an increase in the marginal rate of substitution between attention and income. With low expected productivity, agents expect a low level of income. When expected income is low, it becomes more valuable to avoid every unit of income loss due to inattention because the expected marginal utility of income is high. When $\gamma > 1$, the income effect on attention dominates and agents pay more attention when they expect lower productivity. The assumption that agents have a GHH preference over consumption and labor separates the income effect on attention from the income effect on labor. This separation permits countercyclical attention to be consistent with procyclical labor, as empirically observed in the data.⁵

Having characterized how equilibrium attention changes in $\bar{\theta}$, I show below that a decrease in $\bar{\theta}$ can lead to increases in all four measures of uncertainty through endogenous increase in attention:

Proposition 1 *Suppose that $\gamma > 1$ (equilibrium attention is decreasing in $\bar{\theta}$), then*

$$\frac{\partial}{\partial \bar{\theta}} SD(\tilde{Y} | \bar{\theta}) < 0.$$

Moreover, given parameters, there exists a threshold $\bar{z} \in \mathbb{R}$ such that if the equilibrium

⁵While the GHH preference is important for separating the income effect on attention and labor, the functional form assumption on attention cost is not essential. In Supplementary Appendix C, I show that Lemma 4 holds for general convex attention cost functions.

attention $\bar{z}(\bar{\theta}) < \check{z}$, then

$$\frac{\partial}{\partial \bar{\theta}} \text{Disp}(\tilde{y}_i | \bar{\theta}) < 0, \quad \frac{\partial}{\partial \bar{\theta}} \text{Disp}(\mathbb{E}_i[\tilde{Y}] | \bar{\theta}) < 0, \quad \text{and} \quad \frac{\partial}{\partial \bar{\theta}} \text{SD}(\tilde{Y} | x_i, \bar{\theta}) < 0, \forall i,$$

up to second-order approximation. The threshold $\check{z} > 0$ if $r > \frac{1}{2}$, and $\check{z} \rightarrow \infty$ as $s \rightarrow 1$.

The proposition states that if attention is countercyclical, a decrease in expected productivity $\bar{\theta}$ leads to an increase in aggregate output volatility; moreover, it leads to increases in cross-sectional output dispersion, forecast dispersion, and subjective uncertainty if equilibrium attention is below a certain threshold $\check{z}$ determined by parameters of the model (and does not depend on $\bar{\theta}$).

To build intuition, recall that in the fixed-attention economy, an increase in attention leads to a rise in responsiveness $\Lambda(z)$. The same is true in the endogenous-attention economy, where agents pay more attention endogenously in response to a decrease in $\bar{\theta}$. As the economy becomes more responsive to aggregate shocks, aggregate output volatility increases accordingly. The three other measures of uncertainty, nevertheless, are affected by agents' attention level through two competing channels. On one hand, when agents pay attention to reduce idiosyncratic noise in signals x_i , agents face lower uncertainty about the exogenous state θ and signal realizations become concentrated around θ (signals are less dispersed). Lower uncertainty about θ and the reduction in signal dispersion directly reduce the subjective uncertainty agents face about aggregate output, as well as the cross-sectional dispersion in forecasts and output. On the other hand, agents are more responsive to their signals when the signals are precise: agents update their beliefs more, and the same amount of idiosyncrasy in their information set generates more dispersion in their forecasts and labor input choices. Moreover, aggregate labor N responds strongly to the exogenous state θ as agents react with higher responsiveness $\Lambda(z)$. A strong response of aggregate labor N not only increases the cross-sectional dispersion in forecasts and output, it also leads to higher subjective uncertainty about aggregate output among agents, because each agent faces more uncertainty about other agents' endogenous response N under dispersed information.

Which of the two channels dominates is determined by agents' initial attention level $\bar{z}(\bar{\theta})$ relative to threshold $\check{z}$. Threshold $\check{z}$ represents the strength of the second channel, and depends on how agents respond to exogenous shocks and the endogenous response of the economy, as captured by derived parameters r and s . To understand why the second channel dominates when equilibrium attention is below a threshold, note that if equilibrium attention is low initially, agents face more uncertainty about the exogenous state θ . As a result, an increase in signal precision leads to more updates in beliefs and higher responsiveness, and consequently generates more dispersion in forecast and output, as well as higher subjective uncertainty. If $r > \frac{1}{2}$ so that the direct response of aggregate labor is strong enough, all four measures of uncertainty increase with attention as long as the

original level of attention is low enough.⁶ Furthermore, if strategic complementarity s is high, an increase in responsiveness of each agent triggers all other agents to respond strongly. As s goes to 1, and an increase in attention always increases the four measures of uncertainty

Finally, the comovement between aggregate output volatility and subjective uncertainty about aggregate output highlights the importance of dispersed information. Consider an economy without dispersed information where all agents receive a common signal x with precision z . In this economy, changes in precision z always move the two measures of uncertainty in opposite directions: an increase in z leads to higher output volatility as agents become more responsive, but always reduces subjective uncertainty about aggregate output. This is because, when sharing a common signal, agents face no uncertainty about others' endogenous actions; the only uncertainty agents face is about exogenous state θ , which reduces as signal precision z increase. By contrast, in a dispersed information economy, agents not only face uncertainty about exogenous shocks, but also face uncertainty about other agents' endogenous actions (i.e., agents face "strategic uncertainty"). Because uncertainty about endogenous actions depends on how strongly other agents react, agents can face higher uncertainty about others' aggregate response despite the fact that each agent receives a more precise signal individually.

Discussion

Generality of the results

The key observation about how attention affects the four measures of uncertainty applies to a general class of "beauty-contest" models characterized by the two reduced-form parameters r and s , including dispersed information New Keynesian models and models of financial markets reviewed by Angeletos and Lian (2016). In this class of models, an increase in the precision of idiosyncratic signals increases the four measures of uncertainty as long as the conditions in Proposition 1 holds, as shown in the Appendix A.

Separating the income effect on attention and risk aversion

Parameter γ plays a dual role of determining agents' risk aversion and the strength of the income effect on attention — the former describes how substitutable are income realizations across different states of the world, whereas the later describes how agents trade off between their utility from income and their disutility from paying attention. In the appendix, I show these two considerations can be separated by a preference, á la Epstein and Zin (1989), that trades off between the disutility of attention and a certainty equivalence of income with one parameter that governs the strength of the income effect, and a different risk-aversion parameter that converts risky income into its certainty equivalence.

⁶The condition $r > \frac{1}{2}$ is only needed for subjective uncertainty. The two dispersion measures are always countercyclical when attention is low. See appendix for detail.

Endogenous attention can amplify exogenous volatility shocks

The mechanism above is complementary to works that study the effect of exogenous volatility shocks, such as Bloom (2009) and others. While I focus on exploring how an endogenous attention choice can transmit first moment shocks to generate fluctuations in second moment variables, the same mechanism also amplifies the response of the four measures of uncertainty to exogenous volatility shocks. In Supplementary Appendix C, I show that, in the fixed-attention economy, the four measures of uncertainty move one to one (in percentage terms) with the volatility of aggregate productivity σ , whereas the four measures of uncertainty move more than one to one with σ in the endogenous-attention economy under the conditions in Proposition 1. Furthermore, a mechanism that generates fluctuations in uncertainty without exogenous volatility shocks is useful because measurement of exogenous volatility shocks is tricky. In Supplementary Appendix D, I show the volatility of aggregate TFP becomes acyclical when TFP is adjusted for utilization á la Fernald (2014).

4. Empirical evidence on fluctuations in uncertainty and attention

Fluctuations in the four measures of uncertainty and attention have been widely studied empirically in the literature. Although approaches vary, the countercyclical patterns of the four measures of uncertainty and attention emerge as prominent features of the data. This section presents empirical evidence on the four measures of uncertainty and attention that closely follows the literature, but in a way that connects directly to the qualitative result in Proposition 1 and the quantitative result in Section 7.

Countercyclical fluctuations in uncertainty

I construct four empirical measures of uncertainty corresponding to the four measures of uncertainty studied in Section 3. The empirical measures are constructed following standard procedures in the literature, but with minor modifications so that they map closely to the model. A short description is provided for each measure, and their fluctuations over the business cycle are presented below. Detailed descriptions of the data are provided in Supplementary Appendix D.

1. *Aggregate output volatility:* To measure aggregate output volatility, I estimate the conditional heteroskedasticity of quarterly real GDP growth with a EGARCH(1,1)-ARMA(2,2) model; the result is robust to different GARCH and mean process specifications. The measure captures how shocks to expected aggregate output feed into the volatility of aggregate output. The measure is closely related to works that study countercyclical fluctuations in aggregate volatility; see Jurado et al. (2015), Ilut et al. (2018), and Adrian et al. (2019), for example. I use the GARCH estimation to construct aggregate output volatility in the data because the estimation does not rely on

variables that are not in the model, and it allows a direct comparison between data and model in Section 7. The comparison between data and model views the GARCH estimation as a pure statistical procedure that captures certain features of the data, and it allows me to examine whether the model-generated process can replicate such features. Denote the empirical measure of aggregate output volatility by

$$\widetilde{SD}_t(\Delta\tilde{Y}_t).$$

2. *Cross-sectional output dispersion*: To measure cross-sectional output dispersion, I calculate the cross-sectional standard deviation of quarterly revenue growth of firms within sectors for each quarter in the Compustat data, and calculate the revenue-weighted average dispersion for each quarter. The measure corresponds to the countercyclical cross-sectional dispersion of output and employment studied in Bloom et al. (2018), Ilut et al. (2018), and others. Denote the empirical measure of cross-sectional output dispersion by

$$\widetilde{Disp}_t(\Delta\tilde{y}_{i,t}).$$

3. *Forecast dispersion about aggregate output*: To measure forecast dispersion about aggregate output, I calculate the cross-sectional standard deviation of one-quarter-ahead estimates of GDP growth from the Survey of Professional Forecasters (SPF) in each quarter. Previous works such as Bachmann et al. (2013) and Kozeniauskas et al. (2018) document the countercyclical fluctuations in forecast dispersion. The empirical measure of forecast dispersion is denoted by

$$\widetilde{Disp}_t(f_{i,t}^{\Delta\tilde{Y}_t}),$$

where $f_{i,t}^{\Delta\tilde{Y}_t}$ is the point forecast of each individual forecaster.

4. *Subjective uncertainty about aggregate output*: To measure subjective uncertainty about aggregate output, I use the probability-range data in the SPF to construct the subjective uncertainty agents face when forecasting aggregate output. The survey asks each forecaster to assign probability weights to different ranges of possible aggregate output growth. The literature has used this data to study countercyclical subjective uncertainty; see Bloom (2014) and Fajgelbaum et al. (2017), for example. Following Engelberg et al. (2009), I fit a parametric distribution to the discrete probability weights submitted by each forecaster, calculate the standard deviation of the distribution, and average over forecasters. Let $\Phi_{i,t}(\Delta\tilde{Y}_t)$ denote the fitted distribution over aggregate output growth for forecaster i at time t . The empirical measure of subjective uncertainty is defined as

$$\overline{\widetilde{SD}_{i,t}(\Delta\tilde{Y}_t)},$$

where $\widetilde{SD}_{i,t}(\Delta\widetilde{Y}_t)$ is the standard deviation of $\Delta\widetilde{Y}_t$ under distribution $\Phi_{i,t}(\Delta\widetilde{Y}_t)$.⁷

Figure 1: Aggregate output volatility, cross-sectional output dispersion, forecast dispersion about aggregate output, and subjective uncertainty about aggregate output. x-axis: time; y-axis: log-deviation of variables from the long-run average; gray areas indicate NBER recessions.

Figure 1 shows the log deviations of the four measures of uncertainty from their long-run averages over time, where the NBER recession periods are marked by the gray areas. All four empirical measures of uncertainty show strongly countercyclical comovement, rising sharply during recessions and declining during booms. In Proposition 1, these fluctuations arise endogenously as a result of economic agents paying more attention to macroeconomic events while getting their information from dispersed sources. Admittedly, the phenomena presented in Figure 1 do not map exactly to the four measures of uncertainty in the economy studied in Section 3 — all these fluctuations are intrinsically dynamic phenomena in the data, whereas the economy studied is placed in a static setup. To bring the model closer to the data, I extend the economy to a dynamic setup below, and compare business cycle

⁷For the probability-range data, the SPF asks for fixed-event forecasts of year-over-year GDP growth. To adjust for seasonality, I control for quarter-of-year fixed effect when constructing the average subjective uncertainty. The same adjustment is applied to the model-generated data in Section 7 so the series are comparable.

statistics calculated from the series in Figure 1 to their counterparts constructed from the model-generated data in Section 7.

Countercyclical fluctuations in attention

Existing empirical measures of attention to economic events fall into two categories: One uses proxies for attention to establish qualitative patterns; the other uses forecast survey data to provide a quantitative but indirect measure of attention.

The empirical finance literature uses internet traffic data as a proxy for the attention of retail and institutional investors; see Da et al. (2011) and Ben-Rephael et al. (2017), for example. Following this approach, I provide evidence in Supplementary Appendix D that people’s attention to economic events is countercyclical. I construct two different proxies for attention to economic events, using Google search frequency of economic news and economic terms. Both attention proxies show strong countercyclical fluctuations. Flynn and Sastry (2021) also provide evidence that firms’ attention to macroeconomic issues are countercyclical, where attention is proxied by the similarity of words between firms’ regulatory filings and macroeconomics textbooks, using a natural-language-processing algorithm.

For a quantitative yet indirect measure of attention, Coibion and Gorodnichenko (2015) provide a method to estimate how much professional forecasters update their beliefs about macroeconomic variables. They show that, during recessions, professional forecasters in SPF update their expectations about macroeconomic variables with larger Kalman gains. This finding is consistent with the mechanism proposed by this paper, where agents pay attention to obtain signals with a higher signal-to-noise ratio during recessions, and use these signals to update their beliefs about the aggregate economy. The quantitative nature of the evidence allows a direct comparison between data and model in Section 7. There, the model generates patterns in forecasts that are similar to those empirically observed in the forecast survey that indicates countercyclical attention.

5. Dynamic model

Fluctuations in uncertainty and attention are intrinsically dynamic phenomena. I now extend the model to a dynamic economy, which allows me to compare the four measures of uncertainty and attention constructed from model simulation directly to that from the U.S. data.

The dynamic economy consists of the same agents and final good producer as before. Time lasts from $t = 0, \dots, \infty$, and each period t splits into three stages in which agents pay attention, decide labor input, and consume. Each agent observes a new signal about unknown aggregate productivity in stage 2, where the signal precision depends on their attention; each agent observes the price of their good in stage 3. Aggregate productivity

is persistent, and the use of intermediate goods in final good production is affected by idiosyncratic shocks. The presence of idiosyncratic shocks introduce additional disturbances to the prices intermediate goods, which prevent agents from making perfect inference about the past aggregate productivity from simply observing the price of their good.

Preferences and technologies

The preference of an agent i over consumption, labor, and attention is given by

$$\mathbb{E}_{i,0} \sum_{t=0}^{\infty} \beta^t \left(u(c_{i,t}, n_{i,t}) - \kappa z_{i,t} \right),$$

where $u(c_{i,t}, n_{i,t})$ is given by the same period utility function as in Section 2. 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 and saving to focus on the dynamics generated by changes in agents' information structures over the business cycle.

A representative competitive final good producer maximizes profit $Y_t - \int p_{i,t} y_{i,t}$ with technology of constant elasticity of substitution $1/\eta$ that combines intermediate goods to produce the final good:

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

where aggregate productivity θ_t follows an AR(1) process with persistence ρ ,

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

and $\omega_{i,t}$ are idiosyncratic productivity shocks for intermediate goods, which are i.i.d. normal over time and across intermediate goods with common variance $\sigma_{\omega_i}^2$.

Timing and information

Each period t consists of three stages. In stage 1, each agent i chooses attention level $z_{i,t}$ given their respective information set $\mathcal{F}_{i,t-1}$, which is to be defined below. In stage 2, according to each agent's attention level, agents receive idiosyncratic signals

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

Agents make labor input decision $n_{i,t}$ 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 own product $p_{i,t}$ and consume $c_{i,t}$ subject to budget constraints. To summarize, each agent i

chooses stochastic processes $z_{i,t}, n_{i,t}, c_{i,t}$ under the following information constraints:

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

where x_i^t, p_i^t denote the histories of respective variables up to time t , and $\sigma(\cdot)$ denotes the σ -algebra generated by the respective processes. Without loss of generality, assume that agents have a common prior $\theta_0 \sim \mathcal{N}(0, \sigma_\omega^2)$ at period $t = 0$, which is inconsequential for the stationary properties of the economy.

Finally, in each period t , the final good producer observes $\{\theta_t, \omega_{i,t}, p_{i,t}\}$, and chooses intermediate goods input $\{y_{i,t}\}$ to maximize profit.

Definition of equilibrium

An equilibrium consists of 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) production of $q_{i,t}, Y_t$ is given by the respective technologies; and (iv) markets clear for all intermediate goods $q_{i,t} = y_{i,t}, \forall i, t$, and final good $Y_t = \int c_{j,t}, \forall t$.

Equilibrium characterization

Similar to the static setup, in this economy the prices of intermediate goods $\{p_{i,t}\}$ can be solved from the final good producer's profit-maximization problem and market clearing, given the distribution of labor input $\{n_{i,t}\}$:

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

Note the price of an intermediate good, $p_{i,t}$, depends not only on the aggregate condition θ_t and N_t , but also on the realization of idiosyncratic shocks $\omega_{i,t}$. Idiosyncratic shocks $\omega_{i,t}$ shifts the final good producer's demand for intermediate good i , but it is not directly observed by agent i . The presence of idiosyncratic shocks prevent agents from making perfect inference about the past realizations of aggregate state $\{\theta_\tau\}_{\tau < t}$. This feature captures the difficulty for agents to make inference about the aggregate economy based on a small set of prices they passively observe; it is quantitatively important for the model to match realistic size of forecast errors in the forecast survey, and allows active learning with endogenous attention choice to play an important role.

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

Lemma 5 *An equilibrium solves the following system:*

$$\begin{aligned} \sum_{\tau=t}^{\infty} \beta^{\tau-t} \mathbb{E} \left[(U_{i,\tau} - \beta \kappa z_{i,\tau+1}) \times \frac{1}{2z_{i,t}} (1 - \epsilon_{i,t}^2) \mid \mathcal{F}_{i,t-1} \right] &= \kappa, \forall i, t, \\ \mathbb{E} \left[\frac{\partial}{\partial n_{i,t}} U_{i,t} \mid \mathcal{F}'_{i,t} \right] &= 0, \forall i, t, \\ N_t &= \left(\int \left(e^{\omega_{i,t}} n_{i,t} \right)^{1-\eta} \right)^{\frac{1}{1-\eta}}, \forall t, \end{aligned}$$

where $U_{i,t}(\theta_t, N_t, n_{i,t}, \omega_{i,t}) := u(e^{\theta_t + \omega_{i,t}} N_t^\eta n_{i,t}^{1-\eta}, n_{i,t})$ denotes agent i 's period utility from consumption and labor in period t .

The first two conditions are necessary for the optimality of attention choice and labor input. The trade-off resembles that in the static model. The only differences are that (i) agents now enter each period with different information set $\mathcal{F}_{i,t-1}$ instead of a common prior, and (ii) the marginal benefit of attention now depends on the use of information for all future decisions, which is captured by how an increase in attention $z_{i,t}$ alters the probability density functions and simplifies to the term $\frac{1}{2z_{i,t}}(1 - \epsilon_{i,t}^2)$. The third condition aggregates individual labor into aggregate labor, taking into account idiosyncratic shocks $\omega_{i,t}$ for different intermediate goods in the final good producer's production function.

Due to the idiosyncratic shocks that prevent agents from making perfect inference about the past aggregate states, the economy contains a feature often referred to as “persistent dispersed information”, with which agents enter each period with different information set $\mathcal{F}_{i,t-1}$, and never reach perfect consensus about the aggregate state. Persistent dispersed information is important because it allows the model to generate quantitatively plausible sizes of forecast errors as in the U.S. forecast survey while producing fluctuations in output and labor consistent with salient business cycle moments of the U.S. economy: without persistent dispersed information, matching the persistence and volatility of output and labor will imply an upper bound for the sizes of forecast errors that is much smaller than that in the U.S. forecast survey, because agents “learn too much” from simply observe the prices of their good. Persistent dispersed information reconciles the sizes of forecast errors and other business cycle moments in the data.

While persistent dispersed information is necessary for the model to generate key features of the data, it poses a challenge known as the “infinite regress problem.” The problem is as follows: Because agents never observe the aggregate state θ_t , their actions will in general depend on the whole history of the signals they received in the past. If agents were to share a common information set, the history can be succinctly summarized by the conditional expectation and variance of θ_t under the Gaussian information structure. However, because information is dispersed, agents each have different conditional expectations and variances

of θ_t . In order to forecast other agents' equilibrium actions, each agent needs to track additional state variables to include relevant information from the history of signals. Yet, this process goes on ad infinitum: Whenever agents track an additional state variable, each agent needs to track even more state variables to form beliefs about that additional state variable. In general, agents would not be able to summarize their past information with a finite number of state variables, and each agent's equilibrium action will be a nonlinear function of a whole infinite-dimension history.

Many methods have been proposed to address the infinite regress problem within the framework of linear rational expectation models.⁸ However, existing methods are constrained to first-order approximation and miss higher-order dynamics of the economy: In particular, existing methods cannot capture fluctuations in attention or the four measures of uncertainty, because these fluctuations are intrinsically higher-order properties of the model. Below I introduce a computational procedure that extends the perturbation method introduced in Section 2 to compute higher-order dynamics of dispersed information models with an infinite regress problem.

Computational procedure

Similar to the approximation of the static economy, equilibrium objects in the economy are approximated by Taylor expansions with respect to a sequence of economies scaled by perturbation parameter δ . For example, the expansion of agent i 's labor input is given by

$$n_{i,t}(\delta) = \bar{n} + \hat{n}_{i,t} \delta + \frac{1}{2} \hat{\hat{n}}_{i,t} \delta^2 + \dots,$$

Different from the static model, however, an equilibrium strategy in a dynamic economy with an infinite regress problem is now a nonlinear function of the infinite-dimension history of signals and prices, $s_i^t = (x_i^t, p_i^{t-1})$, and the expansions are multilinear functions of the history $\hat{s}_i^t$:

$$\hat{n}_{i,t} = \mathbf{n}_s \hat{s}_i^t, \quad \hat{\hat{n}}_{i,t} = \hat{s}_i^{t\top} \mathbf{n}_{ss} \hat{s}_i^t + 2\mathbf{n}_{s\delta} \hat{s}_i^t + \dots$$

In order to solve the expansions numerically, one needs to look for a finite-dimension approximation. Note that expansions are multilinear in $\hat{s}_i^t$, which are themselves linear in shocks and noises $\omega_t, \omega_{i,t}, \epsilon_{i,t}$. An analogy of the spectrum theorem motivates the search for optimal factors $f_{i,t}$'s (for a given dimension) in the class of factors with linear dynamics,

⁸For example, Lorenzoni (2009) truncates the history of signals with a fixed time window; Nimark (2017) truncates belief hierarchy above a certain order; and Huo and Pedroni (2020) and Huo and Takayama (2015) show that for a certain class of information structure, the history of signals can actually be summarized by a finite number of state variables. Huo and Pedroni (2020) and Huo and Takayama (2015) show analytic solution for cases in which a finite number of state variables exist, and provide a numerical method for cases where the analytic solution does not apply.

along with some corresponding multilinear functions Φ 's:

$$\hat{n}_{i,t} \approx \Phi_f f_{i,t}^{(1)}, \quad \hat{n}_{i,t} \approx f_{i,t}^{(2)\top} \Phi_{ff} f_{i,t}^{(2)} + 2\Phi_{f\delta} f_{i,t}^{(2)} + \dots,$$

where

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

The optimal factors for each order of approximation and the corresponding multilinear functions can be solved recursively: Given the $(m-1)^{th}$ order approximation, consider the following two-step procedure: (1) For a given factor structure $A^{(m)}$, $C^{(m)}$, simulate the economy and solve the corresponding multilinear functions Φ 's that minimize the residuals of the expanded equilibrium conditions; (2) optimize over factor structure $A^{(m)}$, $C^{(m)}$ to look for the optimal factor structure.

The procedure is reminiscent of that in Krusell and Smith (1998), but the approaches differ in subtle ways. The infinite-dimension state problem here comes from both the time dimension — due to infinite regress (the infinite-dimension history), and the cross-sectional dimension — due to agents having heterogeneous information and beliefs. The two-step procedure described above searches for finite-dimension factors that summarize infinite-dimension histories along the time dimension; this is in contrast to methods that solve heterogeneous agent models where summarizing cross-sectional distribution is key.⁹ The complication of cross-sectional heterogeneity, on the other hand, is alleviated by the perturbation approach due to the multilinear nature of the Taylor expansions and the Gaussian structure of shocks: expansions of aggregate variables can be solved easily in closed forms given a factor structure $A^{(m)}$, $C^{(m)}$ and the corresponding multilinear functions Φ 's. This aggregation result is common in linear dispersed information models, which extends naturally to higher order expansions.

The procedure is similar to existing methods addressing the infinite regress problem in that sense that it summarizes the infinite dimensional history with finite dimension factors. The difference is that instead of using factors proposed base on heuristics, as in Lorenzoni (2009) or Nimark (2017), the procedure here attempts to solve for the optimal factor for a given dimension within a class of factor specifications. In comparison to Maćkowiak and Wiederholt (2015, 2009) and the computational method in Huo and Takayama (2015), the procedure here does not conjecture the dynamics of the equilibrium objects directly, but solves them as multilinear functions of underlying factors with linear dynamics. This difference is inconsequential for first-order approximation, yet it becomes crucial for approximation of higher-order dynamics where a good conjecture of equilibrium stochastic process is not necessarily obvious a priori.

⁹See, for example, Ahn et al. (2018), Bhandari et al. (2018) and Fernández-Villaverde et al. (2019) for methods that address cross-sectional heterogeneity.

6. Calibration

I calibrate the model to target key moments from the U.S. aggregate data and the SPF. Each period in the model corresponds to a quarter. Agents in the model represent managers of firms whose forecasts are represented by forecasters in the SPF. Stages 1 and 2 of each period occur at the beginning of the corresponding quarter, at which point agents pay attention, receive information, make forecasts and make labor input decisions. Stage 3 occurs at the end of the quarter at which point production of goods realizes and output is recorded. The equilibrium is approximated to the second order with the method described above, where computational details are provided in Appendix B.

External parameters

Discount rate β is set at .99 to reflect that events happen at the quarterly frequency. Elasticity of substitution between intermediate goods $1/\eta$ is set at 5 so that the average markup over the marginal cost of labor is 25% in the steady state. The standard deviation of idiosyncratic shocks σ_{ω_i} is set at 5.1%, which corresponds to the quarterly standard deviation of microvolatility shocks in Bloom et al. (2018). Parameter γ plays a dual role of determining risk aversion and the strength of the income effect on attention, and cannot be distinguished without departure from the standard preference specification.¹⁰ I set γ to 10, which implies a strong income effect on attention that generates countercyclical attention choices that are consistent with data.

Internal calibration

The remaining parameters are calibrated internally. Parameters to be calibrated are the persistence of aggregate productivity process ρ , the standard deviation of aggregate productivity innovation σ_ω , the convexity of the labor cost ν , and the marginal cost of attention κ . Parameters ρ, σ_ω, ν are directly linked to the input and output of production. To calibrate these parameters, I target (i) the persistence of aggregate output, (ii) the (unconditional) aggregate output volatility, and (iii) the relative volatility of hours to output. The marginal cost of attention κ is directly linked to the information structure of the economy, which is reflected in the forecast survey data. To calibrate κ , I target the size of the average forecast error of aggregate output from the SPF. Let $\bar{f}e_t := \tilde{Y}_t - \int \mathbb{E}'_{i,t-1}[\tilde{Y}_t] di$ denote the one-quarter-ahead average forecast error of aggregate output among forecasters in period t . The size of the average forecast error $||\bar{f}e_t||$ is defined as the standard deviation of $\bar{f}e_t$ over time. Table 1 shows the calibrated parameters with targeted moments from the data and the model.

¹⁰Parameter γ affects business cycle moments only through its effect on attention fluctuations. This is because the level of risk aversion matters, but it only affects the level of labor and output up to second order approximation; moreover, the model does not contain an intertemporal trade-off because agents live hand-to-mouth.

Table 1: Calibration results

Parameter	ρ	σ_ω	ν	$1/\sqrt{\bar{z}(\kappa)}$
	.86	.002	.07	.002
Moment	$\rho_1 Y_t^\dagger$	$sd Y_t^\dagger$	$\frac{sd N_t^\dagger}{sd Y_t^\dagger}$	$ fe_t $
Data	.93	2.1	.91	.46
Model	.91	2.1	.91	.47

†: Band-pass filtered with frequency corresponding to 6-32 quarters.

Parameters ν and κ are of special importance. Parameter ν is closely related to the level of strategic complementarity $s = \frac{\eta}{\eta + \nu}$, and κ determines the average attention level. The calibrated model features a large elasticity of labor supply $1/\nu$ because the volatility of hours is large relative to output, which leads to a high level of strategic complementarity in the model. Although the elasticity of labor supply is large if an agent is interpreted as an individual, it can be rationalized if an agent is interpreted as a collection of manager and a group of workers, and the elasticity of labor supply captures responses from the extensive margin á la Rogerson (1988). Under this interpretation, the disutility of labor in the GHH preference represents the labor-cost function a manager faces, and the consumption-labor composite corresponds to the profit the manager receives after subtracting labor cost from revenue. Parameter κ is calibrated to target the size of the average forecast error. The size of the average forecast error is directly related to the average attention level in the model: Given the same level of (unconditional) aggregate output volatility, a larger size of the average forecast error implies a lower attention level on average, which corresponds to a high attention cost. The last column shows signal precision $\bar{z}(\kappa)$ instead of attention cost κ because κ is unit of utils, which does not have meaning per se. As discussed in Section 3, the level of strategic complementarity and the average attention level are key properties of the model that determine how fluctuations in attention affect fluctuations in the measures of uncertainty.

7. Quantitative result

I simulate the dynamics of the calibrated economy to generate business cycle statistics that capture the fluctuations in uncertainty and attention, and compare these statistics to corresponding statistics for the U.S. economy to evaluate the quantitative potential of the mechanism. For fluctuations in the four measures of uncertainty, I calculate aggregate output volatility, cross-sectional dispersion of output, forecast dispersion about aggregate output, and subjective uncertainty about aggregate output, as described in Section 4. For

fluctuations in attention, I compare measures of attention constructed from forecasts in the model to those from the SPF, following the method provided by Coibion and Gorodnichenko (2015).

Fluctuations in uncertainty: Cyclicalities

Table 2 compares the cyclicalities of the four uncertainty measures generated by the calibrated model to corresponding statistics in the data.

Table 2: Cyclicalities of measures of uncertainty

		$\widetilde{SD}_t(\Delta\tilde{Y}_t)$	$\widetilde{Disp}_t(\Delta\tilde{y}_{i,t})$	$\widetilde{Disp}_t(f_{i,t}^{\Delta\tilde{Y}_t})$	$\overline{\widetilde{SD}_{i,t}(\Delta\tilde{Y}_t)}$
$cor(\cdot, \tilde{Y}_t)$	Data	-.69	-.53	-.58	-.65
	Model	-.98	-.66	-.65	-.85

Series are band-pass filtered with frequency corresponding to 6-32 quarters.

The calibrated model generates countercyclical comovements in all four uncertainty measures. Note this result is not a mechanical outcome of the model setup. The model contains no a priori assumption about how the four uncertainty measures should fluctuate over the business cycle: Had the target moments taken on different values in the data, the model would have generated measures of uncertainty with the opposite cyclical patterns. Countercyclical fluctuations in the four measures of uncertainty is a quantitative result when key features of the model are pinned down by empirical features of the data in the calibration. Corroborating the theoretical insight in Section 3, countercyclical attention generates countercyclical comovements in the measures of uncertainty only because the calibrated model features a high level of strategic complementarity and a low level of attention on average, resulting from targeting the volatility in hours and the size of the average forecast error as discussed in Section 6.

Fluctuations in uncertainty: Levels, magnitudes, and persistence

Table 3 shows the average levels, the magnitudes of fluctuations, and the persistence of the four measures of uncertainty generated by the model and the corresponding statistics in the data.

Table 3: Levels, magnitudes, and persistence of the measures of uncertainty

		$\widetilde{SD}_t(\Delta\tilde{Y}_t)$	$\widetilde{Disp}_t(\Delta\tilde{y}_{i,t})$	$\widetilde{Disp}_t(f_{i,t}^{\Delta\tilde{Y}_t})$	$\overline{\widetilde{SD}_{i,t}(\Delta\tilde{Y}_t)}$
<i>avg</i>	Data	.96	10.8	.22	.41
	Model	1.1	7.3	.57	.83
<i>sd</i> [†]	Data	.20	.64	.11	.03
	Model	.05	.01	.03	.01
$\rho_1^\dagger$	Data	.89	.86	.87	.87
	Model	.88	.89	.86	.89

†: Band-pass filtered with frequency corresponding to 6-32 quarters.

The calibrated model generates time series for the four measures of uncertainty with levels, magnitudes of fluctuations, and persistence that are roughly consistent with the data, despite the fact that none of the moments in Table 3 are targeted in the calibration. For aggregate output volatility, forecast dispersion, and subjective uncertainty, the levels of uncertainty generated by the model are slightly higher than those observed in the data; whereas the size of fluctuations in these measures are around 25% to 33% of their counterparts in the data at business cycle frequency.

The fluctuations in the cross-sectional dispersion of output are what the model misses the most. The model generates cross-sectional output dispersion that is slightly lower than in the data, and the size of fluctuations is a magnitude lower. This should not come as a surprise, however. In the data, there is tremendous heterogeneity among firms in terms of their ages, sizes, balance sheets, etc, that affects how they respond to economic conditions. The model abstracts away from all these heterogeneities and does not feature any interaction between agents' dispersed information and the idiosyncratic productivity shocks. An exploration of how dispersed information interacts with other types of heterogeneity in the cross section can potentially generate larger fluctuations in cross-sectional output dispersion, bringing the model closer to the data.

The model generates persistence in all four measures of uncertainty that is similar to that in the data. This comes from the fact that fluctuations in the measures of uncertainty are tightly connected to the fluctuations in aggregate output through the endogenous response of attention. Whereas models that feature exogenous volatility shocks often need to assume persistent dynamics for the volatility process exogenously, the mechanism in this paper provides a parsimonious way to explain the joint dynamics of the four measures of uncertainty as integrated phenomena of the business cycle.

Finally, the magnitudes of fluctuations in the four measures of uncertainty depend on how much agents' attention varies over the business cycle, which in turn hinges on the strength of the income effect on attention in the model. Without deviating from standard preference assumptions, the income effect on attention is inevitably governed by the same parameter γ as risk aversion and was set externally at a value commonly used in the literature. I show below that parameter γ chosen for the calibrated model appears to generate less fluctuations in attention than that in the SPF, using a measure of attention constructed from agents' forecasts over the business cycle. This is consistent with the fact that the four measures of uncertainty fluctuate less in the model than in the data, and suggests that a generalization of preferences that separate income effect on attention from risk aversion (as suggested in Supplementary Appendix C) can be an important direction to understand the fluctuations in attention and uncertainty.

Fluctuations in attention: Evidence from forecast survey data

Although agents' attention to aggregate events is not directly observable, how much attention they pay affects their information and beliefs about macroeconomic variables and leaves traces in forecast survey data. Coibion and Gorodnichenko (2015) show that people's attention level is closely related to the regression coefficient from regressing the average forecast error on the average forecast revision: When people pay more attention, the average forecast in each period contains more currently available information about the economy; therefore, the average forecast error becomes less predictable and is associated with a smaller regression coefficient. By comparing how the coefficient changes over the business cycle, we can learn about how people's attention adjusts to different economic conditions.

Specifically, let $f_{i,t,s}^{\Delta Y} := \mathbb{E}'_{i,t-1}[\Delta \tilde{Y}_{t,s}]$ denote agent i 's forecast at period t about aggregate output growth between period t and period s , and $\bar{f}^{\Delta Y}_{t,s} := \int f_{i,t,s}^{\Delta Y} di$ denote the average forecast among forecasters. The average forecast error and average forecast revision about aggregate output growth between period t and s are given respectively by

$$\bar{f}e_{t,s} := \Delta \tilde{Y}_{t,s} - \bar{f}^{\Delta Y}_{t,s}, \quad \text{and} \quad \bar{f}r_{t,s} := \bar{f}_{t,s} - \bar{f}_{t-1,s}.$$

Coibion and Gorodnichenko (2015) consider the following regression:

$$\bar{f}e_{t,s} = \alpha + \beta_{CG} \bar{f}r_{t,s} + \text{residual}_{t,s}, \forall t, s,$$

where they allow coefficient β_{CG} to depend on the time between period t and the closest NBER recession start date before period t . They show that the coefficient β_{CG} declines after recessions start and increases as the economy recovers. Moreover, β_{CG} is closely related to the inverse of the Kalman gain agents use to update their beliefs about aggregate variables in a dispersed information model, where the decline of coefficient β_{CG} is associated with

agents receiving idiosyncratic signals about aggregate variables that are more precise (that is, agents pay more attention to aggregate variables). To make the exercise applicable to the model-generated data without relying on recession dating, I split the time periods into periods with output above trend, $\tilde{Y}_t^\dagger > 0$, and periods with output below trend, $\tilde{Y}_t^\dagger < 0$. For each group, I calculate β_{CG} as a measure of average attention in the respective time periods. Table 4 reports the coefficient β_{CG} for periods above and below trend, respectively, for for the U.S. forecast survey and the model-generated data.

Table 4: Measures of attention from forecast data

		$\tilde{Y}_t^\dagger > 0$	$\tilde{Y}_t^\dagger < 0$
β_{CG}	Data	.76	.45
	Model	.77	.67

†: Band-pass filtered with frequency corresponding to 6-32 quarters.

For both the U.S. data and the model, coefficients β_{CG} are smaller when output is low, and are larger when output is high. This qualitative similarity suggests that agents change how they update beliefs between good times and bad times in a manner similar to that of the forecasters in the SPF. Through the length of the model, this pattern indicates countercyclical fluctuations in attention, consistent with the evidence provided by Coibion and Gorodnichenko (2015), as well as the countercyclical movements in other attention proxies discussed in Section 4.

The gap between coefficients β_{CG} of good and bad periods is informative of how much agents’ attention fluctuates over the business cycle. The gap is smaller in the model than its empirical counterpart. This suggests that the calibrated model generates less fluctuations in attention than in the data. This is consistent with the fact that uncertainty measures also fluctuate less in the model than in the data, and suggests that a stronger income effect on attention will generate larger fluctuations in attention and all four measures of uncertainty and bring the model closer to the data.

8. Concluding remarks

This paper provides a theory that connects economic agents’ attention choice to four prominent business cycle phenomena related to the fluctuations in macroeconomic uncertainty. Countercyclical fluctuations in attention turns homoskedastic shocks into heteroskedastic responses, and generate countercyclical fluctuations in aggregate output volatility, cross-

sectional output dispersion, forecast dispersion and subjective uncertainty as integrated phenomena of the business cycle. When calibrated to match basic features of the U.S. economy, countercyclical attention generates countercyclical fluctuations in all four measures of uncertainty, consistent with evidence from the data.

In the model, uncertainty rises endogenously as agents pay attention to macroeconomic events while receiving dispersed information. Exploring the normative aspect of the mechanism will be valuable for understanding how macroeconomic policies can mitigate the surge in macroeconomic uncertainty during economic crises. The model in this paper is deliberately stylized so that the mechanism can be kept transparent. This is at the cost of abstracting from other interesting features of the business cycle and left limited room for analysis of macroeconomic policies. Incorporating these elements will be useful for both a comprehensive quantitative assessment of the mechanism and an examination of its policy implication. I leave this for future work.

A. Derivations and proofs

Proofs of Lemma 1

Rewrite agents' preference as

$$\frac{1}{1-\gamma} v_i^{1-\gamma} - \kappa(z_i),$$

where

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

Proof of Lemma 1. The third condition in Lemma 1 results from standard CES algebra. For the optimality condition of labor n_i given a realization of signal x_i and attention z_i , substituting $c_i = e^\theta N^\eta n_i^{1-\eta}$ and taking the first order condition with respect to n_i gives:

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

Dividing both sides by $R_i := \mathbb{E}_i[v_i^{-\gamma} \mathbf{1}_{\{v_i > \underline{u}\}}]^{-\frac{1}{\gamma}}$, and moving $\nu'(n_i)$ to the right hand side gives the second condition in Lemma 1:

$$\mathbb{E}_i \left[\left((1-\eta) e^\theta N^\eta n_i^{-\eta} \right) \left(\frac{v_i}{R_i} \right)^{-\gamma} \mathbf{1}_{\{v_i > \underline{u}\}} \right] = \nu'(n_i), \quad \forall x_i.$$

On the other hand, let $U(z_i)$ denote agent i 's value function given attention z_i :

$$V(z_i) := \max_{n(\cdot)} \mathbb{E} \left[\frac{1}{1-\gamma} \left(\max\{e^\theta N(\theta)^\eta n(x_i)^{1-\eta} - \nu(n(x_i)), \underline{u}\} \right)^{1-\gamma} \middle| z_i \right] - \kappa(z_i)$$

where the equilibrium aggregate labor $N(\theta)$ is taken as given.

Optimality of attention choice implies that $V'(z_i) = 0$. By envelope theorem, $V'(z_i) = 0$ is equivalent to

$$\int \frac{1}{1-\gamma} \left(\max\{e^\theta N(\theta)^\eta n(x_i)^{1-\eta} - \nu(n(x_i)), \underline{u}\} \right)^{1-\gamma} \frac{\partial}{\partial z_i} \varphi(\theta, x_i | z_i) d\theta dx_i - \kappa'(z_i) = 0,$$

which corresponds to the first condition in Lemma 1. $\square$

Proofs of Lemma 2, 3, and 4

Below I first provide details of the perturbation method, and proceed to show Lemma 2, 3, and 4.

Consider a sequence of economy parameterize by δ such that

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

where

The equilibrium in economy δ is described by the conditions in Lemma 1:

$$\begin{aligned} \mathbb{E} \left[\frac{1}{1-\gamma} v_i(\delta)^{1-\gamma} \times \frac{1-\epsilon_i^2}{2z(\delta)} \Big| z(\delta) \right] - \delta^2 \kappa(z(\delta)) &= 0 \\ \mathbb{E} \left[v_i(\delta)^{-\gamma} \left((1-\eta)e^\theta N^\eta n_i^{-\eta} - \nu'(n_i(\delta)) \right) \mathbf{1}_{\{v_i(\delta) > \underline{u}\}} \Big| x_i(\delta), z(\delta) \right] &= 0, \\ \log N(\delta) &= \frac{1}{1-\eta} \log \left(\int \exp((1-\eta) \log n_i(\delta)) \right) \end{aligned}$$

where $\log v_i(\delta) = \log(\max\{c_i(\delta) - \nu(n_i(\delta)), \underline{u}\})$, and the term $\frac{1-\epsilon_i^2}{2z(\delta)}$ comes from explicitly calculate the term $\frac{\partial}{\partial z_i} \varphi(\theta, x_i|z_i)$ in Lemma 1:

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

Assume that the equilibrium can be approximated with the following Taylor expansions:

$$\begin{aligned} \log n_i(\delta) &= \bar{n} + \hat{n}_i\delta + \delta^2(\dots), \\ \log N(\delta) &= \bar{N} + \hat{N}\delta + \delta^2(\dots), \\ \log z(\delta) &= \log \bar{z} + \hat{z}\delta + \delta^2(\dots), \\ \log v_i(\delta) &= \bar{v} + \hat{v}_i\delta + \delta^2(\dots). \end{aligned}$$

0th order expansion

The 0th order expansions comes from evaluating the equilibrium conditions at $\delta \rightarrow 0$:

$$\begin{aligned} \mathbb{E} \left[\frac{1}{1-\gamma} e^{(1-\gamma)\bar{v}} \times \frac{1-\epsilon_i^2}{2\bar{z}} \Big| \bar{z} \right] &= 0, \\ e^{\bar{\theta} + \eta\bar{N} - \eta\bar{n}}(1-\eta) &= e^{\nu\bar{n}} \\ \bar{N} &= \bar{n} \end{aligned}$$

where

$$\bar{v} = \log \left(e^{\bar{\theta} + \eta\bar{N} + (1-\eta)\bar{n}} - \nu(e^{\bar{n}}) \right) = \log \left(\frac{\nu + \eta}{1 + \nu} e^{\bar{\theta} + \bar{n}} \right).$$

Conditions above pin down $\bar{n}, \bar{N}, \bar{v}$, but note that the attention level $\bar{z}$ is not pinned down because

$$\mathbb{E} \left[\frac{1}{1-\gamma} e^{(1-\gamma)\bar{v}} \frac{1-\epsilon_i^2}{2\bar{z}} \middle| \bar{z} \right] \equiv 0, \quad \forall \bar{z}.$$

1st order expansion

Differentiate the equilibrium conditions with respect to δ and evaluate at $\delta \rightarrow 0$ gives:

$$\mathbb{E} \left[e^{(1-\gamma)\bar{v}} \hat{v}_i \frac{1-\epsilon_i^2}{2\bar{z}} - \frac{1}{1-\gamma} e^{(1-\gamma)\bar{v}} \frac{1-\epsilon_i^2}{2\bar{z}} \hat{z} \middle| \bar{z} \right] = 0, \quad (1)$$

$$(\nu + \eta) \hat{n}_i = \mathbb{E} [\hat{\theta} + \eta \hat{N} \mid \hat{x}_i], \quad (2)$$

$$\hat{N} = \int \hat{n}_i \quad (3)$$

where $\hat{x}_i = \theta + \frac{1}{\sqrt{\bar{z}}} \epsilon_i$.

In order to characterize the expansions $\hat{n}_i, \hat{N}, \dots$, write

$$\log n_i(\delta) = \mathbf{n}(x_i(\delta), \delta), \quad \log N(\delta) = \mathbf{N}(\theta(\delta), \delta), \quad \log v_i(\delta) = \mathbf{v}(\theta(\delta), \log n_i(\delta), \delta).$$

Taking derivatives with respect to δ and evaluating at $\delta \rightarrow 0$ gives

$$\hat{n}_i = \mathbf{n}_x \hat{x}_i + \mathbf{n}_\delta, \quad \hat{N} = \mathbf{N}_\theta \hat{\theta} + \mathbf{N}_\delta, \quad \hat{v}_i = \mathbf{v}_\delta + \mathbf{v}_\theta \hat{\theta} + \mathbf{v}_n \hat{n}_i,$$

where $\hat{x}_i = \hat{\theta} + \frac{1}{\sqrt{\bar{z}}} \epsilon_i$.

The coefficients $\mathbf{n}_x, \mathbf{n}_\delta, \mathbf{N}_\delta, \mathbf{N}_\theta, \mathbf{v}_\delta, \mathbf{v}_\theta, \mathbf{v}_n$ can be solved from the expansion of equilibrium conditions in eqs. (1) to (3):

$$\mathbf{n}_x = \mathbf{N}_\theta = \frac{r\lambda(\bar{z})}{1-s\lambda(\bar{z})}, \quad (4)$$

and $\mathbf{v}_\theta = \frac{1+\nu}{\eta+\nu}(1+\eta\mathbf{N}_\theta)$, $\mathbf{v}_n = \mathbf{n}_\delta = \mathbf{N}_\delta = \mathbf{v}_\delta = 0$, where

$$\lambda(\bar{z}) := \frac{\sigma^2}{\sigma^2 + 1/\bar{z}}, \quad r := \frac{1}{\nu + \eta}, \quad s := \frac{\eta}{\nu + \eta}.$$

As discussed in Section 3, parameters r, s respectively capture the responsiveness to exogenous shock and the level of strategic complementarity in the economy. Function $\lambda(\bar{z})$ is the Kalman gain given the zeroth order expansion of attention $\bar{z}$.

Note that $\bar{z}$ is not determined by the first order expansion of attention optimality because

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

In order to pin down $\bar{z}$, expand the optimality condition for attention the second order and evaluate at $\delta \rightarrow 0$:

$$\mathbb{E}\left[\left((1-\gamma)\hat{v}_i^2 + \hat{v}_i\right)e^{(1-\gamma)\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}} \mid \bar{z}\right] - 2\kappa'(\bar{z}) = 0, \quad (5)$$

where

$$\hat{v}_i = \mathbf{v}_{nn}\hat{n}_i^2 + \mathbf{v}_n\hat{n}_i + (\mathbf{v}_{\theta\theta}\hat{\theta}^2 + \mathbf{v}_{\delta\delta} + 2\mathbf{v}_{\theta n}\hat{\theta}\hat{n}_i + 2\mathbf{v}_{\theta\delta}\hat{\theta} + 2\mathbf{v}_{n\delta}\hat{n}_i).$$

From direct calculation, $\mathbf{v}_{nn} = -(1-\eta)(1+\nu)$ and $\mathbf{v}_n = 0$ as before. The term contained in the parenthesis above is orthogonal to $1-\epsilon_i^2$ because of the normality and independence of $\epsilon_i, \hat{\theta}$. This observation is useful for derivation below, and is stated as the lemma below:

Lemma 6 $\mathbb{E}[(1-\epsilon_i^2)\epsilon_i^k\hat{\theta}^h] = 0, \quad \forall k, h \in \{2m+1 \mid m \in \mathbb{N}\} \text{ or } k = 0.$

Substitute the expression for $\hat{v}_i$ back into eq. (5) and use the expression for $\hat{n}_i$ in eq. (4), we obtain an express that pins down $\bar{z}$:

$$e^{(1-\gamma)\bar{v}}|\mathbf{v}_{nn}|\left(\frac{\mathbf{n}_x}{\bar{z}}\right)^2 - 2\kappa'(\bar{z}) = 0. \quad (6)$$

With linear attention cost $\kappa(z) = \kappa z$, we can solve for $\bar{z}$ explicitly:

$$\bar{z} = \frac{1}{1-s} \left(\left(\frac{r^2 e^{(1-\gamma)\bar{v}} |\mathbf{v}_{nn}|}{2\kappa} \right)^{\frac{1}{2}} - \frac{1}{\sigma^2} \right). \quad (7)$$

2nd order expansion

Below I solve for the second order expansions of action $\hat{n}_i$ and $\hat{N}$, which are jointly determined with the first order expansion of attention $\hat{z}$.

$$0 = \mathbb{E}[-\hat{n}_i + s\hat{N} \mid \hat{x}_i] + 2\frac{d}{d\delta}\mathbb{E}[-\hat{n}_i + r\hat{\theta} + s\hat{N} \mid x_i(\delta)] \Big|_{\delta=0} - \tilde{\gamma} \text{Var}(\theta + \eta\hat{N} \mid \hat{x}_i), \quad (8)$$

$$\hat{N} = \int \hat{n}_i + (1-\eta) \text{Var}(\hat{n}_i \mid \theta), \quad (9)$$

where $\tilde{\gamma} = \frac{2(1+\nu)}{(\eta+\nu)^2}\gamma - 1$, and

$$\frac{d}{d\delta}\mathbb{E}[-\hat{n}_i + r\hat{\theta} + s\hat{N}|x_i(\delta)]\Big|_{\delta=0} = \frac{1}{2}\mathbf{n}_x(1 - \lambda(\bar{z}))\hat{x}_i \hat{z}.$$

The first order expansion of attention $\hat{z}$ is given by the third order expansion of the attention optimality condition:

$$\begin{aligned} & \mathbb{E}\left[\left((1-\gamma)^2\hat{v}_i^3 + 3(1-\gamma)\hat{v}_i\hat{v}_i + \hat{v}_i\right)e^{(1-\gamma)\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big|\bar{z}\right] \\ & + 3\mathbb{E}\left[-\hat{z}\left((1-\gamma)\hat{v}_i^2 + \hat{v}_i\right)e^{(1-\gamma)\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big|\bar{z}\right] - 2\kappa''(\bar{z})\bar{z}\hat{z} = 0 \end{aligned}$$

By lemma 6,

$$\mathbb{E}\left[\frac{1-\epsilon_i^2}{2\bar{z}}\hat{v}_i^3\Big|\bar{z}\right] = 0, \text{ and } \mathbb{E}\left[\frac{1-\epsilon_i^2}{2\bar{z}}\hat{v}_i\hat{v}_i\Big|\bar{z}\right] = 0.$$

Using eq. (5), the third order expansion of attention optimality reduces to

$$\mathbb{E}\left[\hat{v}_i e^{(1-\gamma)\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big|\bar{z}\right] - 6\kappa'(\bar{z})\hat{z} - 2\kappa''(\bar{z})\bar{z}\hat{z} = 0,$$

where

$$\hat{v}_i = \mathbf{v}_n\hat{n}_i + 3\mathbf{v}_{nn}\hat{n}_i\hat{n}_i + \mathbf{v}_{nnn}\hat{n}_i^3 + 3\mathbf{v}_{nn\theta}\hat{\theta}\hat{n}_i^2 + 3\mathbf{v}_{nnN}\hat{N}\hat{n}_i^2 + \dots$$

Substituting $\hat{v}_i$ back, using $\mathbf{v}_n = 0$, and using lemma 6 to reduce terms that are orthogonal to $1 - \epsilon_i^2$, we have:

$$3\mathbb{E}\left[\mathbf{v}_{nn}\hat{n}_i\hat{n}_i e^{(1-\gamma)\bar{v}}\frac{1-\epsilon_i^2}{2\bar{z}}\Big|\bar{z}\right] - 6\kappa'(\bar{z})\hat{z} - 2\kappa''(\bar{z})\bar{z}\hat{z} = 0. \quad (10)$$

Note the second order expansion of labor input and aggregate labor are given by:

$$\hat{n}_i = \mathbf{n}_{xx}\hat{x}_i^2 + 2\mathbf{n}_{x\delta}\hat{x}_i + \mathbf{n}_{\delta\delta} - \mathbf{n}_x\frac{\epsilon_i}{2\sqrt{\bar{z}}}\hat{z}, \quad \hat{N} = \mathbf{N}_{\theta\theta}\hat{\theta}^2 + 2\mathbf{N}_{\theta\delta}\hat{\theta} + \mathbf{N}_{\delta\delta}.$$

Substituting the expression for $\hat{n}_i, \hat{N}$ to eq. (8), and equalizing the coefficients in front of $\hat{x}_i^2$ and $\hat{x}_i$ gives:

$$(1 - s\lambda^2)\mathbf{n}_{xx} = 0, \quad (11)$$

$$2(1 - s\lambda)\mathbf{n}_{x\delta} = \frac{3}{2}\mathbf{n}_x(1 - \lambda(\bar{z}))\hat{z}. \quad (12)$$

Equation (12) implies the LHS of eq. (10) is homogeneous in $\hat{z}$, and therefore

$$\hat{z} = 0, \quad \mathbf{n}_{x\delta} = 0.$$

Moreover, Equation (11) implies $\mathbf{n}_{xx} = 0$. As a result,

$$\hat{Y} = N_{\delta\delta} \in \mathbb{R}, \quad \hat{y}_i = n_{\delta\delta} \in \mathbb{R}.$$

Note that in the static setup studied above, we have $\hat{z} = 0$ for a given $\bar{\theta}$, and the effect of $\bar{\theta}$ on attention and measures of uncertainty is derived below as a comparative statics result on $\bar{z}(\bar{\theta})$. In the dynamic model below, agents' expected productivity changes with the realization of past productivity shocks. The first order expansion of attention $\hat{z}$ (derived from the third order expansion of attention optimality) picks up the effect of past shocks on attention, just as $\bar{z}(\bar{\theta})$ picks up the effect of $\bar{\theta}$ on attention in the static model.

Expansions of the four measures of uncertainty

Consider the four measures of uncertainty in the economy indexed by δ :

$$\begin{aligned} SD(\tilde{Y}(\delta) | \bar{\theta}) &= \left(\mathbb{E} \left[(\tilde{Y}(\delta) - \mathbb{E}[\tilde{Y}(\delta) | \bar{\theta}])^2 | \bar{\theta} \right] \right)^{\frac{1}{2}} \\ Disp(\tilde{y}_i(\delta) | \bar{\theta}) &= \left(\int (\tilde{y}_i(\delta) - \int \tilde{y}_i(\delta))^2 di \right)^{\frac{1}{2}} \\ Disp(\mathbb{E}_{i,\delta}[\tilde{Y}(\delta)] | \bar{\theta}) &= \left(\int (\mathbb{E}_{i,\delta}[\tilde{Y}(\delta)] - \int \mathbb{E}_{i,\delta}[\tilde{Y}(\delta)] di)^2 \right)^{\frac{1}{2}} \\ SD(\tilde{Y}(\delta) | x_i(\delta), \bar{\theta}) &= \left(\mathbb{E}_{i,\delta} \left[(\tilde{Y}(\delta) - \mathbb{E}_{i,\delta}[\tilde{Y}(\delta)])^2 \right] \right)^{\frac{1}{2}}, \quad \forall i \end{aligned}$$

where $\mathbb{E}_{i,\delta}[\cdot] = \mathbb{E}[\cdot | x_i(\delta), \bar{z}(\delta), \bar{\theta}]$.

Since the four measures of uncertainty share a similar structure, let $U(\delta)$ denote the four measures of uncertainty above generically. A measure of uncertainty $U(\delta)$ has the following form:

$$U(\delta) = \left(\int \left(f(\delta) - \int f(\delta) \varphi(\theta, \delta) \right)^2 \varphi(\theta, \delta) \right)^{\frac{1}{2}},$$

where $f(\delta) \in \{\tilde{Y}(\delta), \tilde{y}_i(\delta), \mathbb{E}_{i,\delta}[\tilde{Y}(\delta)]\}$ such that $f(\delta) = \bar{f} + \hat{f}\delta + \frac{1}{2}\hat{\hat{f}}\delta^2 + \dots$, and $\varphi(\theta, \delta)$ is either a probability density over which expectations $\mathbb{E}, \mathbb{E}_{i,\delta}$ are taken over, or simply the an uniform distribution over the cross section $i \in [0, 1]$.

From direct calculation, the 0^{th} , 1^{st} and 2^{nd} order expansion of $U(\delta)$ are given by:

$$\bar{U} = \left(\int \left(\bar{f} - \int \bar{f} \varphi(\theta, 0) \right)^2 \varphi(\theta, 0) \right)^{\frac{1}{2}} = 0, \quad (13)$$

$$\hat{U} = \left(\int \left(\hat{f} - \int \hat{f} \varphi(\theta, 0) \right)^2 \varphi(\theta, 0) \right)^{\frac{1}{2}}, \quad (14)$$

$$\hat{\hat{U}} = \frac{1}{\hat{U}} \int \left(\hat{f} - \int \hat{f} \varphi(\theta, 0) \right) \left(\hat{f} - \int \hat{f} \varphi(\theta, 0) - \frac{d}{d\delta} \int \hat{f} \varphi(\theta, \delta) \Big|_{\delta=0} \right) \varphi(\theta, 0). \quad (15)$$

Because $\hat{z} = 0$, we have

$$\hat{Y}, \quad \hat{y}_i, \quad \frac{d^2}{d\delta^2} \mathbb{E}_{i,\delta}[\tilde{Y}(\delta)] \Big|_{\delta=0} \in \mathbb{R},$$

and

$$\frac{d}{d\delta} \mathbb{E}[\hat{Y}|\bar{\theta}] \Big|_{\delta=0} = \frac{d}{d\delta} \int \hat{y}_i di \Big|_{\delta=0} = \frac{d}{d\delta} \int \left(\frac{d}{d\delta} \mathbb{E}_{i,\delta}[\tilde{Y}(\delta)] \Big|_{\delta=0} \right) di \Big|_{\delta=0} = \frac{d}{d\delta} \mathbb{E}_{i,\delta}[\hat{Y}] \Big|_{\delta=0} = 0.$$

As a result, the term $\hat{f} - \int \hat{f} \varphi(\theta, 0) - \frac{d}{d\delta} \int \hat{f} \varphi(\theta, \delta) \Big|_{\delta=0}$ in eq. (15) equals to zero, and $\hat{\hat{U}} = 0$ for all four measures of uncertainty. The second order approximations of the four measures of uncertainty can simply be calculated from eq. (14):

$$SD(\tilde{Y}|\bar{\theta}) \approx (1 + \Lambda(\bar{z})) \sigma, \quad (16)$$

$$Disp(\tilde{y}_i|\bar{\theta}) \approx \Lambda(\bar{z}) \frac{1}{\sqrt{\bar{z}}}, \quad (17)$$

$$Disp(\mathbb{E}_i[\tilde{Y}]|\bar{\theta}) \approx (1 + \Lambda(\bar{z})) \lambda(\bar{z}) \frac{1}{\sqrt{\bar{z}}}, \quad (18)$$

$$SD(\tilde{Y}|x_i, \bar{\theta}) \approx (1 + \Lambda(\bar{z})) \sqrt{1 - \lambda(\bar{z})} \sigma, \forall i. \quad (19)$$

With the results derived above, Lemma 2-4 follows immediately.

Proof of Lemma 2. In a fixed-attention economy, the equilibrium is characterized by eqs. (2) to (3) with an exogenous signal precision z . The first order expansion is given by

$$\hat{n}_i = \mathbf{n}_x \hat{x}_i, \quad \hat{N} = \mathbf{N}_\theta \hat{\theta},$$

where $\mathbf{n}_x, \mathbf{N}_\theta$ correspond to the function $\Lambda(z)$ in Lemma 2 and is given by eq. (4)

$$\Lambda(z) := \mathbf{n}_x = \mathbf{N}_\theta = \frac{r\lambda(z)}{1 - s\lambda(z)} > 0. \quad (20)$$

The function $\Lambda(z)$ does not depend on $\bar{\theta}$, and it is increasing in $\lambda(z)$, which is an increasing function of z . $\square$

Proof of Lemma 3. In a fixed-attention economy, the second order approximations of the four measures of uncertainty are given by eqs. (16) to (19), with $\bar{z}$ replaced by an exogenous level of attention $\bar{z}$. Since $\Lambda(z)$ does not depend on $\bar{\theta}$, the four measures of uncertainty are constant in $\bar{\theta}$ up to second-order approximation. $\square$

Proof of Lemma 4. From eq. (7), attention $\bar{z}$ is decreasing in $\bar{v}$ if and only if $\gamma > 1$. Because $\bar{v}$ is increase in $\bar{\theta}$, the statement in Lemma 4 follows. $\square$

Proofs of Proposition 1

Given $\gamma > 1$, Lemma 4 implies $\frac{\partial}{\partial \bar{\theta}} \bar{z}(\bar{\theta}) < 0$.

The following lemma characterize how the four uncertainty measures depend on $\bar{z}$:

Lemma 7 *The four uncertainty measures depends on $\bar{z}$ as follows: Aggregate output volatility*

$$SD(\tilde{Y}|\bar{\theta}) \approx (1 + \Lambda(\bar{z}))\sigma$$

is increasing in $\bar{z}$. Cross-sectional output dispersion and forecast dispersion of aggregate output

$$Disp(\tilde{y}_i|\bar{\theta}) \approx \Lambda(\bar{z})\frac{1}{\sqrt{\bar{z}}} \quad \text{and} \quad Disp(\mathbb{E}_i[\tilde{Y}]|\bar{\theta}) \approx (1 + \Lambda(\bar{z}))\lambda(\bar{z})\frac{1}{\sqrt{\bar{z}}}$$

are hump-shaped functions of $\bar{z}$, which equal zero if $\bar{z} = 0$, strictly positive if $\bar{z} \in (0, \infty)$, and go to zero as $\bar{z} \rightarrow \infty$. Subjective uncertainty about aggregate output

$$SD(\tilde{Y}|x_i, \bar{\theta}) \approx (1 + \Lambda(\bar{z}))\sqrt{1 - \lambda(\bar{z})}\sigma$$

is either decreasing or hump-shaped in $\bar{z}$. The measure equals to σ if $\bar{z} = 0$, and goes to 0 as $\bar{z} \rightarrow \infty$. It is a hump-shaped function if and only if $r > \frac{1}{2}$.

Proof of Lemma 7. The fact that aggregate output volatility is increasing in $\bar{z}$ follows directly from that $\Lambda(\bar{z})$ is increasing in $\bar{z}$.

To characterize how the other three measures of uncertainty depends on $\bar{z}$, let $\zeta(\lambda)$ denote the inverse function of $\lambda(z)$, which maps a given level of Kalman gain $\lambda \in [0, 1]$ to signal

precision z :

$$\zeta(\lambda) := \frac{\lambda}{(1-\lambda)\sigma^2}. \quad (21)$$

Let $d_y(\lambda)$ denote the cross-sectional output dispersion as a function of λ :

$$d_y(\lambda) := \Lambda(\zeta(\lambda)) \frac{1}{\sqrt{\zeta(\lambda)}}.$$

The following condition characterizes $d_y(\lambda)$:

$$\frac{\partial}{\partial \lambda} d_y(\lambda) > 0 \iff g_y(\lambda) < s,$$

where

$$g_y(\lambda) := \frac{2\lambda - 1}{\lambda}.$$

Since $g_y(0) = -\infty$, $g_y(1) = 1$ and $g_y(\lambda)$ is increasing, there exists

$$\lambda_{d_y} := \inf \{ \lambda | g_y(\lambda) > s \} \in (0, 1),$$

$d_y(\lambda)$ has a single peak at λ_{d_y} . Moreover, as $s \rightarrow 1$, $\lambda_{d_y} \rightarrow 1$ and $z(\lambda_{d_y}) \rightarrow \infty$.

Similarly, let $f_d(\lambda)$ denote the forecast dispersion as a function of λ :

$$f_d(\lambda) := (1 + \Lambda(\zeta(\lambda))) \lambda \frac{1}{\sqrt{\zeta(\lambda)}}.$$

The following condition characterizes how $f_d(\lambda)$ depends on $\bar{z}$:

$$\frac{\partial}{\partial \lambda} f_d(\lambda) > 0 \iff \begin{cases} 0 < \lambda < \frac{1}{2}, \\ \frac{1}{2} \leq \lambda < 1, \quad g_d(\lambda) < s, \quad h_d(\lambda) < r. \end{cases}$$

where

$$g_d(\lambda) := \frac{4\lambda - 3}{2\lambda^2 - \lambda}, \quad h_d(\lambda) := \frac{(2\lambda - 1)(s\lambda - 1)^2}{\lambda(2s\lambda^2 - (4 + s)\lambda + 3)}.$$

Since $g_d(0) = -\infty$, $g_d(1) = 1$, and $g_d(\lambda)$ is increasing in λ , there exists

$$u_{f_d} := \inf \{ \lambda | g_d(\lambda) > s \} \in (0, 1).$$

From direct calculation, $h_d(\frac{1}{2}) = 0$, $h_d(u_{f_d}) = \infty$ and $\forall \lambda \in (\frac{1}{2}, u_{f_d})$,

$$h'_d(\lambda) > 0 \quad \text{if} \quad h_d(\lambda) > r > s.$$

As a result, there exists

$$\lambda_{f_d} := \inf \{ \lambda \mid h_d(\lambda) > r, \lambda < u_{f_d} \} \in (0, 1),$$

such that $f_d(\lambda)$ has a single peak at λ_{f_d} . Moreover, because

$$h_d(\lambda) > r > s \implies \lambda > \frac{1}{2-s},$$

the infimum $\lambda_{f_d} \rightarrow 1$ as $s \rightarrow 1$.

Finally, let $f_u(\lambda)$ denote the subjective uncertainty as a function of λ :

$$f_u(\lambda) := (1 + \Lambda(\zeta(\lambda)))\sqrt{1-\lambda} \sigma.$$

The following condition characterize how $f_u(\lambda)$ depends on λ :

$$\frac{\partial}{\partial \lambda} f_u(\lambda) > 0 \iff g_u(\lambda) < s, \quad h_u(\lambda) < r,$$

where

$$g_u(\lambda) := \frac{3\lambda - 2}{\lambda^2}, \quad h_u(\lambda) := \frac{(s\lambda - 1)^2}{\lambda(s\lambda - 3) + 2}.$$

Because $g_u(0) = -\infty$, $g_u(1) = 1$, and $g_u(\lambda)$ is increasing in λ , there exists

$$u_{f_u} := \inf \{ \lambda \mid g_u(\lambda) > s \} \in (0, 1).$$

From direct calculation, $h_u(u_{f_u}) = \infty$, and $\forall \lambda \in [0, u_{f_u})$,

$$h'_u(\lambda) > 0 \quad \text{if} \quad h_u(\lambda) > r > s.$$

As a result, there exists

$$\lambda_{f_u} := \inf \{ \lambda \mid h_u(\lambda) > r, \lambda < u_{f_u} \},$$

such that $f_u(\lambda)$ has a single peak at λ_{f_u} .

Moreover, because

$$h_u(\lambda) > r > s \implies \lambda > 2 - \frac{1}{s},$$

the infimum $\lambda_{f_u} \rightarrow 1$ as $s \rightarrow 1$.

Finally, because $h_u(0) = \frac{1}{2}$, we have $r > \frac{1}{2} \iff \lambda_{f_u} > 0$. $\square$

Proof of Proposition 1. Let $\check{\lambda} = \min\{\lambda_{d_y}, \lambda_{f_d}, \lambda_{f_u}\}$, and define $\check{z} := \zeta(\check{\lambda})$, where $\zeta(\lambda)$ is the inverse function of $\lambda(z)$ given by eq. (21). The statement of Proposition 1 follows lemma 7. $\square$

Dynamic model

Below I derive equilibrium conditions for the dynamic economy provided in Lemma 5, and then I derive a system of equations (the expansion of equilibrium conditions) that characterize the first and second order Taylor expansions of the equilibrium objects. The system is then solved by the computational procedure described in Section 5 and Appendix B.

Proofs of Lemma 5

Let S_i^t be the collection of possible history of signals and prices agent i receive up to the beginning of period t , and write s_i^t for typical element of S_i^t

$$s_i^t := \{x_{i,t}^t, \tilde{p}_i^t\}, \quad \forall t \geq 1$$

where $\tilde{p}_{i,t} = \theta_t + \eta \log N_t + \omega_{i,t}$ is a transformation of $p_{i,t}$ that contains the same information. Let $S_i'^t$ be a collection of $s_i'^t$, which denote the history up to period t without the price signal $x_{i,t}$: $s_i'^t := \{s_i^{t-1}, x_{i,t}\}, \forall t \geq 0$.

A strategy (z_i, n_i) is a sequence of mappings $\{z_{i,t}, n_{i,t}\}_{t=0}^\infty$ such that

$$z_{i,t} : S_i^{t-1} \rightarrow \mathbb{R}_+, \quad n_{i,t} : S_i'^t \rightarrow \mathbb{R}_+.$$

Rewrite agents' period payoff as $U(\theta, N, n, \omega) := u(c(\theta, N, n, \omega), n)$, the variance of noise as $\sigma^2(z) := \frac{1}{z}$, and the attention cost $\kappa(z) := \kappa z$.

Proof. A strategy $\{n_t, z_t\}_{t=0}^\infty$ is an equilibrium only if, for all $\tilde{n}, \tilde{z} \in \mathbb{R}_+$ and after any history s_i^t :

$$\begin{aligned} & \sum_{\tau=t}^{\infty} \beta^{s-t} \int U(\theta_\tau, N_\tau, n_\tau(s_i'^\tau), \omega_{i,\tau}) - \beta \kappa(z_{\tau+1}(s_i^\tau)) d\Phi(\theta^\tau, \omega_i^\tau, s_i^\tau | s_i^{t-1}, z^t) - \kappa(z_t(s_i^{t-1})) \\ & \geq \sum_{\tau=t}^{\infty} \beta^{s-t} \int U(\theta_\tau, N_\tau, n_\tau(s_i'^\tau), \omega_{i,\tau}) - \beta \kappa(z_{\tau+1}(s_i^\tau)) d\Phi(\theta^\tau, \omega_i^\tau, s_i^\tau | s_i^{t-1}, \{z^{t-1}, \tilde{z}\}) - \kappa(\tilde{z}) \end{aligned}$$

and

$$\begin{aligned} & \sum_{\tau=t}^{\infty} \beta^{s-t} \int U(\theta_{\tau}, N_{\tau}, n_{\tau}(s'_{i,\tau}), \omega_{i,\tau}) - \beta \kappa(z_{\tau+1}(s_i^{\tau})) d\Phi(\theta^{\tau}, \omega_i^{\tau}, s_i^{\tau} | s_i'^t, z^t) \\ & \geq \sum_{\tau=t}^{\infty} \beta^{s-t} \int U(\theta_{\tau}, N_{\tau}, \tilde{n}, \omega_{i,\tau}) - \beta \kappa(z_{\tau+1}(s_i^{\tau})) d\Phi(\theta^{\tau}, \omega_i^{\tau}, s_i^{\tau} | s_i'^t, z^t) \end{aligned}$$

where

$$N_t = \left(\int \left(e^{\omega_{i,t}} n_t(s_i'^t) \right)^{1-\eta} d\Phi(\omega_{i,t}, s_i'^t | \theta^t, z^t) \right)^{\frac{1}{1-\eta}} \in \sigma(\theta^t).$$

The following two first order condition follows

$$\begin{aligned} & \frac{\partial}{\partial \tilde{z}} \left(\sum_{\tau=t}^{\infty} \beta^{s-t} \int U(\theta_{\tau}, N_{\tau}, n_{i,s}, \omega_{i,\tau}) - \beta \kappa(z_{i,\tau+1}) d\Phi(\theta^{\tau}, \omega_i^{\tau}, s_i^{\tau} | s_i^{t-1}, \{z^{t-1}, \tilde{z}\}) - \kappa(\tilde{z}) \right) \Big|_{\tilde{z}=z_{i,t}} \\ & = 0, \\ & \frac{\partial}{\partial \tilde{n}} \left(\int U(\theta_t, N_t, \tilde{n}, \omega_{i,t}) d\Phi(\theta^t, \omega_i^t | s_i^t, z^t) \right) \Big|_{\tilde{n}=n_{i,t}} = 0. \end{aligned}$$

The F.O.C. with respect to $\tilde{n}$ gives the second condition in Lemma 5.

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

$$\begin{aligned} & \frac{\partial}{\partial \tilde{z}} \varphi(\theta^{\tau}, \omega_i^{\tau}, s_i^{\tau} | s_i^{t-1}, \{z^{t-1}, \tilde{z}\}) \\ & = \frac{\partial}{\partial \tilde{z}} \varphi(h_{i,t+1} | \tilde{z}, s_i^{t-1}, \theta^t, \omega_i^t) \prod_{l=t+1}^{\tau} \varphi(h_{i,l+1} | \tilde{z}_i^l, h_i^l, \theta^l, \omega_i^l) \varphi(\theta^{\tau}, \omega_i^{\tau} | s_i^{t-1}, \{z^{t-1}, \tilde{z}\}) \\ & = \frac{\partial}{\partial \tilde{z}} \varphi(x_{i,t} | \tilde{z}, \theta_t) \prod_{l=t+1}^{\tau} \varphi(x_{i,l} | z_{i,l}, \theta_l) \varphi(\theta^{\tau}, \omega_i^{\tau} | s_i^{t-1}, \{z^{t-1}, \tilde{z}\}). \end{aligned}$$

Taking derivative and evaluating at $\tilde{z} = z_{i,t}$ gives

$$\frac{\partial}{\partial \tilde{z}} \varphi(x_{i,t} | \tilde{z}, \theta_t) \Big|_{\tilde{z}=z_{i,t}} = \frac{\partial}{\partial \tilde{z}} \phi \left(\frac{x_{i,t} - \theta_t}{\sigma(\tilde{z})} \right) \Big|_{\tilde{z}=z_{i,t}} = \frac{\sigma'(z_{i,t})}{\sigma(z_{i,t})} (\epsilon_{i,t}^2 - 1) \varphi(x_{i,t} | z_{i,t}, \theta_t),$$

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

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

and the F.O.C. for $\tilde{z}$ is given by

$$\sum_{\tau=t}^{\infty} \beta^{s-t} \int \frac{\sigma'(z_{i,t})}{\sigma(z_{i,t})} (\epsilon_{i,t}^2 - 1) \left(U(\theta_{\tau}, N_{\tau}, n_{i,s}, \omega_{i,\tau}) - \beta \kappa(z_{i,\tau+1}) \right) d\Phi(\theta^{\tau}, \omega_i^{\tau}, s_i^{\tau+1} | s_i^{t-1}, z_i^t) = \kappa'(z_{i,t}).$$

□

Characterization of expansion sequence

Similar to the static economy, I approximate the equilibrium of the dynamic economy with perturbation around the deterministic steady state. Specifically, consider a sequence of economy parameterized by δ such that

$$\omega_t(\delta) = \delta \omega_t, \quad \omega_{i,t}(\delta) = \delta \omega_{i,t}, \quad \epsilon_{i,t}(\delta) = \delta \epsilon_{i,t}, \quad \kappa(\delta) = \delta^2 \kappa,$$

The equilibrium “actions” (labor input and aggregate labor) are approximated by Taylor expansion up to the second order:

$$\log n_{i,t}(\delta) = \bar{n} + \hat{n}_{i,t} \delta + \frac{1}{2} \hat{\hat{n}}_{i,t} \delta^2 + \dots, \quad \log N_t(\delta) = \bar{N} + \hat{N}_t \delta + \frac{1}{2} \hat{\hat{N}}_t \delta^2 + \dots,$$

and, as in the static model, the first and second order approximation of actions are accompanied by the zeroth and first order expansion of attention choices:

$$\log z_{i,t}(\delta) = \log \bar{z} + \hat{z}_{i,t} \delta + \dots$$

Taking derivatives of the equilibrium conditions in Lemme 5 with respect to δ to corresponding orders and evaluate at $\delta = 0$ gives the systems that characterize the Taylor expansions of equilibrium objects.

1st order expansion

As in the static problem, expanding the attention optimality condition to the second order, and expanding the optimality condition of labor choice and the aggregation condition to the first order gives:

Lemma 8

$$\begin{aligned} \sum_{\tau=t}^{\infty} \beta^{s-t} \mathbb{E} \left[\hat{U}_{i,\tau} \times \frac{1}{2\bar{z}} (1 - \epsilon_{i,t}^2) \middle| \bar{\mathcal{F}}_{i,t-1} \right] - 2\kappa &= 0, \\ \mathbb{E} [\hat{U}_{n,i,t} | \bar{\mathcal{F}}'_{i,t}] &= 0, \\ \hat{N}_t &= \int \hat{n}_{i,t}. \end{aligned}$$

where

$$\hat{U}_{n,i,t} = \nabla \bar{U}_n \begin{pmatrix} \theta_t \\ \hat{N}_t \\ \hat{n}_{i,t} \\ \omega_{i,t} \end{pmatrix}, \quad \hat{U}_{i,t} = \nabla \bar{U} \begin{pmatrix} 0 \\ \hat{N}_t \\ \hat{n}_{i,t} \\ 0 \end{pmatrix} + (\theta_t \hat{N}_t \hat{n}_{i,t} \omega_{i,t}) \bar{H}(U) \begin{pmatrix} \theta_t \\ \hat{N}_t \\ \hat{n}_{i,t} \\ \omega_{i,t} \end{pmatrix}.$$

As before, $\bar{U}_n = 0$ and $\hat{N}_t \perp \epsilon_{i,t}$ implies the first order expansions of actions $\hat{n}_{i,t}$ and $\hat{N}_t$ are jointly characterized with the zeroth order expansion of attention $\bar{z}$.

2nd order expansion

In order to characterize the second order expansion of actions, $\hat{n}_{i,t}$ and $\hat{N}_t$, and the first order expansion of attention $\hat{z}_{i,t}$, expand the optimality condition of attention to the third order and the other two conditions to the second order:

Lemma 9

$$\begin{aligned} \sum_{\tau=t}^{\infty} \beta^{\tau-t} \mathbb{E} \left[U_{i,\tau}^{(3)} \times \frac{1}{2\bar{z}} (1 - \epsilon_{i,t}^2) \middle| \bar{\mathcal{F}}_{i,t-1} \right] - 6\kappa \hat{z}_{i,t} &= 0, \\ \mathbb{E} [\hat{U}_{n,i,t} | \bar{\mathcal{F}}'_{i,t}] + 2 \frac{d}{d\delta} \mathbb{E} [\hat{U}_{n,i,t} | \mathcal{F}'_{i,t}(\delta)] \Big|_{\delta=0} &= 0, \\ \hat{N}_t = (1 - \eta) \int \left(\hat{n}_{i,t} - \int \hat{n}_{i,t} \right)^2 + \int \hat{n}_{i,t}, \end{aligned}$$

where direct calculation gives

$$\begin{aligned} &\mathbb{E} [U_{i,\tau}^{(3)} (1 - \epsilon_{i,t}^2) \mid \bar{\mathcal{F}}_{i,t-1}] \\ &= \mathbb{E} [(3\hat{U}_{n,i,t} \hat{n}_{i,\tau} + \hat{n}_{i,\tau}^2 (3\bar{U}_{nn\theta} \theta_\tau + 3\bar{U}_{nnN} \hat{N}_\tau + \bar{U}_{nnn} \hat{n}_{i,\tau})) (1 - \epsilon_{i,t}^2) \mid \bar{\mathcal{F}}_{i,t-1}]. \end{aligned}$$

Therefore, the second order expansions of actions $\hat{n}_{i,t}$ and $\hat{N}_t$ are jointly characterized with the first order expansion of attention $\hat{z}_{i,t}$ by the system above.

Using the notation defined above, $s_i^t = \{x_i^t, p_i^{*t-1}\}$, and let $\phi(s'_{i,t}(\delta) | \theta^t, \delta)$ denote the density function of $s'_{i,t}$ conditional on θ^t , we have

Lemma 10 Consider a generic random variable ξ_t such that $\xi_t \in \sigma(\theta^t)$, then

$$\frac{d}{d\delta} \mathbb{E} [\xi_t \mid \mathcal{F}'_{i,t}(\delta)] \Big|_{\delta=0} = \sum_{\tau=0}^t Cov \left(\xi_t, \frac{\frac{d}{d\delta} \phi(s'_{i,t-\tau}(\delta) | \theta^{t-\tau}, \delta) \Big|_{\delta=0}}{\phi(\bar{s}'_{i,t-\tau} | \theta^{t-\tau}, 0)} \mid \bar{\mathcal{F}}'_{i,t} \right)$$

where $\phi(s'_{i,t}(\delta)|\theta^t, \delta)$ denote the conditional density of signals $s'_{i,t}(\delta)$ conditional on θ^t in the δ -economy.

Proof. Rewrite the expectation as integral over probability density, we have

$$\frac{d}{d\delta} \mathbb{E}[\xi_t | \mathcal{F}'_{i,t}(\delta)] = \int_{\theta^t} \xi(\theta^t) \frac{d}{d\delta} \phi(\theta^t | s'^t_{i,t}(\delta), \delta). \quad (22)$$

Bayes rule implies

$$\phi(\theta^t | s'^t_{i,t}(\delta), \delta) = \frac{\phi(s'^t_{i,t}(\delta) | \theta^t, \delta) \phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)}{\int_{\tilde{\theta}^t} \phi(s'^t_{i,t}(\delta) | \tilde{\theta}^t, \delta) \phi(\tilde{\theta}^t | s'^{t-1}_{i,t}(\delta), \delta)}.$$

Taking derivative with respect to δ , we have

$$\begin{aligned} \frac{\frac{d}{d\delta} \phi(\theta^t | s'^t_{i,t}(\delta), \delta)}{\phi(\theta^t | s'^t_{i,t}(\delta), \delta)} &= \frac{\frac{d}{d\delta} \phi(s'_{i,t}(\delta) | \theta^t, \delta)}{\phi(s'_{i,t}(\delta) | \theta^t, \delta)} - \mathbb{E} \left[\frac{\frac{d}{d\delta} \phi(s'_{i,t}(\delta) | \theta^t, \delta)}{\phi(s'_{i,t}(\delta) | \theta^t, \delta)} \middle| \mathcal{F}'_{i,t}(\delta) \right] \\ &+ \frac{\frac{d}{d\delta} \phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)}{\phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)} - \mathbb{E} \left[\frac{\frac{d}{d\delta} \phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)}{\phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)} \middle| \mathcal{F}'_{i,t}(\delta) \right]. \end{aligned}$$

Therefore, eq. (22) can be written as

$$\frac{d}{d\delta} \mathbb{E}[\xi_t | \mathcal{F}'_{i,t}(\delta)] = Cov \left(\xi_t, \frac{\frac{d}{d\delta} \phi(s'_{i,t}(\delta) | \theta^t, \delta)}{\phi(s'_{i,t}(\delta) | \theta^t, \delta)} \middle| \mathcal{F}'_{i,t}(\delta) \right) + Cov \left(\xi_t, \frac{\frac{d}{d\delta} \phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)}{\phi(\theta^t | s'^{t-1}_{i,t}(\delta), \delta)} \middle| \mathcal{F}'_{i,t}(\delta) \right).$$

Iterate backward, we have

$$\frac{d}{d\delta} \mathbb{E}[\xi_t | \mathcal{F}'_{i,t}(\delta)] = \sum_{\tau=0}^t Cov \left(\xi_t, \frac{\frac{d}{d\delta} \phi(s'_{i,t-\tau}(\delta) | \theta^t, \delta)}{\phi(s'_{i,t-\tau}(\delta) | \theta^t, \delta)} \middle| \mathcal{F}'_{i,t}(\delta) \right).$$

Finally, because $\theta^t_{t-\tau+1} \perp \bar{s}'_{i,t-\tau}(\delta) |_{\theta^{t-\tau}}$, we have

$$\phi(\bar{s}'_{i,t-\tau}(\delta) | \theta^t, \delta) = \phi(\bar{s}'_{i,t-\tau}(\delta) | \theta^{t-\tau}, \delta),$$

and the lemma follows from evaluating the expression at $\delta = 0$. $\square$

It is useful to clarify the expression in lemma 10. Note that $s'_{i,t-\tau}(\delta), \tau \geq 0$ are signals in the information set $\mathcal{F}'_{i,t}(\delta)$ that agent i condition on when formation expectations, whereas θ^t is a running variable integrated over the probability density function. Therefore, give

any path of realizations of shocks $\{\theta_{t-\tau}^*, \epsilon_{i,t-\tau}^*, \omega_{i,t-\tau}^*\}$, we have

$$s'_{i,t-\tau}^*(\delta) = \begin{pmatrix} \delta\theta_{t-\tau}^* + \frac{\delta\epsilon_{i,t-\tau}^*}{\sqrt{z_{i,t-\tau}^*(\delta)}} \\ \delta\theta_{t-\tau-1}^* + \eta \log N(\delta\theta_{t-\tau-1}^*, \delta) + \delta\omega_{i,t-\tau-1}^* \end{pmatrix}$$

and

$$\begin{aligned} \left. \frac{d}{d\delta} \phi(s'_{i,t-\tau}^*(\delta) | \theta^t, \delta) \right|_{\delta=0} &= \frac{-\bar{z}}{2} (x_{t-\tau}^* - \theta_{t-\tau}) (\theta_{t-\tau}^* - \theta_{t-\tau}) \hat{z}_{i,t-\tau}^* \\ &\quad + \frac{-\eta}{2\sigma_{\omega_i}^2} (p_{t-\tau-1}^* - (\theta_{t-\tau-1} + \eta \hat{N}_{t-\tau-1})) (\hat{N}_{t-\tau-1}^* - \hat{N}_{t-\tau-1}). \end{aligned}$$

B. Computational method

Taylor expansion of the equilibrium objects $\log n_{i,t}$ and $\log z_{i,t}$ gives the following:

$$\begin{aligned} \hat{n}_{i,t} &= \mathbf{n}_s \hat{s}_i^t + \mathbf{n}_\delta, \quad \bar{z}_{i,t} = \bar{z}, \\ \hat{n}_{i,t} &= \hat{s}_i^{t\top} \mathbf{n}_{ss} \hat{s}_i^t + 2\mathbf{n}_{s\delta} \hat{s}_i^t + \mathbf{n}_{\delta\delta} + \mathbf{n}_s \hat{s}_{i,t}, \quad \hat{z}_{i,t} = \mathbf{z}_s \hat{s}_i^t \end{aligned}$$

Because derivatives $\mathbf{n}_s, \mathbf{n}_{ss}$, etc, are multilinear, spectral theory motivates the search for a finite dimensional approximation with linear factors:

$$\begin{aligned} \tilde{n}_{i,t}^{(1)} &= \Phi_f^n f_{i,t}^{(1)} + \Phi_\delta^n, \quad \tilde{z}_{i,t}^{(0)} = \Phi^z, \\ \tilde{n}_{i,t}^{(2)} &= f_{i,t}^{(2)\top} \Phi_{ff}^n f_{i,t}^{(2)} + 2\Phi_{f\delta}^n f_{i,t}^{(2)} + \Phi_{\delta\delta}^n + \Phi_f^n f_{i,t}^{(1,1)}, \quad \tilde{z}_{i,t}^{(1)} = \Phi_f^z f_{i,t}^{(2)} + \Phi_\delta^z \end{aligned}$$

where Φ_f^n, Φ_{ff}^n , etc, are finite dimension multilinear functions and $f^{(1)}, f^{(2)}, f^{(1,1)}$ are linear factors give by

$$\begin{aligned} f_{i,t+1}^{(1)} &= A^{(1)} f_{i,t}^{(1)} + C^{(1)} \hat{s}_{i,t}, \\ f_{i,t+1}^{(2)} &= A^{(2)} f_{i,t}^{(2)} + C^{(2)} \hat{s}_{i,t}, \quad f_{i,t+1}^{(1,1)} = A^{(1)} f_{i,t}^{(1,1)} + C^{(1)} \hat{s}_{i,t} \end{aligned}$$

for some $A^{(1)}, C^{(1)}$ and $A^{(2)}, C^{(2)}$. The goal of the computational procedure is to solve for multilinear functions Φ 's and factor dynamics A 's and C 's that minimizes the expansion of equilibrium conditions characterized in Lemma 8 and Lemma 9.

Using the aggregation conditions

$$N_t^{(1)} = \int n_{i,t}^{(1)}, \quad N_t^{(2)} = \int n_{i,t}^{(2)} + (1 - \eta) \int \left(n_{i,t}^{(1)} - \int n_{i,t}^{(1)} \right)^2,$$

the optimality conditions in Lemma 8 and Lemma 9, respectively for the first and second order, can be expressed as

$$\begin{aligned} 0 &= \mathbb{E}[G^{(1)}(\tilde{n}_{i,t}^{(1)}, \tilde{z}_{i,t}^{(0)}) | \bar{\mathcal{F}}_{i,t}], \\ 0 &= \mathbb{E}[G^{(2)}(\tilde{n}_{i,t}^{(2)}, \tilde{z}_{i,t}^{(1)}, \tilde{n}_{i,t}^{(1)}, \tilde{z}_{i,t}^{(0)}) | \bar{\mathcal{F}}_{i,t}], \end{aligned}$$

where $G^{(1)}(\cdot)$, $G^{(2)}(\cdot)$ are multilinear functions that depends on partial derivatives of agents' preferences and technology, evaluated at the steady state.

Computational procedure

I solve for $\Phi^{(m)}$ and $A^{(m)}, C^{(m)}$ successively for each order of approximation m . Specifically, given the $(m-1)^{th}$ order approximation $\{\tilde{n}_{i,t}^{(k)}, \tilde{z}_{i,t}^{(k-1)}\}, \forall k \leq m-1$, consider the following two-step procedure:

Outer loop:

- (i) Fix a dimension n for the factors $f_{i,t}^{(m)}$, and specify matrices

$$A^{(m)} \in \mathbb{R}^{n \times n}, C^{(m)} \in \mathbb{R}^{n \times |s|}$$

such that the process $f_{i,t}^{(m)}$ is stable.

- (ii) Given factor specifications $A^{(m)}, C^{(m)}$, solve for coefficients $\Phi^{*(m)}$'s that minimize sum of squared residuals of equilibrium conditions, $R(A^{(m)}, C^{(m)}, \Phi^{(m)})$, with the procedure described below in the inner loop.
- (iii) Solve for $A^{(m)}, C^{(m)}$ that minimize $R(A^{(m)}, C^{(m)}, \Phi^{*(m)})$.

Inner loop:

- (i) For any given $A^{(m)}, C^{(m)}$, simulate $\{\omega_t, \epsilon_{i,t}\}_{t=0}^T$. Construct factors $f_{i,t}^{(m)}$:

$$f_{i,t+1}^{(m)} = A^{(m)} f_{i,t}^{(m)} + C^{(m)} \bar{s}_{i,t}, \quad f_{i,0}^{(m)} = 0.$$

- (ii) Find multilinear functions $\Phi^{(m)}$'s to minimize the sum of squares residuals

$$R(A^{(m)}, C^{(m)}, \Phi^{(m)}) := \sum_{t=0}^T residual_{i,t}^{(m)},$$

where

$$residual_{i,t}^{(m)} = \mathbb{E}[G^{(m)}(\tilde{n}_{i,t}^{(m)}, \tilde{z}_{i,t}^{(m-1)}, \{\tilde{n}_{i,t}^{(k)}, \tilde{z}_{i,t}^{(k-1)}\}_{k \leq m-1}) | \bar{\mathcal{F}}_{i,t}],$$

and

$$\tilde{n}_{i,t}^{(2)} = \Phi_n^{(m)}(f_{i,t}^{(m)}, \{f_{i,t}^{(k,l)}\}_{k+l=m}), \quad \Phi_z^{(m-1)}(f_{i,t}^{(m-1)}, \{f_{i,t}^{(k,l)}\}_{k+l=m-1}).$$

References

- [1] Tobias Adrian et al. “Vulnerable growth”. In: *American Economic Review* 109.4 (2019), pp. 1263–89.
- [2] SeHyoun Ahn et al. “When inequality matters for macro and macro matters for inequality”. In: *NBER macroeconomics annual* 32.1 (2018), pp. 1–75.
- [3] G-M Angeletos and Chen Lian. “Incomplete information in macroeconomics: Accommodating frictions in coordination”. In: *Handbook of macroeconomics*. Vol. 2. Elsevier, 2016, pp. 1065–1240.
- [4] George-Marios Angeletos and Zhen Huo. “Myopia and anchoring”. In: *American Economic Review* 111.4 (2021), pp. 1166–1200.
- [5] George-Marios Angeletos and Jennifer La’O. “Noisy business cycles”. In: *NBER Macroeconomics Annual* 24.1 (2010), pp. 319–378.
- [6] George-Marios Angeletos and Jennifer La’O. “Sentiments”. In: *Econometrica* 81.2 (2013), pp. 739–779.
- [7] George-Marios Angeletos and Chen Lian. “Forward guidance without common knowledge”. In: *American Economic Review* 108.9 (2018), pp. 2477–2512.
- [8] George-Marios Angeletos and Karthik Sastry. *Inattentive Economies*. Tech. rep. National Bureau of Economic Research, 2019.
- [9] George-Marios Angeletos et al. “Quantifying confidence”. In: *Econometrica* 86.5 (2018), pp. 1689–1726.
- [10] Rüdiger Bachmann et al. “Uncertainty and economic activity: Evidence from business survey data”. In: *American Economic Journal: Macroeconomics* 5.2 (2013), pp. 217–49.
- [11] Susanto Basu and Brent Bundick. “Uncertainty shocks in a model of effective demand”. In: *Econometrica* 85.3 (2017), pp. 937–958.
- [12] Azi Ben-Rephael et al. “It depends on where you search: Institutional investor attention and underreaction to news”. In: *The Review of Financial Studies* 30.9 (2017), pp. 3009–3047.

- [13] Jess Benhabib et al. “Endogenous information acquisition and countercyclical uncertainty”. In: *Journal of Economic Theory* 165 (2016), pp. 601–642.
- [14] Anmol Bhandari et al. *Inequality, business cycles, and monetary-fiscal policy*. Tech. rep. National Bureau of Economic Research, 2018.
- [15] Nicholas Bloom. “Fluctuations in uncertainty”. In: *Journal of Economic Perspectives* 28.2 (2014), pp. 153–76.
- [16] Nicholas Bloom. “The impact of uncertainty shocks”. In: *econometrica* 77.3 (2009), pp. 623–685.
- [17] Nicholas Bloom et al. “Really uncertain business cycles”. In: *Econometrica* 86.3 (2018), pp. 1031–1065.
- [18] Jaroslav Borovicka and Lars Peter Hansen. “Robust Preference Expansions”. In: (2013).
- [19] Lawrence J Christiano et al. “Risk shocks”. In: *American Economic Review* 104.1 (2014), pp. 27–65.
- [20] 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.
- [21] Zhi Da et al. “In search of attention”. In: *The Journal of Finance* 66.5 (2011), pp. 1461–1499.
- [22] Zhi Da et al. “The sum of all FEARS investor sentiment and asset prices”. In: *The Review of Financial Studies* 28.1 (2015), pp. 1–32.
- [23] Joseph Engelberg et al. “Comparing the point predictions and subjective probability distributions of professional forecasters”. In: *Journal of Business & Economic Statistics* 27.1 (2009), pp. 30–41.
- [24] Larry G Epstein and Stanley E Zin. “Substitution, Risk Aversion, and the Temporal Behavior of Consumption”. In: *Econometrica* 57.4 (1989), pp. 937–969.
- [25] Pablo D Fajgelbaum et al. “Uncertainty traps”. In: *The Quarterly Journal of Economics* 132.4 (2017), pp. 1641–1692.
- [26] John Fernald. “A quarterly, utilization-adjusted series on total factor productivity”. In: Federal Reserve Bank of San Francisco. 2014.
- [27] Jesús Fernández-Villaverde et al. *Financial frictions and the wealth distribution*. Tech. rep. National Bureau of Economic Research, 2019.
- [28] Jesús Fernández-Villaverde et al. “Risk matters: The real effects of volatility shocks”. In: *American Economic Review* 101.6 (2011), pp. 2530–61.
- [29] Joel P Flynn and Karthik Sastry. “Attention Cycles”. In: *Available at SSRN 3592107* (2021).
- [30] Simon Gilchrist et al. *Uncertainty, financial frictions, and investment dynamics*. Tech. rep. National Bureau of Economic Research, 2014.
- [31] Christian Hellwig and Laura Veldkamp. “Knowing what others know: Coordination motives in information acquisition”. In: *The Review of Economic Studies* 76.1 (2009), pp. 223–251.

- [32] Zhen Huo and Marcelo Pedroni. “A single-judge solution to beauty contests”. In: *American Economic Review* 110.2 (2020), pp. 526–68.
- [33] Zhen Huo and Naoki Takayama. “Rational expectations models with higher order beliefs”. In: *Manuscript, Yale Univ* (2015).
- [34] Cosmin Ilut et al. “Slow to hire, quick to fire: Employment dynamics with asymmetric responses to news”. In: *Journal of Political Economy* 126.5 (2018), pp. 2011–2071.
- [35] Kenneth L Judd. *Numerical methods in economics*. MIT press, 1998.
- [36] Kyle Jurado et al. “Measuring uncertainty”. In: *American Economic Review* 105.3 (2015), pp. 1177–1216.
- [37] Kenneth Kasa. “Forecasting the forecasts of others in the frequency domain”. In: *Review of Economic Dynamics* 3.4 (2000), pp. 726–756.
- [38] Nicholas Kozeniauskas et al. “What are uncertainty shocks?” In: *Journal of Monetary Economics* 100 (2018), pp. 1–15.
- [39] 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.
- [40] Giovanni Lombardo and Harald Uhlig. “A Theory Of Pruning”. In: *International Economic Review* 59.4 (2018), pp. 1825–1836.
- [41] Guido Lorenzoni. “A theory of demand shocks”. In: *American Economic Review* 99.5 (2009), pp. 2050–84.
- [42] Guido Lorenzoni. “Optimal monetary policy with uncertain fundamentals and dispersed information”. In: *The Review of Economic Studies* 77.1 (2010), pp. 305–338.
- [43] Robert E Lucas. “Expectations and the Neutrality of Money”. In: *Journal of economic theory* 4.2 (1972), pp. 103–124.
- [44] Bartosz Maćkowiak and Mirko Wiederholt. “Business cycle dynamics under rational inattention”. In: *The Review of Economic Studies* 82.4 (2015), pp. 1502–1532.
- [45] Bartosz Maćkowiak and Mirko Wiederholt. “Optimal sticky prices under rational inattention”. In: *American Economic Review* 99.3 (2009), pp. 769–803.
- [46] Bartosz Maćkowiak et al. “Dynamic rational inattention: Analytical results”. In: *Journal of Economic Theory* 176 (2018), pp. 650–692.
- [47] Bartosz Maćkowiak et al. “Rational inattention: A disciplined behavioral model”. In: *CEPR Discussion Papers* 13243 (2018).
- [48] Taneli Mäkinen and Björn Ohl. “Information acquisition and learning from prices over the business cycle”. In: *Journal of Economic Theory* 158 (2015), pp. 585–633.
- [49] David P Myatt and Chris Wallace. “Endogenous information acquisition in coordination games”. In: *The Review of Economic Studies* 79.1 (2012), pp. 340–374.
- [50] Kristoffer Nimark. “Dynamic higher order expectations”. In: (2017).
- [51] Kristoffer P Nimark. “Man-bites-dog business cycles”. In: *American Economic Review* 104.8 (2014), pp. 2320–67.
- [52] Anna Orlik and Laura Veldkamp. *Understanding uncertainty shocks and the role of black swans*. Tech. rep. National Bureau of Economic Research, 2014.

- [53] Phelps. *Microeconomic foundations of employment and inflation theory*. WW Norton New York, 1970.
- [54] Richard Rogerson. “Indivisible labor, lotteries and equilibrium”. In: *Journal of monetary Economics* 21.1 (1988), pp. 3–16.
- [55] Christopher A Sims. “Implications of rational inattention”. In: *Journal of monetary Economics* 50.3 (2003), pp. 665–690.
- [56] Christopher A Sims. “Rational inattention and monetary economics”. In: *Handbook of monetary economics*. Vol. 3. Elsevier, 2010, pp. 155–181.
- [57] Robert M Townsend. “Forecasting the forecasts of others”. In: *Journal of Political Economy* 91.4 (1983), pp. 546–588.
- [58] Stijn Van Nieuwerburgh and Laura Veldkamp. “Learning asymmetries in real business cycles”. In: *Journal of monetary Economics* 53.4 (2006), pp. 753–772.
- [59] Joseph Vavra. “Inflation dynamics and time-varying volatility: New evidence and an ss interpretation”. In: *The Quarterly Journal of Economics* 129.1 (2014), pp. 215–258.
- [60] Michael Woodford. *Imperfect common knowledge and the effects of monetary policy*. Tech. rep. National Bureau of Economic Research, 2001.

C. Supplementary Appendix: Extensions

General attention cost

The zeroth order expansion of attention $\bar{z}(\theta)$ is characterized by eq. (6) in Appendix A for a general attention cost $\kappa(z)$:

$$e^{(1-\gamma)\bar{v}} |\mathbf{v}_{nn}| \left(\frac{\mathbf{n}_x}{\bar{z}} \right)^2 - 2\kappa'(\bar{z}) = 0.$$

Write $\bar{v} = \bar{v}(\bar{\theta})$ and $\bar{z} = \bar{z}(\bar{\theta})$, use the expression for $\mathbf{n}_x$, and totally differentiate with respect to $\bar{\theta}$

$$\bar{z}'(\bar{\theta}) = \frac{r^2 \sigma^4 (\sigma^2 \bar{z}(\bar{\theta}) + 1) (1 - \gamma) |\mathbf{v}_{nn}| e^{(1-\gamma)\bar{v}} \bar{v}'(\bar{\theta})}{2 \left(r^2 \sigma^6 e^{(1-\gamma)\bar{v}} |\mathbf{v}_{nn}| + (s\lambda - 1)^2 (\sigma^2 \bar{z}(\bar{\theta}) + 1)^3 \kappa''(\bar{z}(\bar{\theta})) \right)},$$

where $\bar{v}'(\bar{\theta}) > 0$. It follows that if $\kappa''(\cdot) > 0$, then

$$\bar{z}'(\bar{\theta}) < 0 \iff \gamma > 1.$$

Attention choice amplifies exogenous changes in volatility σ

I show below that endogenous attention amplifies the economy's response to exogenous changes in σ . To study the effect of a changes in σ , consider a slightly different formulation of signal structure and attention cost: an agent receives signal

$$x_i = \theta + \epsilon_i, \quad \epsilon_i \stackrel{iid}{\sim} \mathcal{N}(0, \frac{\sigma^2}{z_i}), \quad \text{with attention cost } \kappa z_i.$$

Compared to the specification in the main text, the formulation here makes sure that given by exerting the same level of disutility κz_i , an agent receives the same amount of information, as measured by the ratio of variance reduction (which is the same as entropy reduction given the Gaussian structure). This change does not affect any of the analysis in the main text because it is only a normalization of the cost function when σ^2 is fixed.

With this specification, equilibrium attention and responsiveness of the economy is still given by eq. (7) in Appendix A and eq. (20)

$$\bar{z}(\sigma) = \frac{1}{1-s} \left(\left(\frac{r^2 e^{(1-\gamma)\bar{v}} |\mathbf{v}_{nn}|}{2\kappa} \right)^{\frac{1}{2}} - \frac{1}{\sigma^2} \right), \quad \Lambda(\bar{z}(\sigma)) = \frac{r\lambda(\bar{z}(\sigma))}{1-s\lambda(\bar{z}(\sigma))},$$

except that the formula of Kalman gain $\lambda(z)$ is now given by

$$\lambda(z) = \frac{1}{1 + \frac{1}{z}}.$$

It follows immediately that an increase in σ leads to an increase in attention:

$$\frac{\partial}{\partial \sigma} \bar{z}(\sigma) > 0.$$

Similar to eqs. (16) to (19), the four measures of uncertainty are now approximated by:

$$SD(\tilde{Y}|\bar{\theta}) \approx (1 + \Lambda(\bar{z}))\sigma \quad (23)$$

$$Disp(\tilde{y}_i) \approx \Lambda(\bar{z}) \frac{\sigma}{\sqrt{\bar{z}}} \quad (24)$$

$$Disp(\mathbb{E}[\tilde{Y}|\bar{\theta}, x_i]) \approx (1 + \Lambda(\bar{z}))\lambda \frac{\sigma}{\sqrt{\bar{z}}} \quad (25)$$

$$SD(\tilde{Y}|\bar{\theta}, x_i) \approx (1 + \Lambda(\bar{z}))\sqrt{(1 - \lambda(z))}\sigma. \quad (26)$$

From equation eqs. (23) to (26) in Appendix A, we can see that if attention $\bar{z}$ is fixed, an increase in σ directly leads to proportional increases in the four measures of uncertainty (that is, the elasticities of the four uncertainty measures with respect to σ would be one). With attention responding endogenously, an increase in σ leads to an increase in attention $\bar{z}(\sigma)$, and amplifies changes in the four measures of uncertainty under the conditions characterized in Lemma 7 in Appendix A:

Lemma 11 *If $\bar{z}(\sigma) < \bar{z}$, then*

$$\begin{aligned} \frac{d}{d \log \sigma} \log (1 + \Lambda(\bar{z}))\sigma &> 1, \\ \frac{d}{d \log \sigma} \log \frac{\sigma}{\sqrt{\bar{z}}} &> 1, \\ \frac{d}{d \log \sigma} \log (1 + \Lambda(\bar{z}))\lambda \frac{\sigma}{\sqrt{\bar{z}}} &> 1, \quad \text{and} \\ \frac{d}{d \log \sigma} \log (1 + \Lambda(\bar{z}))\sqrt{(1 - \lambda(z))}\sigma &> 1. \end{aligned}$$

Separating income effect on attention from risk aversion

I show below that income effect on attention and risk aversion can be separated by generalizing the preference in Section 2 in the manner of Epstein and Zin (1989). Under the generalized preference, income effect on attention governs how equilibrium attention to $\bar{\theta}$,

where as risk aversion has no effect on $\bar{z}(\bar{\theta})$.

Consider agents choosing consumption, labor, and attention choice under preference

$$\frac{1}{1-\psi} \mathcal{V}_i^{1-\psi} - \kappa(z_i), \quad \text{where} \quad \mathcal{V}_i = \mathbb{E}[v_i^{1-\gamma}]^{\frac{1}{1-\gamma}}$$

and $v_i = \max\{c_i - \nu(n_i), \underline{u}\}$, $\nu(n_i) = \frac{n_i^{1+\nu}}{1+\nu}$, and $\kappa(z_i) = \kappa z_i$. Parameter γ governs how different realizations of consumption-labor composite v_i translate into its certainty equivalence $\mathcal{V}_i$, whereas ψ governs how agents trade off the certainty equivalence $\mathcal{V}_i$ with attention cost. In particular, when $\psi = \gamma$, the preference here coincide to that in Section 2.

Solving consumption c_i as a function of n_i, N, θ as in section 2, agents' optimization problem is given by:

$$\max_{z_i, n(x_i)} \frac{1}{1-\psi} \mathcal{V}_i^{1-\psi} - \kappa(z_i).$$

Point-wise optimization for each realization of x_i implies first order conditions with respect to $n(x_i)$ that is identical to that in Lemma 1:

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

Similarly, optimality of attention choice is given by:

$$\frac{1}{1-\gamma} \mathcal{V}_i^{\gamma-\psi} \int v_i^{1-\gamma} \frac{\partial}{\partial z_i} \varphi(\theta, x_i | z_i) d\theta dx_i - \kappa'(z_i) = 0.$$

Similar to the derivation in Appendix A, the zeroth order approximation of attention $\bar{z}$ is pinned down by the second order expansion with respect to δ :

$$e^{(\gamma-\psi)\bar{v}} \mathbb{E} \left[\left((1-\gamma)\hat{v}_i^2 + \hat{v}_i \right) e^{(1-\gamma)\bar{v}} \frac{1 - \epsilon_i^2}{2\bar{z}} \middle| \bar{z} \right] - 2\kappa'(\bar{z}) = 0,$$

where $\bar{v}, \hat{v}_i, \hat{v}_i$ are the same as in Appendix A. Using lemma 6 to simplify the expression, we have the following condition that characterize $\bar{z}$:

$$e^{(1-\psi)\bar{v}} |\mathbf{v}_{nn}| \left(\frac{\mathbf{n}_x}{\bar{z}} \right)^2 - 2\kappa'(\bar{z}) = 0.$$

It follows immediately that how $\bar{z}$ changes with $\bar{\theta}$ depends only on ψ and not γ .

To understand the result, note that risk aversion affects agents' payoff through the term $(1-\gamma)\hat{v}_i^2$, but since attention has no first order effect on agent's payoff (i.e., $\mathbf{v}_n = 0$), γ does not affect agents' attention choice up to second order approximation.

D. Supplementary Appendix: Empirical details

Data description

Empirical evidence provided in the main text and the supporting evidence below comes from four data sources.

First, aggregate output, hours, and TFP (with and without utilization-adjustment) series come from John Fernald’s website (<https://www.johnfernald.net/TFP>), where series are adjusted as described in Fernald (2014).

Second, firm level revenue data is constructed from Compustat (Capital IQ-North America). The data set provides standardized balance sheet information for over 80,000 publicly traded companies at quarterly frequency.

Third, the U.S. forecast survey data comes from the SPF (<https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters>). SPF is a survey of professional forecasters currently run by the Federal Reserve Bank of Philadelphia. SPF conducts quarterly survey around the end of the second month in the quarter with around 40 forecasters. It provides forecaster-level data, in which forecasters report forecasts for outcomes in the current and next four quarters, typically about the level of the variable in each quarter. Besides asking forecasters for point estimates of economic variables, SPF also asks forecasters to report probabilistic forecasts for fixed-event year-over-year percentage changes in the GDP growth (that is, in each quarter, the SPF asks forecasters to provide forecasts about the GDP growth between two years). 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.

Finally, Google Trend data set (<https://trends.google.com/>) provides query share of any group of terms relative to the total amount of queries on Google in a period of time, with the maximum query share in that period normalized to a hundred. Query share of search terms are assigned into different categories using a natural-language-processing algorithm by Google. The categories include Art & Entertainment, Business & Industry, Food, Travel, etc.

Sample selection and variable construction

I describe below the construction of the four measures of uncertainty reported in Figure 1, Table 2, and Table 3.

1. Aggregate output volatility, $\widetilde{SD}_t(\Delta\widetilde{Y}_t)$, is measured as the conditional heteroskedasticity of quarterly real GDP growth with a univariate EGARCH(1,1)-ARMA(2,2) model.
2. Cross-sectional output dispersion $\widetilde{Disp}_t(\Delta\widetilde{y}_{i,t})$, is calculated using the Compustat

data. I restrict the sample to U.S. firms, excluding firms in the financial and utility sector because these sectors are heavily regulated. I also exclude firms (1) with negative sales, (2) with acquisitions of size larger than 5% of its asset value, and (3) of history less than 10 quarters. To construct dispersion of revenue-growth within each sector, I use four-digit SIC code to define a sector, and exclude sectors with less than 5 firms and sectors that appear in less than 10 quarters. I construct quarterly real revenue growth for each firm in the sample as:

$$\Delta \tilde{y}_{i,t} = \frac{\tilde{y}_{i,t} - \tilde{y}_{i,t-1}}{(\tilde{y}_{i,t} + \tilde{y}_{i,t-1})/2},$$

where $\tilde{y}_{i,t}$ denotes the real revenue of firm i at time t , and real revenue is deflated by the GDP deflator from NIPA. I calculate the standard deviation of revenue growth across firms within each sector with 90% winsorization. I control for quarter-of-year fixed effect for each sector to adjust for seasonality, and calculate $\widetilde{Disp}_t(\Delta \tilde{y}_{i,t})$ as the average standard deviation of revenue growth, where the average is taken across sectors and weighted by the total revenue of each sector in period t .

3. Forecast dispersion about aggregate output, $\widetilde{Disp}_t(f_{i,t}^{\Delta \tilde{Y}_t})$, is calculated using the SPF data from 1969Q1 to 2020Q2. For each time period t and forecaster i , I calculate the forecast of real GDP growth as

$$f_{i,t}^{\Delta \tilde{Y}_t} = \log(f_{i,t}^{\tilde{Y}_{t+1}} / f_{i,t}^{\tilde{Y}_t}),$$

where $f_{i,t}^{\tilde{Y}_s}$ denotes forecaster i 's time t forecast of real GDP for time period s . The forecast dispersion, $\widetilde{Disp}_t(f_{i,t}^{\Delta \tilde{Y}_t})$, is calculated as the standard deviation of $f_{i,t}^{\Delta \tilde{Y}_t}$ for each period across forecasters with 90% winsorization.

4. Subjective uncertainty about aggregate output, $\widetilde{SD}_{i,t}(\Delta \tilde{Y}_t)$, is calculated using the SPF probability-range data from 1992Q1 to 2020Q2.¹¹ Although the data exists for a longer time period, the SPF made major changes to the real GDP growth intervals in the survey before 1992Q1 and after 2020Q2. I restrict the sample to the periods within these two dates in order to avoid being affected by changes in the conduct of the survey. I fit a Beta distribution with parameter a, b and support $[l, r]$ to the response of each forecaster in each period. Specifically, let $\{t_k\}_{k=1}^n$ denote the end points of intervals specified by the SPF where $t_1 = -\infty$ and $t_n = \infty$, and $F_{i,t}(t_l)$ denote the empirical CDF provided by forecaster i . I look for parameters $a_{i,t}, b_{i,t}$

¹¹The data exists since 1968Q4, however, the Federal Reserve Bank of Philadelphia took over the survey only since 1992Q1. I use data starting from 1992Q1 for data quality reason.

and bounds $l_{i,t}, r_{i,t}$ that solves

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

such that

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

where l_{min} and r_{max} some bounds on the support. In other words, if a forecaster put 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 assume that l_{min} (r_{max}) are 3% below (above) the upper (lower) bound of the end intervals specified by the SPF. 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 , and construct a time series from taking averages of the standard deviations in each period t . Finally, since the survey asks forecasters to predict the same year-over-year GDP growth in different quarters throughout the year, the timing of survey introduce seasonal variations in the average of the standard deviations of the fitted distributions. To control for the effect due to the timing of survey, I regress the series on quarter-of-the-year dummies, and take the residuals as the measure of the measure of subjective uncertainty.

I now describe the construction of two measures of attention using the Google Trend data for Figure 2 and Figure 3, and a measure of attention using the SPF for Table 4.

5. For the first proxy of attention to macroeconomic events, I calculate the Google search share of 30 major U.S. media, such as CNN, Fox News, etc, as given by the Pew Research Center.¹² The data is available at monthly frequency since 2004. To focus on searches related to economic issues, I restrict the sample to searches under the “Business and Industrial” category: some examples of queries that contains the term “CNN” are “CNN Dow Jones,” “CNN premarket,” and more recently, “CNN coronavirus” and “CNN stimulus check.” I sum the query shares of all 30 media and construct the measure as the log of total query shares, controlling for month-of-year fixed effect.

6. For the second proxy of attention to macroeconomic events, I calculate the Google

¹²The list can be found at <https://www.pewresearch.org/fact-tank/2020/01/24/qa-how-pew-research-center-evaluated-americans-trust-in-30-news-sources/>

search share of a group of words classified as “economic words” in the textual-analytics literature,¹³ including terms such as “bank,” “unemployment,” and “gold.” I sum the query shares of all words on the list and construct the measure as the log of total query shares, controlling for month-of-year fixed effect.

7. To estimate the coefficient β_{CG} as a measure of attention, I use the SPF data from 1969Q1 to 2020Q2 to construct forecast error and forecast revision with horizon 1 to 4 quarters in each period for each forecaster as described in Section 7.

Supporting evidence: Countercyclical attention

Figure 2 shows the search share of major news media in the Business and Industrial category. Searches for news increases by 100% during the past two recessions — the pandemic crisis in 2020 and financial crisis in 2008. This finding indicates that during recessions, people endogenous choose to shift their focus to economic issues that concern the economy as a whole and is consistent with the model implication that attention is countercyclical.

Figure 2: Google search share of 30 major U.S. media in the Business and Industrial category increases by 100% during the past two recessions. x-axis: time; y-axis: log Google search share, demean.

Figure 3 shows the Google search share of the group of words classified as “economic words” in the textual-analytics literature. This measure reflect household’s concern about economic issues more broadly, which reflects both their information demand about macroeconomic condition and information demand for idiosyncratic purpose. The measure is countercyclical and suggests that the demand for economic information is higher during recessions.

¹³See Da et al. (2015) for a description of the construction of the list of words.

Figure 3: x-axis: time; y-axis: log-deviation of Google search share of “economic words” constructed following Da et al.(2015); gray area indicates NBER recessions.

Supporting evidence: TFP volatility with utilization adjustment

I construct aggregate TFP volatility as the conditional heteroskedasticity of quarterly TFP growth with a univariate EGARCH(1,1)-ARMA(2,2) model, and similarly for utilization-adjusted TFP series.

Figure 4 plots the conditional volatility of TFP with and without utilization adjustment. Both series are estimated conditional volatility from a EGARCH(1,1)-ARMA(2,2) model, except that the red line comes from the unadjusted TFP series and the green line comes from the utilization-adjusted TFP, both from Fernald (2014)’s data set.

Figure 4: Conditional heteroskedasticity of TFP estimated. Red line: TFP without utilization adjustment; green line: TFP with utilization adjustment. Both series are from Fernald (2014)'s data set., estimated with EGARCH(1,1)-ARMA(2,2) and robust to different GARCH specifications. x-axis: time; y-axis: log-deviation of variables from the long-run average; gray area indicates NBER recessions.

When both series are band-pass filtered at business cycle frequency, the conditional volatility of TFP without adjust has a correlation of -0.61 with output. On the other hand, the utilization-adjusted series has a correlation of -0.06 with output.