THE UNIVERSITY OF CHICAGO

ESSAYS ON MACROECONOMIC MODELS WITH DISPERSED INFORMATION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS
AND
THE FACULTY OF THE UNIVERSITY OF CHICAGO
BOOTH SCHOOL OF BUSINESS
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
YU-TING CHIANG

CHICAGO, ILLINOIS
JUNE 2021

獻給我的母親孫慧雅、父親江逸群與妹妹江宇珊

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

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 my friends Kai-Hao Yang, Piotr Żoch, Mikayel Sukiasyan, Ezra Karger, and Hyunmin Park. All errors are my own.

ABSTRACT

I study a dispersed information economy in which agents choose how much attention to pay to macroeconomic events. I show that under certain conditions, attention and four widely studied measures of uncertainty are countercyclical: agents pay attention when they expect the economy to be in a bad state, and this increase in attention leads to higher (i) conditional volatility of aggregate output, (ii) dispersion of individual output, (iii) forecast dispersion about aggregate output, and (iv) forecast uncertainty about aggregate output. As agents pay attention, they react more to an event and their aggregate response generates high volatility. Because information is dispersed, agents' beliefs and reactions diverge and each agent faces higher uncertainty about others' aggregate response. All these implications are consistent with data. I evaluate the mechanism quantitatively in a dynamic dispersed information economy calibrated to U.S. forecast-survey data. In the calibrated economy, countercyclical attention generates countercyclical fluctuations in all four measures of uncertainty with cyclicality, magnitude, and persistence consistent with untargeted moments in the data. The analysis of the dynamic dispersed information economy requires a new solution method. Due to dispersed information, the economy features an "infinite regress problem" under which the equilibrium lacks a finite recursive state space. Existing methods addressing the problem are constrained to first-order approximations. These methods cannot capture attention and uncertainty fluctuations because these fluctuations are higher-order dynamics of the model. I develop a higher-order approximation method for dispersed information economies based on perturbation techniques to capture higher-order dynamics.

CHAPTER 1

STRATEGIC UNCERTAINTY OVER BUSINESS CYCLES

1.1 Introduction

Economic agents, such as firms and households, face uncertainty about macroeconomic events. Under uncertainty, agents form beliefs and make decisions based on information they receive about the economy. Yet, the information agents receive is often *dispersed* in the sense that no one person in the economy possesses exactly the same information as another person. A long tradition in macroeconomics has acknowledged this friction, starting from Phelps (1970) and Lucas (1972); see Angeletos and Lian (2016) for a recent review.

This paper studies a dispersed information economy in which agents observe unknown macroeconomic events with idiosyncratic noise, and make production decisions based on their individual observations. Agents can exert costly effort to reduce the noise in their information. I call this effort their *attention*, following Sims (2003, 2010). The key feature of the model is that depending on aggregate condition, agents can endogenously pay different levels of attention to the economy over the business cycle.

The main result of the paper is that, under certain conditions, agents' attention and *four measures of uncertainty* are countercyclical: agents pay attention when they expect the economy to be in a bad state, and this increase in attention *alone* increases

1. the volatility of aggregate output,
2. the cross-sectional dispersion of individual output,
3. the forecast dispersion about aggregate output, and
4. the forecast uncertainty about aggregate output (i.e., forecast errors expected by each agent).

The first measure, volatility of aggregate output, captures the size of movements in aggregate production. The second measure, cross-sectional dispersion of individual output, captures the difference in outcomes among agents. The third measure, forecast dispersion about aggregate output, captures the different views held by agents about the economy. And the fourth measure, forecast uncertainty about aggregate output, captures the subjective uncertainty faced by each agent in the economy based on their information.

The mechanism behind these phenomena is as follows. First, an income effect on attention drives countercyclical attention: when agents expect the economy to be in a bad state, they pay more attention to avoid bad decisions, because each unit of resource becomes more valuable when expected income is low. As agents pay attention, they react more to the unknown state, and their aggregate response increases aggregate volatility. Because information is dispersed, each agent updates their belief with different signals, and their expectations and reactions diverge. Moreover, as all agents pay attention and react strongly, each agent faces higher uncertainty about others' aggregate response, because they are uncertain about others' beliefs and reactions under dispersed information.

All these phenomena are prominent features of the data. Countercyclical fluctuations of the four uncertainty measures are well-known business cycle phenomena studied extensively in the literature;[1] countercyclical attention also has ample empirical support.[2] Existing works either rely on exogenous heteroskedasticity shocks to generate these fluctuations or only focus on a few phenomena. The main contribution of this paper is to show that, when information is dispersed, *a single mechanism* of people varying their attention to macroeconomic events provides a unified explanation for all these facts.

I characterize the aforementioned results with a static model to illustrate the economic insights. To evaluate the mechanism quantitatively, I extend the model to a dynamic dispersed information economy where the aggregate state follows a persistent process and agents make dynamic attention and production decisions under persistent dispersed information.

Solving the equilibrium fluctuations of attention and uncertainty poses a significant challenge to existing solution methods. The challenge stems from an *infinite regress*

1. Bloom (2014, 2009), Bachmann et al. (2013), Vavra (2014), Jurado et al. (2015), Ilut et al. (2018), and Bloom et al. (2018) document related empirical facts.

2. Coibion and Gorodnichenko (2015) and Flynn and Sastry (2020) provides empirical support for countercyclical attention to macroeconomic events. I provide further evidence in the appendix of this paper.

problem: under persistent dispersed information, agents need to not only solve a filtering problem about the unknown aggregate state, but also form beliefs about others' beliefs about the state, and others' beliefs about others' beliefs, and so on. In general, the economy lacks a finite recursive state space. Recent progress in the literature has found ways to handle the infinite regress problem under first-order approximations.[3] Yet, first-order approximations miss important *higher-order properties* of dispersed information economies: among other well-known limitations, existing methods are constrained to static information structures and linear "actions dynamics" — they cannot capture the fluctuations of attention and uncertainty measures, which are all higher-order properties of the model.

In Chapter 2, I address this problem under a general framework. I develop a perturbation method that solves higher-order approximation of equilibrium dynamics for a large class of dispersed information economies. The method generalizes existing first-order methods to arbitrary high orders of approximation, and overcomes the infinite regress problem under which standard non-linear methods are not applicable. Although the method applies to a large class of dispersed information economies and addresses a broad set of limitations faced by first order methods, its application to this paper captures the fluctuations of attention and uncertainty measures that existing methods would have left out.

With the new method, I calibrate the model to match salient business cycle moments

3. See Huo and Takayama (2015), Nimark (2014), Huo and Pedroni (2020), and Angeletos and Huo (2018) for models with exogenous information structure, and Maćkowiak and Wiederholt (2015) and Mackowiak and Wiederholt (2009) for attention choice with quadratic approximation of payoff around the steady state.

from the aggregate data and key moments from U.S. forecast-survey data. These moments jointly pin down the level of strategic complementarity and the long-run attention level in the model — two key features of the model that determine the conditions under which the mechanism is at work.

Once these key features are pinned down by the data, endogenous attention choices in the calibrated model generate countercyclical fluctuations in all four measures of uncertainty with cyclicality and persistence comparable to the data, and of magnitudes roughly consistent with data. Moreover, the model generates forecast patterns consistent with evidence from forecast survey documented by Coibion and Gorodnichenko (2015) that is indicative of countercyclical attention to macroeconomic variables. All these fluctuations are driven by the single economic mechanism in the model: endogenous attention response under dispersed information.

Literature

This paper is related to several strands of literature.

First, the model builds on the dispersed information and rational inattention literature. Since Phelps (1970) and Lucas (1972), macroeconomists have long recognized the potential of information frictions in explaining macroeconomic phenomena. The literature has been revived since Sims (2003) and Woodford (2001), followed by Angeletos and La'O (2013) and Angeletos and La'o (2010), Maćkowiak and Wiederholt (2015) and Mackowiak and Wiederholt (2009), and Huo and Takayama (2015) among many others. However, the information structure in these models, exogenous or endogenous, has been kept constant over time, both due to technical constraints and their focus on explaining sluggish adjustment in the data. This paper shifts the focus to a different set of business cycle phenomena — the fluctuations of uncertainty over business cycles. I show that once the information structure is allowed to respond to economic conditions, a canonical model in this literature provides unified explanation for a broad set of phenomena related to uncertainty fluctuations. Furthermore, I develop a methodology that is crucial to analyzing higher-order properties of dispersed information economies. The method can be easily applied to revisit earlier works where higher-order properties of the models were neglected. Two papers in the literature study how information about aggregate variables varies over time. Mäkinen and Ohl (2015) study the efficiency of information acquisition over business cycles in a model where information choice and state of the world are both binary. Flynn and Sastry (2020) study how a flexible stochastic attention choice model generates countercyclical variation in attention and aggregate volatility without explicit modeling of information structure. By contrast, this paper shows attention response in a canonical dispersed information economy explains a broad set of phenomena related uncertainty fluctuations, including but beyond aggregate volatility.

On the other hand, countercyclical fluctuations of the four uncertainty measures has been the focus of an extensive literature in the past decade. Bloom (2009), Fernández-Villaverde et al. (2011), Christiano et al. (2014), Gilchrist et al. (2014), Vavra (2014), Basu and Bundick (2017), and Bloom et al. (2018), among many others, generate these fluctuations with exogenous countercyclical heteroskedasticity shocks. Nimark (2014) and Kozeniauskas et al. (2018) study heteroskedasticity productivity shocks in dispersed information models. Others provide explanations for these phenomenon without relying on heteroskedasticity shocks, but focus only on a few phenomena. Van Nieuwerburgh and Veldkamp (2006), Benhabib et al. (2016), and Fajgelbaum et al. (2017) generate countercyclical uncertainty with procyclical information. Orlik and Veldkamp (2014) explain uncertainty fluctuations by allowing forecasters to estimate parameters of exogenous GDP process. Ilut et al. (2018) explain variations in aggregate volatility and cross-sectional dispersion of employment with time-varying responsiveness from a reduced-form concave hiring rule. By contrast, I show attention choice under dispersed information alone provides a unified explanation for these facts and illustrate how these phenomena are related to the fundamental production and information structure of an economy.

Third, the methodology I develop in Chapter 2 and apply to this paper is related to works that use perturbation methods to address economic problems, starting from Judd (1998). It is most closely related to the small shock expansion in the space of stochastic processes introduced by Lombardo and Uhlig (2018), robust preference expansion in Borovicka and Hansen (2013), and perturbation with cross-sectional heterogeneity in Bhandari et al. (2018). I build on their insight to develop a perturbation method for solving dispersed information models to arbitrary high order, including models with the infinite regress problem where state variables of an economy consist of a whole distribution of histories of signals.

Finally, this paper is related to the empirical literature on macroeconomic expectations, most closely related to Coibion and Gorodnichenko (2015). They show the regression coefficient of mean forecast error on mean forecast revision is closely related to the size of information frictions through the length of many models with information frictions. The quantitative strategy of this paper builds on this insight to pin down key parameters related to the information structure in the model. Angeletos and Huo (2018) also adopt a similar strategy.

Outline

The rest of this paper is organized as follows. Section 1.2 illustrates how endogenous attention affects the four measures of uncertainty in a static dispersed information economy. Section 1.3 discusses empirical evidence related to the four uncertainty measures and attention. Section 1.4 extends the economy to a dynamic setup, discusses the computational challenge posed by the model, and gives an overview of the solution method. Section 1.5 discusses calibration. Section 1.6 shows quantitative results. Section 1.7 concludes. All proofs, details of the data, and details of numerical implementation are provided in the appendix.

1.2 Model

Preference and technology

Consider an economy with a continuum of agents indexed by i. Agents have preference over final good consumption c_i and labor n_i and attention z_i:

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

where they trade off consumption and labor with the Greenwood–Hercowitz–Huffman (GHH) preference:

$$u(c_i, n_i) = \frac{1}{1-\gamma}\Big(c_i - \frac{\psi}{1+\nu} n_i^{1+\nu} \Big)^{1-\gamma},$$

and agents derive disutility from attention with an linear additive cost $\kappa\ z_i$.[4]

Each agent uses labor to produce a unique intermediate good with linear technology:

$$q_i = n_i,$$

where q_i denotes the quantity of the intermediate good produced.

Let the final good be the numeraire, and let p_i denotes the relative price of intermediate good i to the final good. Agent i faces budget constraint:

$$c_i = p_i q_i.$$

A representative competitive final good producer produce Y units of the final good with $\{y_i\}$ units of intermediate goods to maximizes profit taking prices $\{p_i\}$ as given:

$$Y - \int p_i y_i.$$

4. Additive attention cost is a common assumption in the literature, see Sims (2010), Myatt and Wallace (2012), Angeletos and Sastry (2019), for example. I discuss the implication of this assumption in the later part of this section, and study a more general formulation in the appendix.

The final good producer produce with a constant elasticity of substitution technology with aggregate productivity θ:

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

Parameter $\eta < 1$ denotes the inverse of elasticity of substitution, and aggregate productivity θ is distributed as:

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

As a prelude, the key theoretical analysis of this paper will be a comparative static exercise with respect to the expected aggregate productivity, $\bar{\theta}$. This exercise captures how the economy responds to changes in "business cycle condition", because in the dynamic model, expected productivity in a period will result from shocks that hit the economy in the past.

Timeline and information

The economy lasts for three periods, $t = 0, 1, 2$, and proceeds as follows.

At the beginning of the economy, aggregate productivity θ is unknown, and agents hold common prior

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

In period $t = 0$, each agent i chooses their attention level z_i simultaneously.

In period $t = 1$, each agent i receives an idiosyncratic signal x_i about θ with precision z_i:

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

Based on the signal x_i, agent i forms an expectation about the price of their product p_i and decides labor input n_i.

In period $t = 2$, 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 their proceeds of selling the intermediate good, and they consume subject to their budget constraints.

Definition of equilibrium

Let $\boldsymbol{\sigma}(X)$ denote the sigma-algebra generated by random variables X.

An equilibrium is a collection of $z_i \in \mathbb{R}$, random variables $n_i, q_i \in \boldsymbol{\sigma}(x_i)$, $p_i, c_i, y_i \in \boldsymbol{\sigma}(x_i, \theta)$ for all i, and $Y \in \boldsymbol{\sigma}(\theta)$ such that the following is true:

1. agent i chooses z_i to optimize expected utility given their prior;
2. agent i chooses n_i to optimize expected utility conditional on x_i;
3. given p_i, agent i consumes c_i subject to the budget constraint;
4. given $\{p_i\}$, final good producers choose $\{y_i\}$ to optimize profit;

5. given $\{n_i\}, \{y_i\}$, production $\{q_i\}, Y$ are determined by their respective technology;

6. markets clear

$$y_i = q_i, \ \forall i, \quad \text{and} \quad Y = \int c_i.$$

Because agents are identical except for their information, I look for a symmetric equilibrium in which $z_i = z, \forall\ i$, and agents adopt an identical strategy that maps signal realizations to labor input.

Equilibrium

Because all uncertainty is resolved in period $t = 2$, the equilibrium in this period can be solved easily given any set of labor input $\{n_i\}$ chosen in the previous period.

The final good producer takes prices $\{p_i\}$ as given and choose y_i to maximize profit. Their profit-maximization problem leads to the standard CES demand for intermediate good:

$$p_i = e^{(1-\eta)\theta}\ Y^{\eta}\ y_i^{-\eta}.$$

Given labor input $\{n_i\}$, markets clearing and production feasibility imply the equilibrium price of intermediate good i relative to the final good can be written as:

$$p(\theta, N, n_i) = e^{\theta} N^{\eta} n_i^{-\eta},$$

where "aggregate labor" N is given by an aggregation condition:

$$N := \left(\int n_i^{1-\eta} \right)^{\frac{1}{1-\eta}}.$$

The price of intermediate good i increases with θ because, when productivity is high, more of the final good is produced. It increases with N because 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, that is, the demand is downward sloped.

To understand the problem agents face at period $t = 1$, I first consider a special case where $\sigma = 0$, so no uncertainty is present in the economy, and $\theta = \bar{\theta}$ with probability one. I then illustrate how uncertainty changes an agent's problem and characterize the equilibrium under uncertainty.

Special case: No uncertainty $\sigma = 0$

When $\sigma = 0$, agents' attention level is irrelevant and they choose their labor n_i without any uncertainty. Their optimization problem simply equalizes marginal revenue to marginal cost:

$$\max_{n_i} \quad u\big(\ p(\bar{\theta}, N, n_i)\ n_i,\ \ n_i\ \big)$$

$$[n_i]: \quad (1-\eta)e^{\bar{\theta}} N^{\eta} n_i^{-\eta} = \psi n_i^{\nu}.$$

Rearranging and taking log, the equilibrium is described the linear best-response function,

$$(\nu + \eta)\ \bar{n}_i = \bar{\theta} + \ \eta\ \bar{N} + constant,$$

and aggregation condition,

$$\bar{N} = \bar{n}_i,\ \forall i,$$

where $\bar{n}_i, \bar{N}$ denote the log of individual and aggregate labor in equilibrium.

In the best-response function above, agent i's labor choice depends on aggregate productivity and aggregate labor through two parameters, ν and η. Parameter ν is the convexity of labor cost. Agents respond to aggregate variables more strongly when the convexity of labor cost is low. Parameter η determines the complementarity between intermediate goods. It has two effects on labor decision. First, when η is higher, agents respond less to aggregate variables because goods are less substitutable and the demand curve for each good is steeper. Second, when η is high, goods are more complementary and an increase in other agents' labor input shifts up the demand for good i more and leads to a stronger response in agent i's labor input. The complementarity between intermediate goods is the source of the general equilibrium effect in the economy — it relates one agent's decision to that of others' decisions. As I show shortly, the strength of general equilibrium effect in the economy is essential for understanding the fluctuations of uncertainty over business cycles.

General case: With uncertainty $\sigma > 0$

When $\sigma > 0$, agents can no longer perfectly predict the aggregate productivity and need to make decisions under uncertainty about θ and other agents' aggregate reaction, N.

Agents form expectations about their marginal product of labor, and the first order condition for labor input n_i is given by:

$$[n_i]: \quad \mathbb{E}_i\left[(1-\eta)e^{\theta}N^{\eta}n_i^{-\eta} \times \partial u_i \right] = n_i^{\nu},$$

where

$$\mathbb{E}_i[\ \cdot\ ] \coloneqq \mathbb{E}[\ \cdot\ |\bar{\theta}, x_i], \quad \text{and} \quad \partial u_i \coloneqq \frac{u_c(\ p(\theta, N, n_i)n_i,\ n_i)}{\mathbb{E}_i[\ u_c(\ p(\theta, N, n_i)n_i,\ n_i)\ ]}$$

denotes the normalized marginal utility with which agent i weights different contingencies.

Under uncertainty, how much agents respond to changes in aggregates θ and N depends not only on parameters ν and η, but also on how much they know about aggregates based on their signals x_i. If agent i pays more attention and receives a more precise signal, they know more about the likely realization of aggregates, and their labor input responds more to them. If all agents pay more attention, each agent expects aggregate labor to respond more to changes in θ as others respond, and their expectations about the aggregate response induce further responses from each agent.

At $t = 0$, given equilibrium attention $\bar{z}$ chosen by other agents, an agent i's optimal of attention z_i solves:

$$[z_i]: \quad \mathbb{E}\Big[-\frac{\epsilon_i^2 - 1}{2z_i} \times u_i \;\Big|\; \bar{\theta} \Big] = \kappa.$$

where

$$u_i = u(p(\theta, N, n_i)n_i, n_i) = \frac{1}{1-\gamma}\Big(e^{\theta} N^{\eta} n_i^{1-\eta} - \frac{\psi}{1+\nu} n_i^{1+\nu} \Big)^{1-\gamma}.$$

and, by envelope theorem, their labor input, $n_i(x_i)$, following realizations of signal x_i, is given by the optimality condition of labor at $t = 1$.

When choosing attention z_i, agents equalize the marginal value of attention to the marginal cost κ. The marginal value of attention can be understood as a projection of an agent's payoff on the size of noise in their signal. The projection captures how much reduction in noise increases their payoff given the optimal use of information described by their labor input strategy, $n_i(x_i)$.

Curvatures in labor cost and demand, ν and η, make it useful for agents to obtain precise signals: the more convex is labor cost and the steeper is demand curve, the more costly it is to make mistakes in labor input. Given ν and η, how costly is a percentage mistake in consumption unit depends on the level of productivity and labor input. When the levels of productivity and labor input are high, the same reduction of uncertainty about θ transforms into more "income" (i.e. an agent's revenue net of labor cost).

On the other hand, parameter γ determines how much an agent's utility changes with different levels of income. Given γ, when agents expect low income, they expect high marginal utility and disutility from consumption and labor, and the marginal rate of substitution between attention and income is high.

Equilibrium approximation

Under uncertainty, solving the equilibrium essentially boils down to solving the fixed-point problem of functional equations: we need to find functions $\boldsymbol{N}(\cdot), \boldsymbol{n}(\cdot)$ and z:

$$\log N = \boldsymbol{N}(\theta), \quad \log n_i = \boldsymbol{n}(x_i), \quad z \in \mathbb{R},$$

that satisfy the optimality of attention, the optimality of labor input and the aggregation condition of aggregate labor. Because the problem generally does not have a closed-form solution, I proceed with a perturbation-based approximation method developed in Chapter 2 to approximate the equilibrium. The method applies to a large class of dispersed information economy and allows one to study *higher-order properties* of dispersed information models, including how attention respond to the business cycle and its implication for uncertainty fluctuations. It delivers a systematic approach to study static dispersed information models analytically, and provides a way to solve higher-order dynamics of dynamic dispersed information economies. I describe how the method applies to the static model below, give an overview of its application to dynamic problems in section 1.4.

To approximate the equilibrium of the economy, I consider a sequence of economies indexed by a perturbation parameter δ. In economy δ, the aggregate productivity, signal noises, and marginal cost of attention are given respectively by:

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

The approximation strategy is to use the limiting equilibrium where δ goes to 0 to construct a Taylor expansion along the sequence of economies to approximate the equilibrium at δ equals to 1.

This sequence of economies has three features. First, when δ equals to 1, the economy is identical to the one we would like to approximate; as δ goes to 0, the equilibrium labor input converges to the special case without uncertainty. Second, fixing an attention level for all economies along the sequence, the information structure is identical in each of economy because signal noises ϵ_i is scaled proportionally to $\hat{\theta}$ along the sequence. Third, the attention cost is scaled by δ^2 along the sequence. This purpose of scaling attention cost is to keep the marginal value and marginal cost of information at the same order along the sequence of economies, so that the information structure in these economies resemble the one we would like to approximate. The marginal value of information is second order for economies with small shocks because agents optimize their labor input, and small change in labor input has no first order effect on their payoff. Therefore, a sequence of economies with attention cost vanishing in second order serves the purpose.

Let $z^\delta, n_i^\delta, N^\delta$ denote the equilibrium in economy δ, and consider the following mappings from δ to equilibrium objects:

$$\log N(\delta) : \delta \to N^{\delta}, \quad \log n_i(\delta) : \delta \to n_i^{\delta}, \quad \log z(\delta) : \delta \to z^{\delta}.$$

The Taylor expansions of these mappings with respect to δ around the limiting economy $\delta \to 0$ is given by:

$$\begin{aligned}
\log N(\delta) &= \bar{N} + \hat{N}\ \delta + \frac{1}{2}\hat{\hat{N}}\ \delta^2 + \cdots \\
\log n_i(\delta) &= \bar{n} + \hat{n}_i\ \delta + \frac{1}{2}\hat{\hat{n}}_i\ \delta^2 + \cdots \\
\log z(\delta) &= \log \bar{z} + \hat{z}\ \delta + \frac{1}{2}\hat{\hat{z}}\ \delta^2 + \cdots ,
\end{aligned}$$

where $\hat{N}, \hat{n}_i, \ldots$ are derivatives of the mappings evaluated at $\delta = 0$. Note these derivatives are random variables because $\log N(\cdot), \log n_i(\cdot)$ maps δ to random variables.

These derivatives are characterized by expanding the equilibrium conditions with respect to δ to corresponding orders, and evaluate at the limiting equilibrium where $\delta \to 0$.

Finally, to see how we can solve these derivatives, take $\hat{N}$ for example. Let $\boldsymbol{N}(\cdot, \delta)$ denote the solution function that maps $\theta(\delta)$ to aggregate labor in economy δ:

$$\log N(\delta) = \boldsymbol{N}(\theta(\delta), \delta).$$

Differentiating with respect to δ and evaluating at $\delta = 0$, we have:

$$\hat{N} = \overline{\boldsymbol{N}}_\theta \ \hat{\theta} + \overline{\boldsymbol{N}}_\delta,$$

where coefficients $\overline{\boldsymbol{N}}_\theta, \overline{\boldsymbol{N}}_\delta \in \mathbb{R}$ are partial derivatives of $\boldsymbol{N}(\cdot, \cdot)$ evaluated at $(0, 0)$, and the task of solving for $\hat{N}$ reduces to solving coefficients $\overline{\boldsymbol{N}}_\theta, \overline{\boldsymbol{N}}_\delta$. Similar is the case for other equilibrium objects and for higher-order expansions. From the expanded equilibrium conditions, we can easily solve for these coefficients. Detail of the derivation is provided in the appendix.

Four measures of uncertainty in the economy

With the approximation method to analyze the equilibrium, I now describe the four measures of uncertainty, and study how they respond to changes in business cycle condition, $\bar{\theta}$.

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

1. Conditional volatility of aggregate output:

$$SD(\ \widetilde{Y}\ |\bar{\theta}) := \Big(\mathbb{E}\Big[\ \big(\widetilde{Y} - \mathbb{E}[\ \widetilde{Y}\ |\bar{\theta}]\big)^2\ \Big|\bar{\theta}\Big]\Big)^{\frac{1}{2}}.$$

2. Dispersion of individual output:

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

3. Forecast dispersion about aggregate output:

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

4. Forecast uncertainty about aggregate output:

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

Conditional volatility of aggregate output is defined as the standard deviation of aggregate output conditional on business cycle condition, $\bar{\theta}$. Dispersion of individual output is the cross-sectional standard deviation of output across agents. Forecast dispersion about aggregate output is the cross-sectional standard deviation of agents' expectations about aggregate output, conditional on their signals x_i. Forecast uncertainty for each agent i is the conditional standard deviation of aggregate output given agent i's signal x_i, which describes the subjective uncertainty agent i face about aggregate output based on their information.

Each of the four measures corresponds to a distinct and widely studied business cycle phenomenon related to uncertainty fluctuations. With the static setup, these four measures are theoretical objects that capture the essence of these phenomena. I use these theoretical objects to illustrate the key insight that these phenomena related uncertainty fluctuations can be understood through a single mechanism: endogenous attention response under dispersed information. I discuss how these theoretical objects map into data in section 1.3, and close the gap between model and data in section 1.6. I focus on measures related to output because empirical evidence are more complete for output related measures; in the appendix, I show all results hold for input as well, consistent with evidence related to input in the literature.

In the analysis below, I characterize the four uncertainty measures up to second order approximation around a given $\bar{\theta}$, and study how they respond to changes in $\bar{\theta}$ as a comparative statics result. In the dynamics setup, how uncertainty measures fluctuate over business cycles will be captured naturally by the second-order equilibrium dynamics when agents' expectations about current economic conditions — and therefore attention level — fluctuates with shocks to the economy.

I show the following two lemmas as intermediate steps.

Lemma 1

Up to second order approximation,

$$SD(\ \widetilde{Y}\ |\bar{\theta}), \qquad Disp(\ \tilde{y}_i\ |\bar{\theta}),$$

$$Disp(\ \mathbb{E}_i[\widetilde{Y}]\ |\bar{\theta}), \quad and \quad SD(\ \widetilde{Y}\ |x_i, \bar{\theta}), \forall i,$$

depends on $\bar{\theta}$ only through $\bar{z}(\bar{\theta})$.

To understand the statement of the lemma, consider an economy identical to the one we have studied, except that agents' signal precision is fixed exogenously at some level z. The lemma says that, in this economy, the four uncertainty measures are constant with respect to changes in $\bar{\theta}$ up to second order approximation.

Note the four measures of uncertainty captures how strongly the economy responds to shocks and noises conditional on $\bar{\theta}$. Because all shocks and noises are homoskedastic and technologies have constant elasticises, a change in $\bar{\theta}$ changes the level of output, but does not affect how much log output responds to shocks by itself. Similarly, because $\bar{\theta}$ changes agents expectation about aggregate output in a uniform and deterministic way, it has no effect on dispersion of outcome, forecast dispersion, and forecast uncertainty. This characterization highlight a single mechanism that drives all the changes in uncertainty measures: the variation of attention in response to $\bar{\theta}$.

The next lemma characterize how equilibrium attention responds a change in $\bar{\theta}$.

Lemma 2

How $\bar{z}(\bar{\theta})$ responds to $\bar{\theta}$ is determined by γ:

$$\frac{\partial}{\partial \bar{\theta}}\,\bar{z}(\bar{\theta}) \lesseqgtr 0 \iff \gamma \gtreqless 1.$$

Attention's response to a change in $\bar{\theta}$ depends on two competing forces: the income and 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. When agents expect low productivity, they expect low labor input. With labor input at a low level, avoiding a 1% mistake saves a fewer units of income. For this reason, a decrease in expected productivity $\bar{\theta}$ discourages agents from paying attention to learn about the 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, the expected marginal utility of income is high, and avoiding every unit of income loss due to inattention becomes more valuable. For this reason, a decrease in expected productivity $\bar{\theta}$ motivates agents to pay attention and reduce the noise in their signals.

When $\gamma > 1$, the curvature of utility over income is large enough such that a decrease in expected productivity causes an increase in marginal utility so that agents pay more attention when they expect a lower productivity: the income effect on attention dominates the substitution effect.

The two lemmas above together describe how attention respond to business cycle and highlight attention response as the single mechanism that affects the four measures of uncertainty in this economy.

The following proposition characterize two key conditions under which countercyclical attention response generates countercyclical comovement between all four measures of uncertainty: (i) a low level of attention, and (ii) a high level of *strategic complementarity* (i.e., how much one agent's decision depends on other agents' decisions).

Proposition 1

Suppose that $\gamma > 1$ (attention is countercyclical). Given η, ν, $\exists\ \check{z}$ such that if $\bar{z}(\bar{\theta}) < \check{z}$, then

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

up to second order approximation.

Moreover, let

$$r := \frac{1}{\eta + \nu}, \quad \textit{and} \quad s := \frac{\eta}{\eta + \nu},$$

then $\check{z} > 0$ if $r > \frac{1}{2}$ and $\check{z} \to \infty$ as $s \to 1$.

With a decrease in $\bar{\theta}$, agents expect lower productivity and lower income. They pay more attention to avoid mistakes in labor input because the marginal value of income is high when income is low — that is, income effect on attention dominates the substitution effect.

As agents pay more attention, aggregate labor becomes more responsive to movements in the unknown aggregate state and conditional volatility of output increases.

When agents pay more attention, they learn about the unknown aggregate productivity θ and face lower uncertainty about it. Yet, under dispersed information, agents face another source of uncertainty: uncertainty about other agents' aggregate response, N. If the economy features strong strategic complementarity and a low level of attention initially, an increase in attention induces high uncertainty because other agents' endogenous response is strong and agents know little about the aggregate state to which other agents respond. Consequently, all agents face higher uncertainty when forecasting aggregate output, despite each having a better understanding of the underlying state, θ.

To understand why forecast dispersion increases, note two effects. First, as agents pay more attention to reduce idiosyncratic noise in their signals, they update their beliefs more with their signals. If agents' initial attention is low, paying attention causes them to incorporate more idiosyncratic noises from signals into beliefs, and their expectations about aggregate output diverge despite the size of idiosyncratic noises in signals is reduced. Second, each agent understands that when others pay more attention, aggregate labor varies more with the unknown state θ. As a result, agents' dispersed beliefs about θ leads to a larger forecast dispersion about aggregate output because other agents' aggregate response is stronger when they pay attention. This channel is stronger when the level of strategic complementarity is high — that is, the general equilibrium effect is strong.

Finally, as agents' expectations about aggregates diverge, so do their expectations about the demand for their goods. Because their decisions reflect their expectations about demand, their labor-input decisions diverge and result in a higher dispersion in output across agents.

None of the comovements between uncertainty measures are mechanically related. Instead, they all underscore the importance of dispersed information in understanding uncertainty fluctuations over business cycles.

To understand why dispersed information is important for the comovement between conditional volatility and forecast uncertainty, consider an economy identical to the one we've studied, except that agents do not observe idiosyncratic signals but receive a public signal with precision z':[5]

$$x = \theta + \frac{\epsilon}{\sqrt{z'}}, \quad \epsilon \sim \mathcal{N}\Big(0, 1\Big).$$

In this economy, agents all receive the same signal and face no uncertainty about aggregate labor. Agents still face forecast uncertainty about aggregate output because they are uncertain about the productivity θ. Yet, changes in z' always cause conditional volatility and forecast uncertainty to move in opposite directions: more precise signals are accompanied with large response and low forecast uncertainty, whereas less precise signals leads to dampen response and higher forecast uncertainty.

5. The external procyclical information mechanism studied in Van Nieuwerburgh and Veldkamp (2006), and Fajgelbaum et al. (2017) corresponds to modeling z' as an increasing function of $\bar{\theta}$.

This emphasize the importance of dispersed information because, when information is dispersed, agents face uncertainty not only about the exogenous fundamental, θ, but also the endogenous response of other agents, N. Therefore, large endogenous response is not only compatible with high forecast uncertainty, but is *the reason* why agents face high uncertainty forecasting aggregate output in a dispersed information economy.

On the other hand, an increase in outcome dispersion is not necessarily associated with an increase in forecast dispersion. For example, in models that generate countercyclical outcome dispersion with an exogenous increase in idiosyncratic productivity shocks,[6] agents all have identical expectations about the aggregate economy because they all receive identical information despite different realizations of their idiosyncratic productivity shocks. These models contain no forecast dispersion across agents.

The comovement between forecast dispersion and forecast uncertainty is also not guaranteed. Consider an economy in which agents have heterogeneous but commonly known beliefs, where forecast dispersion varies with shocks to their beliefs. Shocks to heterogeneous beliefs generates forecast dispersion, but does not by itself lead to higher forecast uncertainty because agents agree to disagree. By contrast, forecast dispersion and forecast uncertainty comove in proposition 1 because agents face uncertainty about others' beliefs under dispersed information.[7]

6. For example, Bloom (2009) and Bloom (2014) among many others.

7. This illustrate the difference between dispersed information and heterogeneous beliefs models when higher order properties, such as forecast uncertainty, are taken into account. By contrast,

Finally, proposition 1 shows countercyclical attention provides a unified explanation for countercyclical movements of all four measures of uncertainty. Yet, the mechanism is also complementary to models that features countercyclical shocks to volatility of exogenous fundamentals. In fact, if the volatility of productivity σ increases exogenously when $\bar{\theta}$ decreases, it will lead to an increase in attention and the mechanism described above will amplify changes in all four measures of uncertainty. I show this amplification mechanism in the appendix.

Remarks

Generality of the result

The result above is obtained in a specific economy and generates attention response with a particular mechanism of income effect on attention. Yet, the key observation about how attention affects the four measures of uncertainty applies to a general class of "beauty-contest models," including dispersed information New Keynesian models and models of financial markets; see Angeletos and Lian (2016) for a survey of this class of models. In this class of models, as long as the level of strategic complementarity is high or the level of attention is low, an increase in private signal precision always leads to an increase in all four measures of uncertainty.

Attention responds to endogenous change in volatility

Although the competing forces between the income effect and the substitution ef-

under first order approximation, Angeletos et al. (2018) obtain an observational-equivalence result between dispersed information economies and economies with heterogeneous beliefs.

fect determines how equilibrium attention changes with $\bar{\theta}$ in the model, it does not mean they are the only reason people pay attention. In fact, a feedback loop exists between people's attention and aggregate volatility: when others pay more attention, aggregate labor fluctuates more with the unknown aggregate productivity; this motivates each agent in the economy to pay more attention and respond to changes in aggregate labor. This feedback loop is reminiscent of a recurring scene in the literature related to attention choice in coordination games, such as Hellwig and Veldkamp (2009),Myatt and Wallace (2012), and Benhabib et al. (2016). In these models, agents' attention choices exhibit strategic complementarity when there is strategic complementarity between their actions.

Separating risk aversion from the income effect on attention

The model assumes an additive attention cost, as is commonly assumed in the rational inattention literature. Under this assumption, parameter γ plays a dual role of determining agents' risk aversion and the strength of the income effect on attention — one trades off income realization across different states of the world and the other trades off utility derived from income to the disutility from paying attention. In the appendix, I show these two considerations can be separated by a preference that trades off between the disutility of attention and the certainty equivalence of income. An elasticity that governs this trade-off will be the strength of the income effect. The certainty equivalence will be separately determined by the level of risk aversion.

Generalization of cost function and signal specifications

All results above hold under a more general class of cost functions. Moreover, when the attention cost is specified as a function of percentage variance reduction, all the analyses above are identical if agents receive signals about endogenous objects such as aggregate labor and output. The reason is that, under the approximation above, aggregate labor and output are linear functions of $\hat{\theta}$ with known coefficients. A signal about aggregate labor and output with a fixed percentage variance reduction leads to the same information as a signal of $\hat{\theta}$ with the same percentage variance reduction. Therefore, the assumption that agents receive information only about the aggregate productivity is without loss of generality in the static setup. Details are discussed in the appendix.

1.3 Uncertainty and attention fluctuations in the data

In this section, I discuss the four uncertainty measures and evidence on attention fluctuations in the data.

Countercyclical fluctuations of uncertainty measures

Countercyclical fluctuations of the four uncertainty measures are well known in the literature. I follow standard procedure in constructing these measures with minor modifications so the measures map closely to the model. I discuss each measure in order.

1. *Conditional volatility of aggregate output*

   Literature has provide ample evidence that recessions are times when macroeconomic variables become more volatile; see Jurado et al. (2015), Ilut et al. (2018), Adrian et al. (2019), for example. For a close mapping to the model, I construct a measure of how aggregate output volatility varies over business cycle by estimating the conditional heteroskedasticity of quarterly real GDP growth series with a GARCH model.[8] The GARCH model captures how shocks to expected aggregate output feed into the volatility of aggregate output — a drop in $\bar{\theta}$ leads to a change in $SD(\ \widetilde{Y}\ |\bar{\theta})$. Denote the empirical measure by

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

2. *Dispersion of individual output*

Countercyclical cross-sectional dispersion of individual outcomes has been documented extensively across various economic variables; see Bloom et al. (2018), Ilut et al. (2018), for example. To construct a measure that corresponds to the dispersion of revenue between agents, $Disp(\ y_i\ )$, I calculate the year-over-year revenue growth standard deviation across firms within sectors for each quarter in the Compustat data, and calculate the revenue-weighted average of these dispersions. Denote the empirical measure of dispersion of individual outcomes by

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

8. I estimate a EGARCH(1,1)-ARMA(2,2) model. The phenomenon is robust to different GARCH specifications and mean process specifications. Countercyclical GARCH phenomenon is closely related to the negative skewness of output growth in the data. In the dynamic setup, negative skewness of output growth is also an important implication of the mechanism.

3. *Forecast dispersion about aggregate output*

It is well known that forecast dispersion about aggregate variables are countercyclical; see Bachmann et al. (2013), Kozeniauskas et al. (2018), for example. For a measure that corresponds to $Disp(\ \mathbb{E}_i[\widetilde{Y}]\ )$, 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. Let $f_{i,t}^{\Delta\widetilde{Y}_t}$ denote the point forecast of each individual forecaster. Denote the empirical measure of forecast dispersion by:

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

4. *Forecast uncertainty about aggregate output*

To measure the subjective uncertainty agents face when forecasting aggregate output, $SD(\ \widetilde{Y}\ |x_i, \bar{\theta})$, I use the probability-range data in the SPF, where 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 forecast uncertainty; see Bloom (2014),Fajgelbaum et al. (2017), for example. I follow Engelberg et al. (2009), and fit a parametric distribution to the discrete probability weights submitted by each individual forecaster. Let $\Phi_{i,t}(\Delta\widetilde{Y}_t)$ denotes the fitted distribution for forecaster i about the possible aggregate output growth at time t. For each individual forecaster, I calculate their forecast uncertainty by taking the standard deviation of aggregate output growth with the fitted distribution, $\Phi_{i,t}(\Delta\widetilde{Y}_t)$. I then average over individuals forecast uncertainty to construct an aggregate measure:[9]

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

where

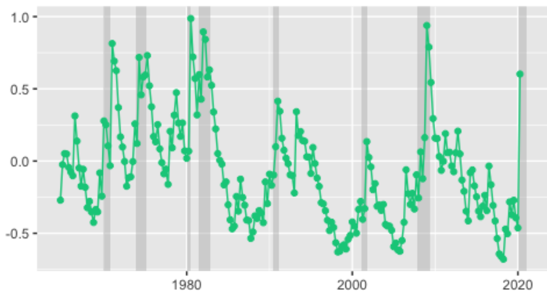
$$\widetilde{SD}_{i,t}(\Delta\widetilde{Y}_t) \coloneqq \left(\int \left(\Delta\widetilde{Y} - \mathbb{E}_{\Phi_{i,t}}[\Delta\widetilde{Y}]\right)^2 d\Phi_{i,t}(\Delta\widetilde{Y})\right)^{1/2},$$

and

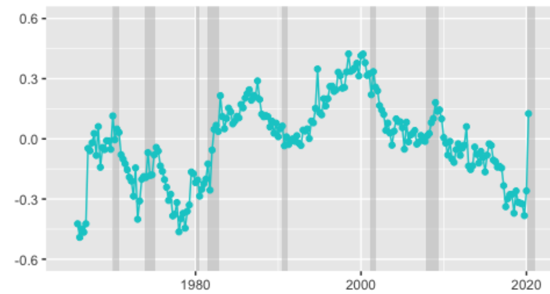
$$\mathbb{E}_{\Phi_{i,t}}[\Delta y] \coloneqq \int \Delta\widetilde{Y}\ d\Phi_{i,t}(\Delta\widetilde{Y}).$$

Although the mapping between model and data is not exact in the static setup, this gap closes in the dynamic model in section 1.4.

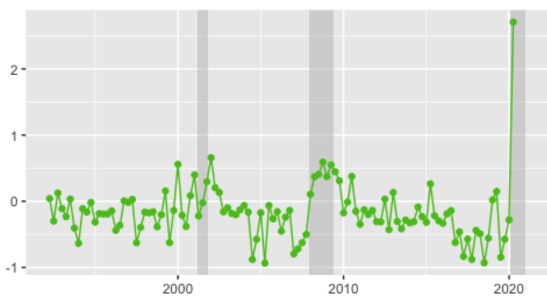
9. For the probability range data, SPF asks for fixed-event forecasts of year-to-year GDP growth. To compare the measure over time, I calculate the deviation of average forecast uncertainty from its quarter-of-year average for each quarter.



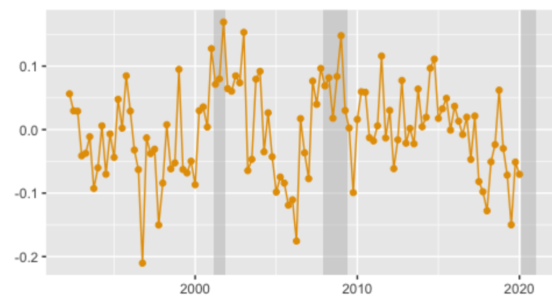
Conditional volatility of aggregate output



Dispersion of individual output



Forecast dispersion about aggregate output



Forecast uncertainty about aggregate output

Figure 1.1: Measures of uncertainty over time

Conditional volatility of GDP growth, forecast uncertainty about GDP growth, forecast dispersion of GDP growth, and dispersion of revenue growth. x-axis: time; y-axis: log-deviation of variables from the long-run average; gray area indicates NBER recessions.

Figure 1.1 show the log deviation of these four measures of uncertainty from their long-run average over time, where the gray area marks the NBER recession periods. All four measures show a strongly countercyclical comovement: rising sharply during recessions and declining during booms.

Countercyclical attention

Literature provides suggestive evidence that recessions are times when people pay more attention to macroeconomic events. Coibion and Gorodnichenko (2015) show that, under structural assumptions, professional forecasters update their beliefs about macroeconomic variables with higher Kalman gain during recessions — that is, the level of "information rigidity" is lower during recessions. This finding is consistent with the mechanism proposed by this paper where agents pay more attention to macroeconomic events and update their beliefs about the aggregate economy during recessions. Flynn and Sastry (2020) also provide evidence that firms' attention to macroeconomic issues are countercyclical, where attention is measured by the similarity of words between firms' SEC filing to introductory macroeconomics textbooks using a natural-language-processing algorithm.

In the appendix, I provide further evidence that people's attention to economic events is countercyclical. I construct three different proxies of attention to economic events using internet traffic data, following recent empirical finance literature. All three attention proxies show strong countercyclical patterns, supporting the prediction of the model that recessions are times in which people pay more attention to the economy.

1.4 Dynamic Model

To close the gap between model and data, I now extend the model to a dynamic setup and evaluate the mechanism quantitatively. With the dynamic model, I discipline key parameters of the model with salient business cycle moments and evidence about expectations from forecast survey data. These disciplines from the data pin down the level of strategic complementarity and the average level of attention in the model — the key features that determine how attention dynamics affects the four uncertainty measures. Along the way, I explain the limitations of existing methods in solving attention dynamics and uncertainty fluctuations — or more generally higher-order dynamics of dispersed information economies, and discuss how the perturbation approach discussed in section 1.2 can be extended to overcome these limitations.

Preference and technology

Time lasts from $t = 0, \cdots, \infty$. A continuum of agents indexed by i make consumption, labor and attention decisions with preference

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

where $c_{i,t}, n_{i,t}, z_{i,t}$ denote agent i's consumption, labor and attention in period t, $u(c_{i,t}, n_{i,t})$ takes the GHH form as before, and β denotes their discount rate.

Agents produce intermediate goods with labor in each period:

$$q_{i,t} = n_{i,t}.$$

and face period-by-period budget constraints,

$$c_{i,t} = p_{i,t} q_{i,t}.$$

A representative competitive final good producer combines intermediate goods to produce final good with CES technology:

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

where aggregate productivity θ_t follows

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

and $\omega_{i,t}$ denotes idiosyncratic productivity shocks for intermediate good i. Assume these shocks are iid normal over time and across agents with variance $\sigma_{\omega_i}^2$.

Profit for the final good producer in period t is given by

$$Y_t - \int p_{i,t} y_{i,t}.$$

I abstract away from capital accumulation and saving decisions in this economy to focus on the dynamics generated purely from agents' endogenous attention fluctuations.

Information

In each period t, agents make decisions with the following order:

1. At the beginning of period t, each agent i chooses their attention level $z_{i,t}$, given information set $\mathcal{F}_{i,t-1}$.

2. Based on their attention level $z_{i,t}$, each agent i observes an idiosyncratic signal

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

   where $\epsilon_{i,t}$ is iid over time and across agents. Agents make labor-input decision $n_{i,t}$ after observing $x_{i,t}$.

3. After agents decide their labor input, price $\{p_{i,t}\}$ realizes. Each agent observes the price of their own product $p_{i,t}$ and consume subject to budget constraints.

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

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

where x_i^t, p_i^t denotes the history of each variable up to time t.

Because I will focus on the stationary behavior of the economy, the exact specification of agents' initial beliefs is inconsequential. Without loss of generality, assume that all agents have a common prior $\theta_0 \sim \mathcal{N}(0, \sigma_0^2)$ at $t = 0$.

Finally, in each period t, the final good producer observes prices $\{p_{i,t}\}$ and choose $\{y_{i,t}\}$ 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

1. $\{z_{i,t}, n_{i,t}, c_{i,t}\}$ optimize agent i's preference, subject to budget constraints and information constraints,
2. the representative final good producer chooses $\{y_{i,t}\}$ to optimize profit taking prices as given,
3. production $q_{i,t}, Y_t$ are given by their respective technologies, and
4. markets clear

$$q_{i,t} = y_{i,t}, \quad \forall i, t \quad \text{and} \quad Y_t = \int c_{j,t} \quad \forall t.$$

Equilibrium

From the final good producer's profit maximization problem, production feasibility, and markets clearing, we can solve the price of intermediate good i in each period t given $\{n_{i,t}\}$:

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

The optimality condition for attention, the optimality condition for labor input, and the aggregation condition lead to the following characterization:

Lemma 3

The necessary conditions for an equilibrium are:

$$[z_{i,t}] : \mathbb{E}_{i,t-1}\left[-\frac{\epsilon_{i,t}^2 - 1}{2z_{i,t}} \sum_{\tau=t}^{\infty} \beta^{\tau-t}\Big(U_{i,\tau} - \beta\kappa\ z_{i,\tau+1}\Big)\right] = \kappa,$$

$$[n_{i,t}] : \mathbb{E}'_{i,t-1}\left[\frac{\partial}{\partial n_{i,t}} U_{i,t} \right] = 0,$$

$$[Agg] : N_t = \left(\int \left(e^{\omega_{i,t}} n_{i,t} \right)^{1-\eta} \right)^{\frac{1}{1-\eta}},$$

where $\mathbb{E}_{i,t-1}, \mathbb{E}'_{i,t-1}$ *denote expectations conditional on* $\mathcal{F}_{i,t-1}$ *and* $\mathcal{F}'_{i,t-1}$*, and*

$$U_{i,\tau} \coloneqq u\big(\ p(\theta_\tau, N_\tau, n_{i,\tau}, \omega_{i,\tau})\ n_{i,\tau},\ n_{i,\tau}\big), \forall \tau \geq t.$$

The equilibrium conditions in the dynamic model resemble the ones in the static model. One important difference exists: whereas agents start with a common prior in the static model, the dynamic model features an *infinite regress problem*:

- Because θ_t is unobservable, agents need to solve a filtering problem to form beliefs about θ_t; and aggregate labor N_t reflects their beliefs.

- Because agents care about aggregate labor N_t but all observe different signals, they will also need to predict others' beliefs about θ_t.

- When all agents try to predict others' beliefs about θ_t, the aggregate labor N_t reflects their beliefs about others' beliefs. All agents now need to also predict others' beliefs about others' beliefs and so on.

In general, each agent's "individual state variable" is the whole history of their signals, their "policy function" is a non-linear function of infinite dimensional history of signals,[10] and the state variable of the economy is the whole cross-sectional distribution of infinite dimensional signal histories.

Recent literature has made progress in dealing with the infinite regress problem under log-linear approximation of equilibrium conditions or quadratic approximation of payoffs around the steady state.[11] Yet, these approaches have an important limitation: they are all "first order methods" in the sense that they all result in

10. See Huo and Takayama (2015) for detail discussion.

11. See Huo and Takayama (2015), Huo and Pedroni (2020), and Angeletos and Huo (2018) for solving log-linearized models with exogenous information structures; and Maćkowiak and Wiederholt (2015), Mackowiak and Wiederholt (2009), and Maćkowiak et al. (2018) for solving models with endogenous information structures with quadratic approximation of payoffs.

linear dynamics in action (labor input) and a static information structure (attention). This limitation constrained the literature from exploring higher-order properties of dispersed information economies — among these higher-order properties are the variation of attention over business cycle and the associated uncertainty fluctuations.

The perturbation-based method developed in Chapter 2 addresses this problem, because it allows me to solve the higher-order equilibrium dynamics of the model and capture attention and uncertainty fluctuations. I outline how the method applies to the dynamic setting in the appendix.

1.5 Calibration

External parameters

I interpret each period in the model as one quarter. Each agent i represents a firm. Agents pay attention, receive information, form forecast and make labor input decisions at the beginning of a quarter, and their outputs realize at the end of a quarter.

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 average markup over the marginal cost of labor is 25% in the steady state. I set σ_{ω_i} to 10% which is roughly the average size of idiosyncratic productivity shock in Bloom et al. (2018). Disutility of work ψ is set so that long-run labor input equals 1/3.

Parameter γ plays a dual role of determining risk aversion and the strength of income effect on attention,[12] and cannot be distinguished without departure from the convention of additive attention cost. I set γ to 10, which implies a strong income effect on attention that generates countercyclical attention choice that is consistent with data. I report results of different choices for this parameter in the appendix, and discuss a general specification of preference to separate the two roles.

Internal calibration

The remaining parameters fall into two groups. The first group concerns production: persistence of aggregate productivity process ρ, volatility of its innovation σ_ω, and the convexity of labor cost ν. The second group consists of the marginal cost of attention κ which determines the information structure of the economy in equilibrium:

$$\underbrace{\rho, \quad \sigma_\omega, \quad \nu,}_{\text{production}} \quad \underbrace{\kappa}_{\text{information}} \quad .$$

To calibrate the production parameters, I target (1) the persistence of aggregate output, (2) the long-run volatility aggregate output, and (3) the relative volatility of hours to output, all at a business cycle frequency of 6-32 quarters. Given the information structure of the economy, matching these moments pins down parameters

12. The model does not contain intertemporal trade off therefore the level of intertemporal elasticity of substitution is irrelevant. Up to second order approximation, the level of risk aversion affects only the level of labor input, which I calibrated to match a long-run level of 1/3 together with the disutility of labor ψ. Yet, one should note that, given the preference and technology specification, the level of labor input does not affect any business cycle property of the model up to second order approximation.

ρ, σ_ω, ν.

To calibrate the marginal cost of attention, I target the size of average forecast error about aggregate output from SPF. Specifically, let $f_{i,t}$ denote the one-quarter-ahead forecast of aggregate output at time t by agent i, and let $\overline{fe}_t$ denote the average forecast error at time t:

$$\overline{fe}_t := \widetilde{Y}_t - \int f_{i,t}\, di, \quad \text{where } f_{i,t} := \mathbb{E}'_{i,t-1}[\widetilde{Y}_t].$$

I follow common practice and measure the size of forecast error by the root-mean-square error, denoted by $||\overline{fe}_t||$. Given the size of fluctuations in aggregate output, large mean forecast errors indicate low attention level on average, which corresponds to a higher attention cost.

Calibration result

The target moments discussed above pins down the production and information parameters, and thereby determines the level of strategic complementarity and the average level of attention in the model — the two essential features that determines whether attention fluctuations can generate countercyclical fluctuations of the four uncertainty measures observed in the data.

Table 1.1 shows the targeted moments in the data and the calibration result.

Parameter	ρ	σ_ω	ν	$1/\sqrt{\bar{z}(\kappa)}$
	.87	.002	.08	.002

Moment	$\rho_1\ Y_t^\dagger$	$sd\ Y_t^\dagger$	$\frac{sd\ N_t^\dagger}{sd\ Y_t^\dagger}$	$\lvert\lvert\overline{fe}_t\rvert\rvert$
Data	.93	2.1	.91	.46
Model	.91	2.2	.91	.46

Table 1.1: Calibrated parameters and target moments

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

Because equilibrium long-run signal precision $\bar{z}$ is a monotonically decreasing function of attention cost κ and the value of attention cost κ does not have meaning per se, I report $\bar{z}$ in the table for ease of comparison.

1.6 Quantitative Results

Once key parameters of the model are pinned down by salient features of aggregate quantity data and long-run features of forecast-survey data, I compare the four uncertainty measures generated from the calibrated model with their counterpart from the data. Recall these measures are respectively:

1. the conditional volatility of aggregate output growth $\widetilde{SD}_t(\Delta\widetilde{Y}_t)$,

2. the dispersion of individual output growth $\widetilde{Disp}_t(\Delta\tilde{y}_{i,t})$.

3. the forecast dispersion of aggregate output $\widetilde{Disp}_t(f_{i,t}^{\Delta\widetilde{Y}_t})$, and

4. the average forecast uncertainty about aggregate output $\overline{\widetilde{SD}_{i,t}(\Delta\widetilde{Y}_t)}$

Fluctuations in uncertainty measures: cyclicality

Table 1.2 compares the cyclicality of the four uncertainty measures generated by the calibrated model to their counterparts in the data.

		$\widetilde{SD}_t(\Delta\widetilde{Y}_t)$	$\widetilde{Disp}_t(\Delta\tilde{y}_{i,t})$	$\widetilde{Disp}_t(f_{i,t}^{\Delta\widetilde{Y}_t})$	$\overline{\widetilde{SD}_{i,t}(\Delta\widetilde{Y}_t)}$
$cor(\cdot, Y_t)$	data	$-.70$	$-.58$	$-.58$	$-.37$
	model	$-.68$	$-.67$	$-.78$	$-.53$

Table 1.2: Measures of uncertainty in the model: cyclicality

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, but a quantitative result when key features of the model are pin downed by empirical features of the data in the calibration process. Although countercyclical attention is an immediate consequence of $\gamma > 1$, the model contains no apriori assumption about how the four uncertainty measures will fluctuate over the business cycle: had the target moments taken on different values, the calibration process would have lead to the opposite cyclical pattern for all these measures.[13]

Resonating the theoretical insight from the static model, these phenomena occur only because key features of the target moments imply a level of strategic complementarity and an average attention level such that countercyclical attention is prone to generate countercyclical uncertainty fluctuations. These features are: (1) a high relative volatility of input to output, (2) a large average size of forecast error of aggregate output relative to its fluctuations.

First, aggregate labor input is as volatile as aggregate output in the real aggregate data. This suggests the response of the endogenous variable to underlying shocks (labor input response to productivity) is strong over the business cycle, and to generate such a response, the calibration leads to a low convexity of labor cost ν. Given the level of complementarity between goods η, a low convexity of labor cost gives rise to a high level of strategic complementarity.

Second, the size of the forecast error about aggregate output growth in the forecast survey pins down the overall informativeness of signals in the model. In the calibrated model, agents update their expectation about aggregate productivity with an average Kalman gain around .6. To understand the meaning of this number, consider

13. In contrast to the static model, higher attention does not always lead to higher conditional volatility in a dynamic economy with public information. Conditional volatility is jointly determined by agents' knowledge about fundamentals and the information update they obtain in a period. If agents did not pay attention in the past because aggregate output has been high, they accumulate uncertainty about the fundamental, and realization of the public signal can induce a large update in beliefs even if agents sill choose a low level of attention.

a counterfactual exercise in which we fix all parameters and varies the attention cost κ. As different attention cost leads to different long-run attention level, the relationship between attention fluctuations and the four uncertainty measures also changes. In this counterfactual exercise, the countercyclical attention only start to generate procyclical movements in forecast uncertainty when the average Kalman gain reaches a level around .9. In this sense, the size of forecast errors suggest an average level of attention that is "relatively low", and countercyclical attention fluctuations leads to countercyclical uncertainty fluctuations in this environment.

Calibrating the model to match these empirical properties of the data give rise to an economy that features

1. a high level of strategic complementarity,

2. a low level of long-run attention, and

All these features explain why an increase in attention during economic downturn generates higher conditional volatility, dispersion of outcomes, forecast dispersion, and forecast uncertainty.

Fluctuations in uncertainty measures: other properties

Table 1.3 compares the average levels, the magnitudes of fluctuations, and the persistence of the four uncertainty measures generated by the calibrated model to their counterparts in the data.

		$\widetilde{SD}_t(\Delta\widetilde{Y}_t)$	$\widetilde{Disp}_t(\Delta\tilde{y}_{i,t})$	$\widetilde{Disp}_t(f_{i,t}^{\Delta\widetilde{Y}_t})$	$\overline{\widetilde{SD}_{i,t}(\Delta\widetilde{Y}_t)}$
avg	data	.91	15	.16	.52
	model	1.3	15	.56	1.1
$sd^{\dagger}$	data	.23	1.4	.09	.02
	model	.07	.04	.03	.04
$\rho_1^{\dagger}$	data	.88	.92	.87	.89
	model	.83	.87	.89	.84

Table 1.3: Measures of uncertainty in the model: average, variation, and persistence

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

Despite none of the moments the table above are targeted in the calibration, the four measures of uncertainty generated by the calibrated model show average levels, magnitudes of fluctuations, and persistence that are roughly consistent with the data.

The only measure that the model does not generate the same magnitude is the dispersion of individual output. This should not come as a surprise. Where as in the data, there are tremendous heterogeneity among firms in terms of age, size, balance sheet, and so on, that affects how responsive they are to business cycle condition, the model abstract away from all these concerns. Rather than a rejection of the model, I view the interaction between dispersed information and the heterogeneity of these "individual states variables" as a promising topic for future research.

Note all these fluctuations in uncertainty measures are generated solely as a consequence of endogenous attention response under dispersed information. The magnitudes of these fluctuations are determined by how much agents' attention level varies over business cycle, and the magnitude of which is closely related to the strength of the income effect on attention in the model. In this paper, I adopt an additive attention cost that is canonical in the literature. With this specification, income effect on attention are inevitably governed by the same parameter as risk aversion: both governs how much agents marginal utility from consumption and labor varies over state of the world. If income effect on attention and risk aversion are indeed closely related, the model links the magnitude of uncertainty fluctuations to the degree of risk aversion commonly obtained in the literature. If income effect on attention and risk aversion are two fundamentally distinct concepts, a more general preference specification will be essential for future research to understand attention fluctuations over the business cycle and quantify its implications. In the appendix, I suggest one way to formulate a preference that contains two parameters to separate the income effect on attention from risk aversion.

Finally, the model generates a similar persistence in all these measures compared with the data. This comes from the fact that fluctuations of uncertainty measures are tightly connected to the fluctuations of aggregate output through endogenous response in attention, information, and beliefs. By contrast, for models that generate second-moment fluctuations through exogenous shocks to volatility, they not only need to assume exogenous shocks but also an exogenous propagation to generate the persistence of these phenomena. This difference again illustrates the virtue of a parsimonious mechanism that explains these facts as integrated phenomena of the business cycle.

Fluctuations in attention: evidence from forecast survey

Although agent's attention to aggregate events is not directly observable in the data, their attention level affects their information and beliefs about macroeconomic variables, and therefore show its trace in forecast survey. Coibion and Gorodnichenko (2015) shows that people's attention level (or, more generally, the degree of information rigidity) is closely related to the regression coefficient of average forecast error on average forecast revision.

Specifically, let $f_{i,t,s} \coloneqq \mathbb{E}'_{i,t-1}[\Delta \widetilde{Y}_{t,s}]$ denote agent i's forecast of aggregate output growth between time t and s at the beginning of time t, and $\overline{f}_{t,s}$ denote the average forecast:

$$\overline{f}_{t,s} \coloneqq \int f_{i,t,s}\, di.$$

Average forecast error and forecast revision at time t about aggregate output growth at time s is given by:

$$\overline{fe}_{t,s} \coloneqq \Delta \widetilde{Y}_{t,s} - \overline{f}_{t,s}, \quad \overline{fr}_{t,s} \coloneqq \overline{f}_{t,s} - \overline{f}_{t-1,s}.$$

Coibion and Gorodnichenko (2015) consider the following regression:

$$\overline{fe}_{t,s} = \alpha + \beta_{CG}\ \overline{fr}_{t,s} + error_{t,s}.$$

To understand why coefficient β_{CG} captures agents' average attention level, consider a positive shock to productivity θ_t. A positive shock leads to a higher average forecast $\overline{f}_{t,s}$ and a positive average forecast revision as agents receive information and update their beliefs. It also leads to a positive average forecast error because the shock increases $\Delta\widetilde{Y}_{t,s}$ more than $\overline{f}_{t,s}$. The reason is that agents take into account the noise in their information, and do not fully incorporate the increase in their signal into their forecasts; as a result, average forecast $\overline{f}_{t,s}$ "underreact", and coefficient β_{CG} is positive. The larger is the noise in their signal (low attention), the less agents incorporate update their forecast, and the higher is the regression coefficient.

Coibion and Gorodnichenko (2015) documented that the regression coefficient β_{CG} is lower during NBER recession periods, and interpret it as an evidence of countercyclical attention (or more generally, a decrease in "information rigidity" during recessions).

In spirit of their finding, I compare the change in agents' information structure between model and data by estimating the regression coefficient β_{CG} separately for periods where output is above or below trend:

The resemblance of these two moments between model and data suggests that the en-

		$\tilde{Y}_t^{\dagger} > 0$	$\tilde{Y}_t^{\dagger} < 0$
β_{CG}	data	.76	.44
	model	.59	.45

dogenous fluctuations in attention in the model generates a similar state-dependent updating in beliefs as in the data. In this way, the model provides a quantitative connection between agents' belief updating behavior and the fluctuations in uncertainty measures.

Note that the fluctuations in attention in the model depends on the strength of income effect on attention, which governed by an parameter γ and will in general depend on the availability of saving technology which I abstract away in the current specification. These concerns are mainly about whether the model can generate enough fluctuations in attention. Had one introduced a richer framework that allows for a new parameter that governs the strength of income effect on attention, these regression coefficients can serve as target moments to calibrate the new parameter, and still allow one to make a quantitative statement about whether endogenous attention choice play an important role in generating fluctuations in uncertainty measures.

1.7 Concluding remark

This paper considers a production economy where agents vary their attention to macroeconomic events and receive information from dispersed sources. In times of economic downturn, agents pay more attention to aggregate economic events as they try to avoid bad decisions when their income is low. As agents pay attention, they react more to the unknown state and increases aggregate volatility. Because information is dispersed, agents' beliefs and reactions diverge as they pay attention, and they each face higher uncertainty about others' aggregate response. These phenomena happens when the economy features a high level of strategic complementarity — that is, a strong general equilibrium effect — and a low level of average attention.

I evaluate the mechanism quantitatively in a dynamic dispersed information model. When key features of the economy are disciplined by salient business cycle moments and evidence about expectations from forecast surveys, agents' countercyclical attention choice generates fluctuations in all four measures of uncertainty consistent with observation in the data.

The framework I adopt to analyze these phenomena builds on a long tradition in macroeconomics that studies how economic agents interact under dispersed information. Dispersed information economies have rich dynamic and cross-sectional implications because agents hold different beliefs about the economy and face uncertainty about others' beliefs. Yet, the richness of the environment causes difficulties in analyzing these economies. Existing works are constrained to first order approximation methods that are unable to study higher-order phenomena of these models: among these phenomena are the fluctuations of attention and uncertainty that I focus on in this paper. To understand these phenomena, I develop a perturbation method that solves arbitrary high-order approximation of a large class of dispersed information economies. The method is not only useful for understanding the phenomena studied in this paper, but also serves as a necessary and easily applicable tool to understand a broad set of mechanisms that build on dispersed information to study monetary policy, sentiment, and coordination. This method is useful for future research because it allows researchers to derive higher-moment empirical implications from existing mechanisms to test and quantify them with data. Moreover, exploring higher-order properties of these models is of crucial policy relevance because these properties are well known to have important policy implications in environments where uncertainty and difference in beliefs matter. I leave these applications to future work.

CHAPTER 2

A HIGHER-ORDER APPROXIMATION METHOD FOR DISPERSED INFORMATION ECONOMIES

2.1 Introduction

This paper develops a higher-order approximation method for dispersed information models. Dispersed information models provide explanations for prominent macroeconomic phenomena and lay grounds for important policy discussions. Early work in this framework goes back to Phelps (1970), Lucas (1972), Townsend (1983), and was revived by more recent work by Woodford (2001), Angeletos and La'o (2010), Huo and Pedroni (2020), and Angeletos and Huo (2021). The literature also overlaps with part of the rational inattention literature that study macroeconomic phenomena, started by Sims (2003) and continued by Mackowiak and Wiederholt (2009) and others. See Angeletos and Lian (2016) for a recent survey.

Despite the success of the literature in many dimensions, analysis in this literature has been constrained to "first order approximations" — either through linearization, assumption of quadratic preference, or working with a log-linear setup. First-order approximations, however, miss important features in these models. With first-order approximations, economic agents don't respond to uncertainty in models featuring strategic uncertainty, the distribution of beliefs has no role beyond its average, and attention choices are static in business cycle models.

A higher-order approximation method of dispersed information models is developed in this paper to overcome the limitations of existing first-order methods. The new method builds on the idea of perturbation technique and generalizes existing first order methods to arbitrarily higher-order approximations. The method allows researchers to characterize higher-order properties of static dispersed information models in closed form, and provides a simple algorithm for solving higher-order dynamics of dynamic dispersed information models.

The perturbation technique that the new method builds on has been widely applied in macroeconomics. The key idea of the perturbation is as follows: when there is a problem that is hard to solve, we can formulate a more general sequence of problems that contains a particular problem that we can solve easily and use it to approximate problems nearby. Judd (1998) provides an introduction to the general framework, and the approach taken by this paper is most closely related to perturbation in infinite-dimension space developed by Bhandari et al. (2018) and small shock expansion in sequence form developed by Lombardo and Uhlig (2018).

Similar to standard perturbation, the new method consists of two steps: (1) constructs a sequence of approximating economies and characterize the Taylor expansions of the equilibrium along the sequence, and (2) solving the Taylor expansions. I elaborate on each of the two steps and contrast them with existing methods below.

The first step of the new method constructs a sequence of approximating economies to approximate an equilibrium of the original economy. The sequence of economies is indexed by a "perturbation parameter" that scales the size of shocks and noises in the economies. Along the sequence of economies, we can characterize the Taylor expansions of the equilibrium by successively differentiating the equilibrium condition with respect to the perturbation parameter and evaluating them at the steady-state economy where shocks and noises vanish.

The second step of the method solves for Taylor expansions of equilibrium. Solving for these expansions demands different procedures for static and dynamic setups. Yet, both procedures exploit two special features of the problem: (1) the multi-linear structure of expansions as functions of shocks and noises, and (2) shocks and noises follows Gaussian distribution in this class of models. In static setups, these two features permit the expansions to be solved in closed form for arbitrarily high order. However, in dynamic setups, solving the expansion entail a non-trivial challenge.

The challenge stems from keeping track of economic agents' beliefs about equilibrium outcome in a dynamic setup where they all receive dispersed information. In dispersed information economies, economic agents need to form beliefs about unobservable exogenous states of the economy when making decisions. However, because agents also care about other agents' decision which they don't observe under dispersed information, they also need to form beliefs about others' beliefs, and about others' beliefs about others' beliefs, and so on. As a result, each agent enters a period with the whole "belief hierarchy" that depends on the whole history of signals they received in dynamic setups, and the relevant information to make decision may not be contained in a finite number of information states. This problem is commonly referred as an *infinite regress problem* in the literature.

In a model with infinite regress problem, an economic agent's equilibrium decision is generally a non-linear function of an infinite-dimension "individual state variable" — the whole history of signals they received. Moreover, because every agents receive different signals due to dispersed information, the equilibrium of the whole economy depends on whole cross-sectional distribution of infinite-dimension signal histories. As a result, expansions of equilibrium dynamics can potentially depend on the infinite-dimension signal histories, and solving for these expansions is a non-trivial task.

In order to make progress, the method in this paper looks for finite-dimension "dynamic factors" to summarize the infinite-dimension signal history. This key idea builds on the insight that complicated high-dimension state space can often be summarized by a few low dimension state variables that flows a simple dynamic structure — an insight emphasized by Huo and Takayama (2015), Huo and Pedroni (2020) in the dispersed information literature, and Krusell and Smith (1998) and Ahn et al. (2018) in the heterogeneous agent macroeconomics literature. This idea works particularly well within the perturbation framework: the original equilibrium may feature state variables with complex dynamics, but the Taylor expansion of equilibrium objects always features linear factors with multi-linear maps.

On the other hand, the complexity resulting from the cross-sectional heterogeneity is addressed by an aggregation result. The aggregation result builds on the two special features discussed above: (1) the multi-linear structure of expansions, and (2) Gaussian distributed shocks and noises. With these two features, the aggregation result establishes a mapping between individual variables to their aggregate counterparts, and thereby sidestep the difficulty of tracking the distribution of individual state variables.

Outline

The rest of this paper is organized as follows. Section 2.2 describes the higher order approximation method for static dispersed information models. Section 2.3 applies the method to a static example, and studies how strategic uncertainty and belief dispersion affects equilibrium behavior. Section 2.4 extends the method to general dynamic setup. Section 2.5 applies the method to a dynamic dispersed information economy with attention choice, and shows that higher order approximation captures endogenous fluctuations in attention absent under standard methods. Section 2.6 concludes.

2.2 Approximation of static dispersed information models

Consider a generic static dispersed information model with a continuum of agents indexed by i. Each agent decides individual choice variables a_i, , which affects their payoff together with aggregate variables A, and exogenous states θ, where

$$\theta \sim \mathcal{N}(0, \Sigma_\theta).$$

Each agent's information set consists of signals x_i such that

$$x_i = \theta + \epsilon_i, \quad \epsilon_i \overset{iid}{\sim} \mathcal{N}(0, \Sigma_\epsilon).$$

Suppose the equilibrium conditions of the economy can be described as a set of optimality conditions and aggregation conditions:

$$\begin{aligned} \text{optimality:} \quad & \mathbb{E}\Big[\ F(a_i, A, \theta)\ \Big|\ x_i\ \Big] = 0,\ \forall i, \\ \text{aggreation:} \quad & \int R(a_i,\ A)\ di = 0 \end{aligned}$$

where the optimality conditions hold under each agent's information set, and the aggregation conditions describe the how individual variables a_i and aggregate variables A are related.

If $F(\cdot), R(\cdot)$ are linear or linearized, the economy belongs to a class of "beauty-contest" game that has been widely studied in the literature. Yet, non-linearity is important for economic phenomena. For example, non-linearity in $F(\cdot)$ captures agents' response to uncertainty, and their potentially asymmetric response to changes in θ; non-linear in $R(\cdot)$ captures how dispersion and skewness of the distribution of individual variables affect the aggregates. These are among the most important properties for understanding macroeconomic phenomena, and are closely related to the central topics studied by the dispersed information literature: strategic uncertainty, attention allocation, and belief dispersion.

Yet, solving nonlinear models is challenging. Essentially, the equilibrium is the solution of a fixed-point problem of a non-linear functional equation. The goal is to find functions $\boldsymbol{a} : \{x_i\} \to \mathbb{R}^{|a|}$ and $\boldsymbol{A} : \ \{\theta\} \to \mathbb{R}^{|A|}$ such that

$$\mathbb{E}\Big[\ F(\boldsymbol{a}(x_i), \boldsymbol{A}(\theta), \theta)\ \Big|\ x_i\ \Big] = 0,$$

$$\int R(\boldsymbol{a}(x_i),\ \boldsymbol{A}(\theta))\ d\Phi(\epsilon_i) = 0.$$

Generally, the problem will not have a closed-form solution and we resort to approximation method.

Approximation: a perturbation approach

The approximation method studied in this paper builds on standard perturbation methods. The key idea of the perturbation method is as follows: when there is a problem that is hard to solve, formulate a more set of general problems, solve a particular problem that is easy to solve, and use it to approximate problems nearby.

Specifically, to approximate the equilibrium of an economy described above, consider a sequence of economies indexed by a parameter δ such that:

$$\theta(\delta) = \delta\ \theta, \quad \epsilon_i(\delta) = \delta\ \epsilon_i$$

This sequence of economies has the following features:

1. When $\delta = 0$, the economy contains no uncertainty, and the equilibrium can be solved easily. When $\delta = 1$, the economy is identical the one of interest.

2. The information structure of the economies is kept identical along the sequence:

$$\mathbb{E}[\, \cdot \,|x_i(\delta)] = \mathbb{E}[\, \cdot \,|x_i], \ \forall \delta.$$

The first feature allows us to take the economy with $\delta = 0$ as a starting point for approximation. The second feature keeps the sequence of approximating economies "similar" to the economy of interest by preserving the information structure.

To look for approximation of the equilibrium objects, let $a_i^\delta, \ A^\delta$ be the equilibrium of economy δ. Suppose that implicit function theorem holds for the problem at hand, then there exists mappings from δ to random variables $a_i^\delta, \ A^\delta$:

$$a_i^\delta = a_i(\delta), \ A^\delta = A(\delta).$$

A Taylor expansion with respect to δ around $\delta = 0$ gives:

$$\begin{aligned} a_i(\delta) &= \bar{a} + a_i^{(1)} \ \delta + \frac{1}{2} \ a_i \ \delta^2 + \cdots \\ A(\delta) &= \bar{A} + A^{(1)} \ \delta + \frac{1}{2} \ A \ \delta^2 + \cdots \end{aligned}$$

where $\bar{a}, \bar{A}$ are equilibrium at $\delta = 0$, and $a_i^{(m)}$, $A^{(m)}$ are derivatives with respect to δ evaluated at $\delta = 0$.

To find the first m^{th} order expansions, differentiate equilibrium conditions w.r.t. δ and evaluated at $\delta = 0$:

$$\frac{d^m}{d\delta^m}\mathbb{E}\Big[\ F(a_i(\delta), A(\delta), \theta(\delta))\ \Big|\ x_i\ \Big]\Big|_\delta = 0,$$
$$\frac{d^m}{d\delta^m}\int R(a_i(\delta),\ A(\delta))\ d\Phi(\epsilon_i)\Big|_\delta = 0.$$

Assume derivatives of $F(\cdot), R(\cdot)$ at $\delta = 0$ are known, then the m^{th} order expansion is given by

$$\bar{F}_a\ a_i^{(m)} = \mathbb{E}\Big[\ \bar{F}_A\ A^{(m)} + p\Big(\{a_i^{(l)}, A^{(l)}\}_{l<m}, \theta\Big)\ \Big|\ x_i\ \Big],$$
$$\bar{R}_A\ A^{(m)} = \int \bar{R}_a\ a_i^{(m)} + q\Big(\{a_i^{(l)}, A^{(l)}\}_{l<m}\Big)\ di$$

where $p(\cdot), q(\cdot)$ are known polynomials that depends on the first $m - 1^{th}$ order expansion.[1]

Note from the formula above, each m^{th}-order expansion of the optimality conditions consists of a linear best-response function of $a_i^{(m)}$ to its aggregate counterpart $A^{(m)}$ and a "modified fundamental" $p\big(\{a_i^{(l)}, A^{(l)}\}_{l<k}, \theta\big)$ that captures higher-order concerns. On the other hand, each m^{th}-order expansion of the aggregation conditions shows that the m^{th}-order aggregate variables are results of linear aggregation of m^{th} order individual variables, and the higher-order properties of $R(\cdot)$ is captured by

1. The polynomials are given by Faà di Bruno's formula.

$q\big(\{a_i^{(l)}, A^{(l)}\}_{l<k}, \theta\big).$

To solve for expansion sequences, write

$$a_i(\delta) = \boldsymbol{a}(x_i(\delta), \delta), \quad \text{where } x_i(\delta) = \epsilon_i(\delta).$$

Take derivative with respect to δ and evaluate at $\delta = 0$, we have the following expressions

$$\begin{aligned}
a_i^{(1)} &= \frac{d}{d\delta}\boldsymbol{a}(x_i(\delta), \delta)\Big|_{\delta=0} = \bar{\boldsymbol{a}}_x \; x_i \; + \; \bar{\boldsymbol{a}}_\delta, \\
a_i &= \frac{d^2}{d\delta^2}\boldsymbol{a}(x_i(\delta), \delta)\Big|_{\delta=0} = \; x_i^\intercal \; \bar{\boldsymbol{a}}_{xx} \; x_i \; + \bar{\boldsymbol{a}}_{x\delta} \; x_i \; + \bar{\boldsymbol{a}}_{\delta\delta}, \\
&\vdots
\end{aligned}$$

where $\bar{\boldsymbol{a}}_x, \bar{\boldsymbol{a}}_\delta, \bar{\boldsymbol{a}}_{xx} \ldots$ are partial derivatives of $\boldsymbol{a}(\cdot, \cdot)$ at $\delta = 0$. In this way, the problem reduces to solving real-valued coefficients $\bar{\boldsymbol{a}}_x, \bar{\boldsymbol{a}}_\delta, \bar{\boldsymbol{a}}_{xx}, \ldots$.

Notice that the approximation described above has the following features:

1. unknown coefficients $\bar{\boldsymbol{a}}_{x\ldots\delta}$ for the m^{th} order expansion enters linearly in the problem,

2. $p(\cdot)$ and $q(\cdot)$ are known polynomials, and $\{a_i^{(l)}, A^{(l)}\}_{l<m}$ are multi-linear forms of Gaussian random signals x_i.

These features imply the following result:

Lemma 4

The m^{th} order coefficients can be solved recursively in closed form from the expanded equilibrium conditions, given the solution of first $m-1^{th}$ order expansion.

2.3 Application: a static example

Consider an economy with a continuum of agents indexed by i, each of them can produce y_i units of good i with cost

$$\nu(y_i) = y_i^{1+\nu}.$$

The demand for good i is given by inverse demand

$$p(\theta, Y, y_i) = e^{\theta} Y^{\alpha} y_i^{-\beta},$$

where θ is an exogenous aggregate state, and Y denotes an aggregate of individual production

$$Y = \Big(\int y_i^{1-\eta}\ di\ Big)^{\frac{1}{1-\eta}}.$$

Agents derive utility from consumption:

$$u(c_i) = \frac{1}{1-\gamma} c_i^{1-\gamma},$$

and their budget constraint is given by:

$$c_i = p(y_i, Y, \theta)y_i - \nu(y_i).$$

Aggregate state θ is unknown to the agents when they decide on production y_i. Agents have common prior

$$\theta \sim N(0, \sigma^2),$$

and each of them receives a private signal x_i before they make decisions:

$$x_i = \theta + \epsilon_i, \quad \epsilon_i \sim N(0, \sigma_\epsilon^2).$$

An equilibrium consists of the aggregation condition and the optimality

$$c'(y_i) = \mathbb{E}_i\Big[\ SDF_i \times \frac{\partial}{\partial y_i} p(y_i, Y, \theta)y_i\ \Big], \quad \text{where} \quad SDF_i = \frac{u'(c_i)}{\mathbb{E}_i[u'(c_i)]}.$$

Let $\{\log y_i, \log Y\}$ correspond to $\{a_i, A\}$ in the general setup, and consider Taylor expansion

$$\log y_i(\delta) = \bar{y} + y_i^{(1)}\ \delta + \frac{1}{2}\ y_i\ \delta^2 + \cdots$$
$$\log Y(\delta) = \bar{Y} + Y^{(1)}\ \delta + \frac{1}{2}\ Y\ \delta^2 + \cdots$$

With the approximation method described above, the 1^{st} expansion of equilibrium conditions is given by:

$$y_i^{(1)} = \mathbb{E}[\ s\ Y^{(1)} + r\ \theta\ |x_i], \quad Y^{(1)} = \int y_i^{(1)},$$

where

$$s := \frac{\alpha}{\beta + \nu}, \quad r := \frac{1}{\beta + \nu}.$$

The 1^{st} expansion leads to exactly the same type of equilibrium conditions as beauty-contest games: a linear best-response function and aggregation.

To capture the 2^{nd}-order properties of the equilibrium, we expand the equilibrium conditions to the 2^{nd} order:

$$\begin{aligned} y_i &= \mathbb{E}[\ s\ Y\ |x_i] - (\tilde{\gamma} - 1)\ Var(\ s\ Y^{(1)} + r\ \theta\ |x_i), \\ Y &= \int y_i + (1 - \eta)\ Var(\ y_i^{(1)}\ |\theta), \end{aligned}$$

where $\tilde{\gamma} = \gamma \times const.$

Combing the two equilibrium conditions, we have

$$Y = \bar{\mathbb{E}}[\ s\ Y\ |x_i] - \underbrace{\tilde{\gamma}\ Var(\ s\ Y^{(1)} + r\ \theta\ |x_i)}_{\text{uncertainty}} - \underbrace{\eta\ Var(\ y_i^{(1)}|\theta)}_{\text{belief dispersion}} + \underbrace{Var(\ s\ Y^{(1)} + r\ \theta\ )}_{\text{log-normal correction}}$$

The 2^{nd} order expansion of equilibrium conditions has a similar form as the 1^{st} order. There are three difference. First, the optimality condition now captures how agents' response to uncertainty given their risk aversion γ. Second, the effect of belief dispersion is captured by the aggregation condition, and depends on the non-linearity of the aggregation technology η. Finally, the log-normal correction is picked up by the 2^{nd} order expansion. All these three 2^{nd} order effects are amplified by the strategic complementary of in the economy: agents not only respond to these adjustments, but also how others respond to them, how others respond to others' response, and so on.

2.4 Approximation of dynamic dispersed information models

Consider a generic dynamic dispersed information model with a continuum of agents indexed by i and time indexed by t. Each agent decides individual control variables $a_{i,t}$ in each period, which affects their payoff together with aggregate variables A_t, and exogenous states θ_t that follows:

$$\theta_{t+1} = \Gamma\theta_t + \omega_{t+1}, \quad \omega_t \overset{iid}{\sim} \mathcal{N}(0, I_{|\omega|}).$$

Furthermore, there are individual and aggregate endogenous state variables $k_{i,t}, K_t$ that describes restrictions on agents' choice set.

Suppose the equilibrium conditions of the economy can be described as a set of optimality conditions, aggregation conditions, and law of motion:

$$\begin{aligned}
\text{optimality:} \quad & \mathbb{E}\Big[\, F(\{a_{i,s}, A_s, \theta_s\}_{s\geq t}) \;\Big|\; \mathcal{F}_{i,t} \,\Big] = 0, \\
\text{aggregation:} \quad & \int R(a_{i,t}, A_t, k_{i,t-1}, K_{t-1})\; di = 0, \\
\text{law of motion:} \quad & k_{i,t} - G(a_{i,t}, A_t, k_{i,t-1}, K_{t-1}, \theta_t) = 0,
\end{aligned}$$

where $\mathcal{F}_{i,t}$ denotes the information set of agent i at time t generated by signals x_i^t:

$$\begin{aligned}
\mathcal{F}_{i,t} &= \boldsymbol{\sigma}\big(\{x_{i,s}\}_{s\leq t}\big), \\
x_{i,t} &= H(a_{i,t-1}, A_{t-1}, k_{i,t-1}, K_{t-1}, \theta_t) + \epsilon_{i,t}, \quad \epsilon_{i,t} \overset{iid}{\sim} \mathcal{N}\big(0, \Sigma)\big).
\end{aligned}$$

Moreover, $a_{i,t}$ is subject to information constraint:

$$a_{i,t} \in \mathcal{F}_{i,t}.$$

In the next section, I will illustrate the approximation method with the following simplifying assumptions (1) $G(\cdot) \equiv 0$, so there is no state variable $k_{i,t}$, (2) $a_{i,t}, A_t, \theta_t, x_{i,t}$ are all one-dimensional, and (3) agents' information structure takes a simple form

$$x_{i,t} = \theta_t + \epsilon_{i,t},\ \ \epsilon_{i,t} \sim \mathcal{N}\big(0, \Sigma\big).$$

After laying out how the method works in this special case, I will then demonstrate how the method extends to the general setup.

Equilibrium approximation: challenge and strategy

Different from the static economy we studied before, in a dynamic setup, agents enter each period with beliefs that depends on the information they observed in the past. If the only thing agents are concerned about is the exogenous state θ, in order to track an agent's belief about θ_t, we only need $\mathbb{E}_{i,t-1}[\theta_t]$ to summary signal history, which is given by the standard Kalman filter. However, when agents' decisions also depends on endogenous aggregates A_t, they not only need to predict the signals other agents receive this period, but also form beliefs about others' beliefs when they enter the period, about other agents' beliefs about others' beliefs, and so on. In principle, the whole belief hierarchy matters, and it is possible that an agents' information set x_i^t cannot be summarized by a finite dimension state space, and the equilibrium of the economy can potentially depends on the whole distribution across agents over their individual histories of signals.

The perturbation method introduced in the previous section provides a useful strategy to approximate the equilibrium of this economy. Consider a sequence of economies indexed by a parameter δ:

$$\omega_t(\delta) = \delta\ \omega_t, \quad \epsilon_{i,t}(\delta) = \delta\ \epsilon_{i,t}.$$

Because $x_{i,t} = \theta_t + \epsilon_{i,t},\ \epsilon_{i,t} \sim \mathcal{N}\big(0, \Sigma\big)$, the information structure does not change along the expansion path:

$$\mathbb{E}[\ \cdot\ |\mathcal{F}_{i,t}(\delta)] = \mathbb{E}[\ \cdot\ |\mathcal{F}_{i,t}], \forall \delta.$$

As before, we look for Taylor expansion with respect to δ around $\delta = 0$

$$a_{i,t}(\delta) = \bar{a} + a_{i,t}^{(1)}\ \delta + \frac{1}{2}\ a_{i,t}\ \delta^2 + \cdots$$
$$A_t(\delta) = \bar{A} + A_t^{(1)}\ \delta + \frac{1}{2}\ A_t\ \delta^2 + \cdots$$

and the equilibrium conditions

$$\mathbb{E}\Big[\ \sum_{s=0}^{\infty}\Big(\bar{F}_{a,s}\ \tilde{a}_{i,t+s}^{(m)} + \bar{F}_{A,s}\ \tilde{A}_{t+s}^{(m)}\Big) + p\Big(\big\{a_{i,t+s}^{(l)}, A_{t+s}^{(l)}, \theta_{t+s}\big\}_{l<m,s\geq 0}\Big)\ \Big|\ \mathcal{F}_{i,t}\ \Big] = 0,$$

$$\int \bar{R}_a\ \tilde{a}_{i,t}^{(m)} + \bar{R}_A\ \tilde{A}_t^{(m)} + q\Big(\big\{a_{i,t}^{(l)}, A_t^{(l)}, \theta_t\big\}_{l<m}\Big)\ di = 0$$

So far, this is all very similar to the procedure for static models. The main difference appears when we solve for expansions $a_{i,t}^{(1)}, A_t^{(1)} ...\ a_{i,t}^{(m)}, A_t^{(m)}$.

When approximating a static model, the Taylor expansion of equilibrium can be written as

$$\boldsymbol{a}(x_i(\delta), \delta) = \bar{a} + (\boldsymbol{a}_x\ x_i + \boldsymbol{a}_\delta)\ \delta + \frac{1}{2}(x_i^\intercal\ \boldsymbol{a}_{xx}\ x_i + \cdots)\ \delta^2 + \cdots,$$

and $\boldsymbol{a}_x,\ \boldsymbol{a}_{xx} \ldots$ are finite-dimension multi-linear maps that can be easily solved from the equilibrium conditions.

In a dynamic setup, an agents' decision potentially depends on the whole history of signals x_i^t:

$$\boldsymbol{a}(x_i^t(\delta), \delta) = \bar{a} + (\boldsymbol{a}_x \ x_i^t + \boldsymbol{a}_\delta) \ \delta + \frac{1}{2}(x_i^{t\intercal} \ \boldsymbol{a}_{xx} \ x_i^t + \cdots) \ \delta^2 + \cdots ,$$

and $\boldsymbol{a}_x, \ \boldsymbol{a}_{xx}, \ldots$ are all infinite dimensional objects that need to be solved for.

One way to make progress is to find finite-dimension "approximate states" to summary the signal histories.

For each order of expansion, we can approximate the history of signals x_i^t at time t and the associated coefficients $a_x, a_{xx}, \ldots$ by n-dimensional "dynamic factors" $f_{i,t}^{(m)}$ and coefficients ϕ's such that:

$$1^{st}\text{order:} \quad \boldsymbol{a}_x \ x_i^t \approx \phi_f f_{i,t}^{(1)}$$

$$2^{nd}\text{order:} \quad x_i^{t\intercal} \ \boldsymbol{a}_{xx} \ x_i^t \approx f_{i,t}^{(2)\intercal} \ \phi_{ff} \ f_{i,t}^{(2)}, \quad \boldsymbol{a}_{x\delta} \ x_i^t \approx \phi_{f\delta} \ f_{i,t}^{(2)},$$

and so on, where factors for each m^{th} order expansion follow a linear structure:

$$f_{i,t+1}^{(m)} = B^{(m)} \ f_{i,t}^{(m)} + D^{(m)} \ x_{i,t+1},$$

and the m^{th} order terms $\tilde{a}_{i,t}^{(m)}$ is approximated as

$$\tilde{a}_{i,t}^{(m)} = \sum_{l=0}^{m} \phi_{f^l \delta^{m-l}}\Big(f_{i,t}^{(m)}\Big),$$

where $\phi_{f^l \delta^{m-l}} : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ is a symmetric l-linear map.

A combination of linear factors, multi-linear structure of expansions, together with Gaussian distributed shocks and noises give rise to two useful results.

The first result is an aggregation result. From the expansion of aggregation condition, we have

$$\tilde{A}_t^{(m)} = \Big(-\bar{R}_A\Big)^{-1} \int \bar{R}_a \; \tilde{a}_{i,t}^{(m)} + q\Big(\{a_{i,t+s}^{(l)}, A_t^{(l)}, \theta_t\}_{l<m}\Big) \; di.$$

Let $\phi^{(m)}$ denote the collection of coefficients of all l-linear maps $\phi_{f^l \delta^{m-l}}$ for $l = 0, \ldots, m$, and let $f_t^{(m)}$ denote the aggregation of individual factors $f_{i,t}^{(m)}$

$$f_{t+1}^{(m)} := \int f_{i,t}^{(m)} \; di,$$

$$f_{t+1}^{(m)} = B^{(m)} \; f_t^{(m)} + D^{(m)} \; \theta_{t+1}.$$

Lemma 5

Approximated aggregates $\tilde{A}_t^{(m)}$ can be expressed as known functions of coefficients $\{\phi^{(l)}\}_{l\le m}$ and aggregate factors $\{f_t^{(l)}\}_{l\le m}$. In particular, it is linear in $\phi^{(m)}$.

For example, with a standard constant elasticity aggregator:

$$A_t = \left(\int a_{i,t}^{1-\eta} \right)^{\frac{1}{1-\eta}},$$

the 1^{st} and 2^{nd} order expansions of aggregates are given by

$$\tilde{A}_t^{(1)} = \int a_{i,t}^{(1)} \; di = \phi_f f_t^{(1)},$$

and

$$\begin{aligned} \tilde{A}_t^{(2)} &= \int \tilde{a}_{i,t}^{(2)} + (1-\eta) \left(\tilde{a}_{i,t}^{(1)} - \tilde{A}_t^{(1)} \right)^2 \; di \\ &= f_t^{(2)\intercal} \phi_{ff} f_t^{(2)} + \phi_{f\delta} f_t^{(2)} + \phi_{\delta\delta} + (1-\eta) tr\big(\; v^{(2)} \; \phi_{ff} \; \big), \end{aligned}$$

where $v^{(2)}$ is the solution of matrix equation

$$v^{(2)} = B^{(2)\intercal} \; v^{(2)} \; B^{(2)} + D^{(2)\intercal} \Sigma \; D^{(2)}.$$

Building on the aggregation result, we can express the expectation of any m^{th} order expansion as function of linear factors:

Lemma 6

Expectation of approximated expansion sequence can be expressed as known functions *of* $\{f_{i,t}^{(l)}, \forall l \leq m\}$*:*

$$\mathbb{E}_{i,t}[\tilde{a}_{i,t+s}^{(m)}] = a^{(m)}(\phi^{(m)}, f_{i,t}^{(m)}, s), \quad \mathbb{E}_{i,t}[\tilde{A}_{t+s}^{(m)}] = A^{(m)}(\{\phi^{(l)}, f_{i,t}^{(l)}\}_{l\leq m}, s), \forall s.$$

Moreover, $a^{(m)}(\cdot), A^{(m)}(\cdot)$ *are linear in coefficients* $\phi^{(m)}$*.*

Lemma 5 and lemma 6 are useful for the computation of equilibrium approximation. These results address the difficulty of cross-sectional heterogeneity that comes from dispersed information. Underlying the result is the recurring insight in macroeconomics: agents in the economy do not care about the complicated cross-sectional distribution per se, what they care about are some particular aggregate variables. The combination of linear factors, the multi-linear structure of expansion sequence, and Gaussian distributed shocks and noises together allow us to express the objects in agents optimality conditions as functions of the factors summarizing their past information. The next section builds on these results, and provide an algorithm to approximate the equilibrium.

Algorithm

Given an order of approximation m, the optimal factors (within the class specified above) and the associated coefficients can be solved recursively given the approximated solution for the first $m-1$ order, $\{a_{i,t}^{(l)}, A_t^{(l)}\}_{l<m}$.

The procedure consists of two layers of optimization:

Outer loop:

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

$$B^{(m)} \in \mathbb{R}^{n \times n}, D^{(m)} \in \mathbb{R}^{n \times |x|}$$

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

(ii) Solve for coefficients ϕ's that minimize sum of squared residuals of equilibrium conditions, $r^*(B^{(m)}, D^{(m)})$ with the procedure described below for the inner loop.

(iii) Solve for $B^{(m)}, D^{(m)}$ that minimize $r^*(B^{(m)}, D^{(m)})$.

Inner loop:

(i) Given $B^{(m)}, D^{(m)}$, simulate $\{\omega_t, \epsilon_{i,t}\}_{t=0}^{T}$. Construct a history of signals $x_{i,t}$ and the associated factors $f_{i,t}$:

$$f_{i,t+1}^{(m)} = B^{(m)}\ f_{i,t}^{(m)} + D^{(m)}\ x_{i,t+1}, \quad f_{i,0}^{(m)} = 0.$$

(ii) Find coefficients ϕ's to minimize the sum of squares residuals of the expanded equilibrium conditions

$$r^* := \sum_t residual_{i,t}^2,$$

where

$$residual_{i,t} = \mathbb{E}\Big[\sum_{s=0}^{\infty} \Big(\bar{F}_{a,s}\ \tilde{a}_{i,t+s}^{(m)} + \bar{F}_{A,s}\ \tilde{A}_{t+s}^{(m)}\Big) + p\Big(\{a_{i,t+s}^{(l)}, A_{t+s}^{(l)}, \theta_{t+s}\}_{l<m,s\geq 0}\Big) \ \Big|\ \mathcal{F}_{i,t} \ \Big],$$

and

$$\mathbb{E}_{i,t}[\tilde{a}_{i,t+s}^{(m)}] = a^{(m)}(\phi^{(m)}, f_{i,t}^{(m)}, s), \quad \mathbb{E}_{i,t}[\tilde{A}_{t+s}^{(m)}] = A^{(m)}(\phi^{(m)}, f_{i,t}^{(m)}, s).$$

Remark

Notice that in the procedure described above, solving coefficients ϕ's in the inner loop involves only quadratic optimization, which can be solved efficiently. The computationally costly step is when we solve for $B^{(m)}$ and $D^{(m)}$ in the outer loop.

However, note that factors $f_{i,t}$ can only be identified up to linear combinations. For example, given any Ω, $\tilde{f}_{i,t} \coloneqq \Omega f_{i,t}$ spans the same space as $f_{i,t}$, and follows

$$\tilde{f}_{i,t+1} = \tilde{B}^{(m)}\tilde{f}_{i,t} + \tilde{D}^{(m)}x_{i,t+1}, \quad \tilde{B}^{(m)} \coloneqq \Omega B^{(m)}\Omega^{-1}, \quad \tilde{D}^{(m)} \coloneqq \Omega D^{(m)}.$$

As a result, we can restrict to matrices in Jordan form when searching for $B^{(m)}$. This reduces the number of variables that involves non-linear optimization to order $\mathcal{O}(n)$, where n is the number of factors.

Extension: endogenous state variables

The method extends easily to setups where there are endogenous state variables with law of motions

$$k_{i,t} - G(a_{i,t}, A_t, k_{i,t-1}, K_{t-1}, \theta_t) = 0.$$

As agents choose $a_{i,t}$ subject to information constraints, endogenous state variables $k_{i,t}$ absorbs the shocks:

$$k_{i,t} = \boldsymbol{k}(x_i^t, \theta^t).$$

The same algorithm still applies, except that we will need more factors to tack the effect of shocks. For example, the 1^{st} order expansion of state variable $k_{i,t}$ will be approximated as:

$$k_{i,t}^{(1)} = \phi_k \begin{pmatrix} f_{i,t}^{(1)} \\ g_t^{(1)} \end{pmatrix}, \text{ where } g_{t+1}^{(1)} = B_k^{(1)} \; g_t^{(1)} + B_k^{(1)} \; \theta_t.$$

Extension: information about endogenous variables

Consider the general case where signals involve endogenous variables:

$$x_{i,t} = H(a_{t-1}, A_{t-1}, \theta_t) + \epsilon_{i,t}, \quad \epsilon_{i,t} \sim \mathcal{N}\big(0, \Sigma\big).$$

The approximation procedure works similarly in this setup, except one important difference: agents' information set now changes with endogenous variables along the sequence of economies

$$\mathbb{E}[\,\cdot\,|\mathcal{F}_{i,t}(\delta)] \neq \mathbb{E}[\,\cdot\,|\mathcal{F}_{i,t}(0)],\ \forall\delta \neq 0.$$

When agents' information structure changes along the sequence of economies, these changes are reflected in the expansion of equilibrium. Changes in information structure is captured by extra terms in the expansion of equilibrium conditions. For example, for the m^{th} order expansion, one needs to compute

$$\frac{d^l}{d\delta^l}\mathbb{E}\Big[\ p^{(m-l)}\Big(\{a_{i,s}^{(m-l)}, A_s^{(m-l)}, \ldots, \theta_s\}_{s\geq t}\Big)\ \Big|\mathcal{F}_{i,t}(\delta)\Big]\ \Big|_{\delta=0}, \forall l < m.$$

Because agents' information set contains non-linear signals, the derivatives above is difficult to calculate in general. One way to calculate these derivatives is to calculate the expectations for a small $\bar{\delta}$ and use finite difference to calculate the derivatives. The expectation can be approximated with standard nonlinear filter, such as particle filter, and signals with endogenous variables truncated at the m^{th} order expansion:

$$\mathbb{E}\Big[\ p^{(m-l)}\Big(\{a_{i,s}^{(l)}, A_s^{(l)}, \ldots, \theta_s\}_{s\geq t}\Big)\ \Big|\mathcal{F}_{i,t}(\bar{\delta})\Big]$$

with signals

$$x_{i,t}^{(m)}(\bar{\delta}) = H\Big(\bar{a} + \bar{\delta}\ a_{i,t-1}^{(1)} + \cdots + \frac{\bar{\delta}^m}{m!}\ a_{i,t-1}^{(m)},$$
$$\bar{A} + \bar{\delta}\ A_{t-1}^{(1)} + \cdots + \frac{\bar{\delta}^m}{m!}\ A_{t-1}^{(m)},\ \bar{\delta}\ \theta_t\Big) + \bar{\delta}\ \epsilon_{i,t}.$$

The use of particle filter is computationally intensive in general, as it involves Monte Carlo simulations. However, the derivatives can be calculated efficiently when signals are linear conditional on past individual choice variables:

$$x_{i,t} = H(a_{i,t-1})\theta_t + \epsilon_{i,t}, \sim \mathcal{N}\big(0, \Sigma\big).$$

When signals are linear conditional on past individual choice variables, standard Kalman filter allows us to express the expectation recursively as a function of past expectations, current signals, and choice variables, which are approximated multi-linear maps of autoregressive factors. As a result, we can derive recursion for the derivatives of expectations can avoid computationally costly Monte Carlo simulation.

2.5 Application: a dynamic example

Consider an economy similar to the static example we studied before. There is a continuum of agents indexed by i. Agents are infinitely-lived, and time is discrete.

Each of them can produce y_i units of good i with cost

$$\nu(y_{i,t}) = y_{i,t}^{1+\nu}.$$

The demand for good i is given by inverse demand

$$p(\theta_t, Y_t, y_{i,t}) = e^{\theta_t} Y_t^{\alpha} y_{i,t}^{-\beta},$$

where θ_t is an exogenous aggregate state that follows

$$\theta_{t+1} = \rho\theta_t + \omega_{t+1}, \quad \omega_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\omega^2),$$

and Y_t denotes an aggregate of individual production

$$Y_t = \left(\int y_{i,t}^{1-\eta} \, di \right)^{\frac{1}{1-\eta}}.$$

Aggregate state θ_t is unknown to the agents. In each period, agents chooses how much "attention", $z_{i,t-1}$, they would like to pay to the aggregate state, and observe a private signal $x_{i,t}$:

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

Agents choose production $y_{i,t}$ and attention subject to information constraints:

$$\begin{aligned} z_{i,t-1} &\in \mathcal{F}_{i,t-1} = \boldsymbol{\sigma}\big(\{x_{i,s}\}_{s\leq t-1}\big), \\ y_{i,t} \quad &\in \mathcal{F}_{i,t} = \boldsymbol{\sigma}\big(\{x_{i,s}\}_{s\leq t}\big), \end{aligned}$$

and they derive utility from their consumption and disutility from attention:

$$u_i = \sum \beta^t \Big(\frac{1}{1-\gamma} c_{i,t}^{1-\gamma} - \kappa(z_{i,t})\Big).$$

Agents are subject to period-by-period budget constraint:

$$c_{i,t} = p(y_{i,t}, Y_t, \theta_t) y_{i,t} - \nu(y_{i,t}).$$

An equilibrium consists of the aggregation condition and the following optimality conditions for attention and production choice:

$$[z_{i,t-1}] : \kappa'(z_{i,t-1}) = \mathbb{E}\Big[\ \frac{1}{2\ z_{i,t-1}}\Big(\epsilon_{i,t}^2 - 1\Big)\ \sum_{s\geq t} \beta^{s-t+1}\Big(\frac{1}{1-\gamma} c_{i,s}^{1-\gamma} - \kappa(z_{i,s})\Big)\ \Big| \mathcal{F}_{i,t}\Big]$$

$$[y_{i,t}] : \quad \nu'(y_i) = \mathbb{E}_i\Big[\ SDF_i \times \frac{\partial}{\partial y_i} p(y_i, Y, \theta) y_i\ \Big], \quad SDF_i = \frac{u'(c_i)}{\mathbb{E}_i[u'(c_i)]}.$$

To approximate the equilibrium, consider the method described in the previous section with the following modification of the approximating sequence of economies:

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

Along the sequence of economies, the attention cost is scaled at rate δ^2. The reason for this modification is because the value of information is 2^{nd} order, and scaling the attention cost accordingly allows us to approximate the economy around the bifurcation point of the sequence.[2] With the modification, the 1^{st} order expansions of production $y_{i,t}^{(1)}, Y_t^{(1)}$ is jointly determined with the steady-state attention level (the 0^{th} order expansion of attention) — a structure identical to the standard dynamic beauty contest setup and the 1^{st} order expansion of an economy with exogenous information. Going higher order, the m^{th} order expansions of production $y_{i,t}^{(m)}, Y_t^{(m)}$ is jointly determined with the $m - 1^{th}$ order expansions of attention. Despite these differences, the approximation method follows the steps described in the previous section closely, and the result is described below.

Approximation result

1^{st} *order approximation*

With 1^{st} order approximation, the economy features a static attention level $\bar{z}$, and

2. See Judd 1996 for detail.

the history of signals can be summarized by a one dimensional factor

$$f^{(1)}_{i,t+1} = B^{(1)} f^{(1)}_{i,t} + D^{(1)} \bar{x}_{i,t+1}, \quad \text{where} \quad \bar{x}_{i,t} = \theta_t + \frac{\epsilon_{i,t}}{\sqrt{\bar{z}}},$$

and $B^{(1)}, D^{(1)} \in \mathbb{R}$. Moreover, $Y^{(1)}_t$ follows an AR(2) process:

$$Y^{(1)}_t = \phi_f f_t, \quad \text{where} \quad f^{(1)}_{t+1} = B^{(1)} f^{(1)}_t + D^{(1)} \theta_{t+1}.$$

The 1^{st} order expansion of the equilibrium conditions is a set of linear restrictions that can be solved by analytical methods developed by Huo and Takayama (2015) and Huo and Pedroni (2020). The computational method described in this paper gives the identical result.

2^{nd} *order approximation*

While existing methods are constrained to 1^{st} order approximation, the method described in the previous section extends to higher order. Higher order expansions capture features of the model 1^{st} order methods miss — in this example, the state-dependent variation of attention choice.

Figure 2.1 shows squared residuals of expanded optimality conditions for production and attention choice respectively.[3] The residuals are in log points of steady-state production and attention. The equilibrium is approximated respectively with 0 (*blue*),

3. In order to find the 2^{nd} and 1^{st} order expansion of production and attention, the corresponding equilibrium conditions are expanded to 2^{nd} and 3^{rd} order respectively.

3 (red), and 5 ($yellow$) factors, where factors are extracted from a simulation of 2000 periods.

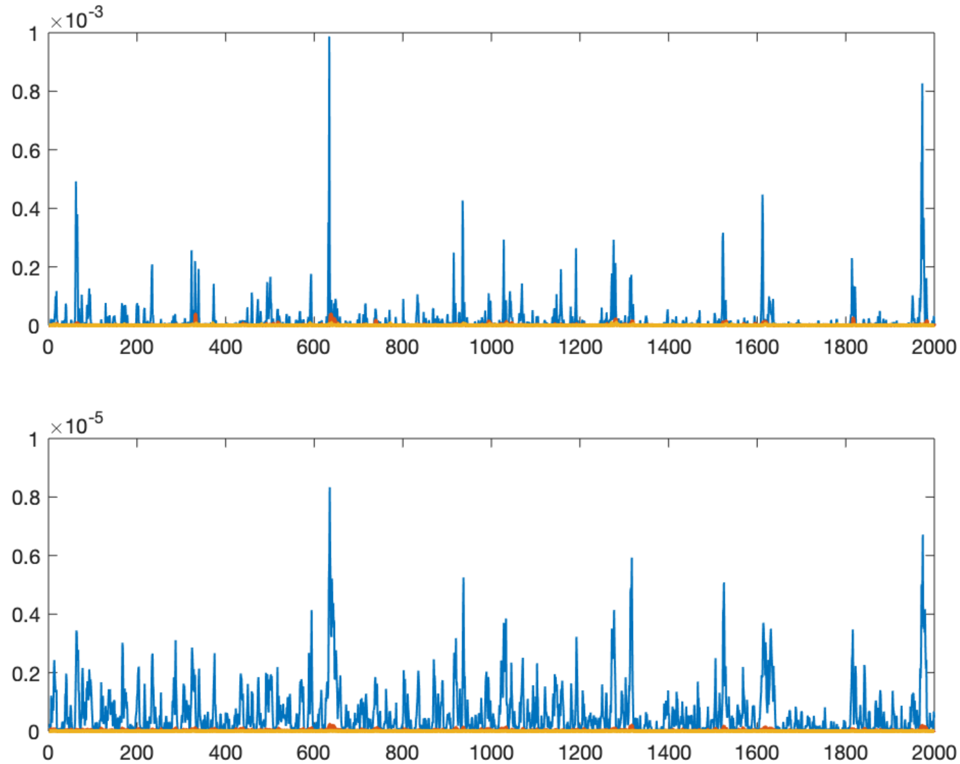


Figure 2.1: Squared residuals of optimality conditions expansion

Top panel: squared residuals of the expansion of labor optimality in each period of simulation. Bottom panel: squared residuals of the expansion of attention optimality in each period of simulation.

Figure 2.2 shows the 0^{th} and 1^{st} order approximation of equilibrium attention level. While the 0^{th} order expansion of attention, $\bar{z}$, is jointly determined with the 1^{st} order expansion of production, $y_{i,t}^{(1)}$, it does not capture any fluctuation in equilibrium attention. By contrast, the 2^{nd} order expansion of production together with the 1^{st} order expansion of production captures the dynamics attention as equilibrium attention responds to fluctuations of the aggregate state.

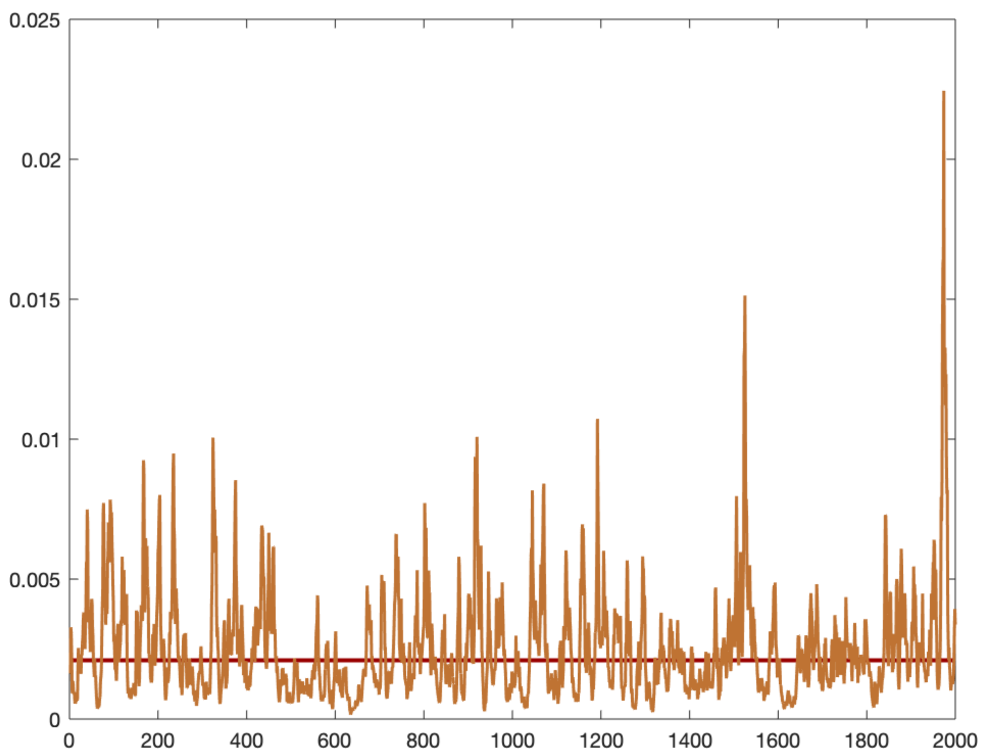


Figure 2.2: Approximated equilibrium attention

The red and brown line show respectively the 0^{th} and 1^{st} order approximation of equilibrium attention level.

2.6 Concluding remark

This paper develops a high-order approximation method of a large class of dispersed information economies. I demonstrate how standard perturbation methods can be applied to analyse dispersed information economies, in particular, how it provides a way to characterize higher-order properties of the model in the presence of infinite regress problem where existing methods are constrained to first order approximations.

The paper illustrate the method with a few simple examples. Yet, the method is by no means restricted to these setup — it can be easily applied to understand a broad set of mechanisms that build on dispersed information to study monetary policy, sentiment, and coordination. This method is useful for future research because it allows researchers to derive higher-moment empirical implications from existing mechanisms to test and quantify them with data. Moreover, exploring higher-order properties of these models is of crucial policy relevance because these properties are well known to have important policy implications in environments where uncertainty and difference in beliefs matter. I leave these applications to future work.

APPENDIX A

DERIVATIONS AND PROOFS

A.1 Static model

Formally, assume agents have preference

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

where

$$v_i = \max\{c_i - \nu(n_i), \underline{c}\}, \quad \nu(n_i) = \frac{\psi}{1+\nu} n_i^{1+\nu},$$

and, from the budget constraint and the CES demand, consumption c_i is given by

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

I assume that agents' utility function has a cutoff $\underline{c}$, which is a small number such that $c_i - \nu(n_i) > \underline{c}$ in the equilibrium without uncertainty. This ensure v_i is positive for any realization of θ and the problem is well-defined.

The first order condition is given by

$$\mathbb{E}_i \left[v_i^{-\gamma} \left(\frac{\partial}{\partial n_i} c_i - \nu'(n_i) \right) \mathbf{1}_{\{v_i \geq \underline{c}\}} \right] = 0.$$

Dividing both sides by $R_i \coloneqq \mathbb{E}_i[v_i^{-\gamma}\mathbf{1}_{\{v_i \geq \underline{c}\}}]^{\frac{1}{-\gamma}}$ gives

$$\mathbb{E}_i\left[\left(\frac{v_i}{R_i}\right)^{-\gamma}\left(\frac{\partial}{\partial n_i}c_i - \nu'(n_i)\right)\mathbf{1}_{\{v_i \geq \underline{c}\}}\right] = 0$$

Approximation

Consider a sequence of economy parameterize by δ such that

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

The equilibrium in economy δ is described by the F.O.C. of attention and labor input, aggregation, and the symmetry of attention:

$$[n_i] : \mathbb{E}\left[v_i(\delta)^{-\gamma}\left(\frac{\partial}{\partial n_i}c_i(\delta) - \nu'(n_i(\delta))\right)\mathbf{1}_{\{v_i(\delta) \geq \underline{c}\}}\middle| x_i(\delta)\right] = 0,$$

$$[Agg] : \log N(\delta) = \frac{1}{1-\eta}\log\left(\int \exp((1-\eta)\log n_i(\delta))\right),$$

$$[z_i] : \mathbb{E}\left[-\frac{\epsilon_i^2 - 1}{2z(\delta)}\frac{1}{1-\gamma}v_i(\delta)^{1-\gamma}\middle|\delta\right] - \delta^2\kappa(z) = 0$$

and

$$\log v_i(\delta) = \log\left(\max\{c_i(\delta) - \nu(n_i(\delta)), \underline{c}\}\right).$$

Assume the solutions have Taylor expansions

$$\log n_i(\delta) = \bar{n} + \hat{n}_i\delta + \delta^2(...)$$
$$\log N(\delta) = \bar{N} + \hat{N}\delta + \delta^2(...)$$
$$\log z(\delta) = \log \bar{z} + \hat{z}\delta + \delta^2(...)$$
$$\log v_i(\delta) = \bar{v} + \hat{v}_i\delta + \delta^2(...)$$

0^{th} order expansion

The 0^{th} order expansion of equilibrium conditions gives

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

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 $\bar{n}, \bar{N}, \bar{v}$. Yet, the attention level $\bar{z}$ is not pinned down because

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

1^{st} order expansion

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

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

where

$$
\bar{x}_i = \theta + \frac{1}{\sqrt{\bar{z}}}\epsilon_i, \quad \hat{v}_i = \boldsymbol{v}_\delta + \boldsymbol{v}_\theta \hat{\theta} + \boldsymbol{v}_n \hat{n}_i.
$$

Write

$$
\begin{aligned}
\log n_i(\delta) &= \boldsymbol{n}(\ x_i(\delta),\ \delta) \\
\log N(\delta) &= \boldsymbol{N}(\ \theta(\delta),\ \delta) \\
\log v_i(\delta) &= \boldsymbol{v}(\ \theta(\delta),\ \log n_i(\delta),\ \delta)
\end{aligned}
$$

then

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

where

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

Expansion of equilibrium conditions characterizes coefficients $\boldsymbol{n}_x, \boldsymbol{n}_\delta, \boldsymbol{N}_\delta, \boldsymbol{N}_\theta, \boldsymbol{v}_\delta, \boldsymbol{v}_\theta, \boldsymbol{v}_n$:

$$\boldsymbol{n}_\delta = \boldsymbol{N}_\delta = 0,\ \boldsymbol{n}_x = \boldsymbol{N}_\theta = \frac{r\lambda}{1-s\lambda},\ \boldsymbol{v}_\delta = 0,\ \boldsymbol{v}_\theta = \frac{1+\nu}{\eta+\nu}(1+\eta\boldsymbol{N}_\theta),\ \boldsymbol{v}_n = 0.$$

where

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

Because

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

the 0^{th} order expansion of attention $\bar{z}$ is still not determined.

2^{nd} order expansion of attention optimality

Expand the attention equilibrium condition to the second order

$$\lim_{\delta\to 0}\frac{d^2}{d\delta^2}\left(\mathbb{E}\left[-\frac{\epsilon_i^2-1}{2z(\delta)}\frac{1}{1-\gamma}e^{(1-\gamma)\log v_i(\delta)}\middle|z(\delta)\right] - \delta^2\kappa'(z)\right) = 0$$

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

where

$$\hat{\hat{v}}_i = \boldsymbol{v}_n \hat{\hat{n}}_i + \boldsymbol{v}_{nn} \hat{n}_i^2 + \cdots ,$$

$$\boldsymbol{v}_n = 0,\ \ \boldsymbol{v}_{nn} = -(1-\eta)(1+\nu),$$

and the omitted part is orthogonal to $\epsilon_i^2 - 1$. This is because, by normality and independence of $\epsilon_i, \hat{\theta}$, we have the following lemma:

Lemma 7

If $k, h \in \{2m+1 | m \in \mathbb{N}\}$, *or* $k = 0$,

$$\mathbb{E}\left[-\frac{\epsilon_i^2 - 1}{2\bar{z}} \ \epsilon_i^k \ \hat{\theta}^h \ \Big| \ \bar{z} \right] = 0.$$

Substitute back, this pins down the 0^{th} order expansion of attention $\bar{z}$:

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

If $\kappa(z) = \kappa \ z$, then

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

2^{nd} order of labor optimality and aggregation

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

$$
\begin{aligned}
0 =& \mathbb{E}[\ -(\eta+\nu)\ \hat{\hat{n}}_i + \eta\ \hat{\hat{N}}\ |\bar{x}_i] \\
&+2\frac{d}{d\delta}\mathbb{E}[-(\eta+\nu)\hat{n}_i + \hat{\theta} + \eta\hat{N}|x_i(\delta)]\Big|_{\delta=0} - \tilde{\gamma}\ Var(\ \theta + \eta\ \hat{N}\ |\bar{x}_i), \\
\hat{\hat{N}} =& \int \hat{\hat{n}}_i + (1-\eta)\ Var(\ \hat{n}_i\ |\theta)
\end{aligned}
$$

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

$$
\begin{aligned}
&\frac{d}{d\delta}\mathbb{E}[-(\eta+\nu)\hat{n}_i + \hat{\theta} + \eta\hat{N}|x_i(\delta)]\Big|_{\delta=0} \\
=&(1+\eta\boldsymbol{n}_x)\frac{d}{d\delta}\Big((\lambda(z(\delta)) - \lambda)\hat{x}_i(\delta)\Big)\Big|_{\delta=0} \\
=&(1+\eta\boldsymbol{n}_x)\lambda'(\bar{z})\hat{z}\ \hat{x}_i
\end{aligned}
$$

3^{rd} order expansion of attention optimality

$$
\lim_{\delta\to 0}\ \frac{d^3}{d\delta^3}\left(\mathbb{E}\left[-\frac{\epsilon_i^2 - 1}{2z(\delta)}\frac{1}{1-\gamma}e^{(1-\gamma)\log v_i(\delta)}\Big|z(\delta)\right] - \delta^2\kappa'(z)\right) = 0
$$

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

By lemma 7,

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

and the expanded attention optimality condition reduces to

$$-6\kappa'(\bar{z})\ \hat{z} + \mathbb{E}\left[-\frac{\epsilon_i^2-1}{2\bar{z}}\ v_i^{(3)}e^{(1-\gamma)\bar{v}} \middle|\ \bar{z}\ \right] - 2\kappa''(\bar{z})\bar{z}\ \hat{z} = 0$$

where

$$v_i^{(3)} = \boldsymbol{v}_n n_i^{(3)} + 3\boldsymbol{v}_{nn}\hat{\hat{n}}_i\hat{n}_i + \boldsymbol{v}_{nnn}\hat{n}_i^3 + 3\boldsymbol{v}_{nn\theta}\hat{\theta}\hat{n}_i^2 + 3\boldsymbol{v}_{nnN}\hat{N}\hat{n}_i^2 + \cdots .$$

Substitute $v_i^{(3)}$ back and cancel out terms that integrate to zero, we have:

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

Note the second order expansion of labor is given by:

$$\hat{\hat{n}}_{i,t} = \boldsymbol{n}_{xx}\hat{x}_i^2 + 2\boldsymbol{n}_{x\delta}\ \hat{x}_{i,t} + \boldsymbol{n}_{\delta\delta} + \boldsymbol{n}_x\frac{-\epsilon_i}{2\sqrt{\bar{z}}}\hat{z}.$$

From expansion of labor optimality:

$$0 = \mathbb{E}[\ -\hat{\hat{n}}_i + s\ \hat{\hat{N}}\ |\bar{x}_i] + 2\ (1+\eta\boldsymbol{n}_x)\lambda'(\bar{z})\hat{z}\ \hat{x}_i + const.$$

Because

$$\mathbb{E}[\ \hat{\hat{N}}\ |\bar{x}_i] = \mathbb{E}[\ \boldsymbol{n}_{xx}\hat{\theta}^2\ |\bar{x}_i] + const = \boldsymbol{n}_{xx}((1-\lambda)\sigma^2 + (\lambda\hat{x}_i)^2) + const,$$

substituting $\hat{\hat{n}}_i$ with its expansion gives

$$\begin{aligned}&\boldsymbol{n}_{xx}\ \hat{x}_i^2 = s\boldsymbol{n}_{xx}\lambda^2\ \hat{x}_i^2\\&(1-s\lambda)\boldsymbol{n}_{x\delta}\ \hat{x}_{i,t} = \left(\frac{\boldsymbol{n}_x(1-\lambda)}{\sqrt{\bar{z}}} + 2\ (1+\eta\boldsymbol{n}_x)\lambda'(\bar{z})\right)\hat{z}\ \hat{x}_i.\end{aligned}$$

Therefore, $\boldsymbol{n}_{xx} = 0$ and the LHS of the expanded optimality condition of attention is homogeneous in $\hat{z}$. Consequently,

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

and

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

Note $\hat{z} = 0$ because the equilibrium is approximated around a deterministic $\bar{\theta}$. How attention $\bar{z}(\bar{\theta})$ responds to business cycle condition is then studied as a comparative statics result of changing $\bar{\theta}$. In the dynamic setup, business cycle condition moves with previous shocks, which are scaled by the same perturbation parameter δ. In that case, the first order expansion of attention $\hat{z}$ (derived from the third order expansion of attention optimality) picks up the effect of previous shocks on attention, just like

$\bar{z}(\bar{\theta})$ picks up the effect of a change in $\bar{\theta}$ in the comparative statics of the static model.

Proof of Lemma 1

The four measures of uncertianty in economy δ is given by:

1. conditional aggregate volatility: $SD(\ \widetilde{Y}(\delta)\ |\bar{\theta}) := \left(\mathbb{E}\Big[\ \big(\widetilde{Y}(\delta) - \mathbb{E}[\ \widetilde{Y}(\delta)\ |\bar{\theta}]\big)^2\ \Big|\bar{\theta}\Big]\right)^{\frac{1}{2}}$
2. dispersion of outcomes: $Disp(\ \tilde{y}_i(\delta)\ |\bar{\theta}) := \left(\int \big(\tilde{y}_i(\delta) - \int \tilde{y}_i(\delta)\big)^2 di\right)^{\frac{1}{2}}$
3. forecast dispersion: $Disp(\ \mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\ |\bar{\theta}) := \left(\int \big(\mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)] - \int \mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\ di\big)^2\right)^{\frac{1}{2}}$
4. forecast uncertainty: $SD(\ \widetilde{Y}(\delta)\ |x_i(\delta), \bar{\theta}) := \left(\mathbb{E}_{i,\delta}\Big[\ \big(\widetilde{Y}(\delta) - \mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\big)^2\Big]\right)^{\frac{1}{2}}$, $\forall\ i$

where

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

Let $U(\delta)$ denote an uncertainty measure above:

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

where $f(\delta) \in \{\widetilde{Y}(\delta), \tilde{y}_i(\delta)\}$ such that $f(\delta) = \bar{f} + \hat{f}\delta + \frac{1}{2}\hat{\hat{f}}\delta^2 + \cdots$ and $\varphi(\theta,\delta)$ denote the probability or cross-section over which the uncertainty measure is calculated, $\int \varphi(\theta,\delta) = 1$.

0^{th} order expansion of uncertainty measures

$$\begin{aligned}
U^{(0)} &= \lim_{\delta\to 0}\left(\int\left(f(\delta)-\int f(\delta)\varphi(\theta,\delta)\right)^2\varphi(\theta,\delta)\right)^{\frac{1}{2}}\\
&=\left(\int\left(\bar{f}-\int \bar{f}\ \varphi(\theta,0)\right)^2\varphi(\theta,0)\right)^{\frac{1}{2}}=0
\end{aligned}$$

1^{st} order expansion of uncertainty measures

$$\begin{aligned}
U^{(1)} &= \lim_{\delta\to 0}\frac{d}{d\delta}\left(\int\left(f(\delta)-\int f(\delta)\varphi(\theta,\delta)\right)^2\varphi(\theta,\delta)\right)^{\frac{1}{2}}\\
&=\lim_{\delta\to 0}\frac{1}{2}\left(\int\left(f(\delta)-\int f(\delta)\varphi(\theta,\delta)\right)^2\varphi(\theta,\delta)\right)^{\frac{-1}{2}}\\
&\quad\times\int 2\left(f(\delta)-\int f(\delta)\varphi(\theta,\delta)\right)\left(\hat{f}-\int\hat{f}\ \varphi(\theta,\delta)-\frac{d}{d\delta}\int\bar{f}\ \varphi(\theta,\delta)\right)\varphi(\theta,\delta)\\
&=\left(\int\left(\hat{f}-\int\hat{f}\ \varphi(\theta,0)\right)^2\varphi(\theta,0)\right)^{\frac{1}{2}}
\end{aligned}$$

2^{nd} order expansion of uncertainty measures

$$\begin{aligned}
U^{(2)} &= \lim_{\delta\to 0}\frac{d^2}{d\delta^2}\left(\int\left(f(\delta)-\int f(\delta)\varphi(\theta,\delta)\right)^2\varphi(\theta,\delta)\right)^{\frac{1}{2}}\\
&=\frac{1}{U^{(1)}}\int\left(\hat{f}-\int\hat{f}\ \varphi(\theta,0)\right)\left(\hat{\hat{f}}-\int\hat{\hat{f}}\ \varphi(\theta,0)-\frac{d}{d\delta}\int\hat{f}\ \varphi(\theta,\delta)\Big|_{\delta=0}\right)\varphi(\theta,0)
\end{aligned}$$

Because the expansion is around a deterministic $\bar{\theta}$, we have $\hat{z} = 0$ and

1. $\frac{d^2}{d\delta^2} SD(\ \widetilde{Y}(\delta)\ |\bar{\theta})\Big|_{\delta=0} = 0, \quad \text{following} \quad \hat{\hat{Y}} \in \mathbb{R}, \quad \frac{d}{d\delta}\ \mathbb{E}[\hat{Y}|\bar{\theta}] = 0,$
2. $\frac{d^2}{d\delta^2} Disp(\ \tilde{y}_i(\delta)\ |\bar{\theta})\Big|_{\delta=0} = 0, \quad \text{following} \quad \hat{\hat{y}}_i \in \mathbb{R}, \quad \frac{d}{d\delta} \int \hat{y}_i di = 0,$
3. $\frac{d^2}{d\delta^2} Disp(\ \mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\ |\bar{\theta})\Big|_{\delta=0} = 0, \quad \text{following}$
   $\frac{d^2}{d\delta^2}\mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\Big|_{\delta=0} = \hat{\hat{Y}}, \quad \frac{d}{d\delta} \int \left(\frac{d}{d\delta}\mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\Big|_{\delta=0} \right) di = 0, \text{ and}$
4. $\frac{d^2}{d\delta^2} SD(\ \widetilde{Y}(\delta)\ |x_i(\delta), \bar{\theta})\Big|_{\delta=0} = 0, \quad \text{following} \quad \hat{\hat{Y}} \in \mathbb{R}, \quad \frac{d}{d\delta}\mathbb{E}_{i,\delta}[\hat{Y}]\Big|_{\delta=0} = 0,$

where $\frac{d^2}{d\delta^2}\mathbb{E}_{i,\delta}[\widetilde{Y}(\delta)]\Big|_{\delta=0} = 0$ and $\frac{d}{d\delta}\mathbb{E}_{i,\delta}[\hat{Y}]\Big|_{\delta=0} = 0$ follows $\hat{z} = 0$.

Approximation of uncertainty measures up to second order

From the derivation above, up to second order approximation, the four uncertainty measures are given by:

Lemma 8

$$
\begin{aligned}
&SD(\widetilde{Y}|\bar{\theta}) \approx \big(1+\Omega(\bar{z})\big)\ \sigma, \\
&Disp(\tilde{y}_i|\bar{\theta}) \approx \Omega(\bar{z})\frac{1}{\sqrt{\bar{z}}}, \\
&Disp(\mathbb{E}_i[\widetilde{Y}]|\bar{\theta}) \approx \big(1+\Omega(\bar{z})\big)\lambda(\bar{z})\frac{1}{\sqrt{\bar{z}}}, \\
&SD(\widetilde{Y}|x_i,\bar{\theta}) \approx \big(1+\Omega(\bar{z})\big)\sqrt{1-\lambda(\bar{z})}\ \sigma, \forall i.
\end{aligned}
$$

where

$$
\Omega(\bar{z}) = \boldsymbol{N}_\theta = \frac{r\lambda(\bar{z})}{1-s\lambda(\bar{z})}, \quad \lambda(\bar{z}) = \frac{\sigma^2}{\sigma^2+1/\bar{z}}.
$$

Proof of Lemma 2

Recall from the derivation above,

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

Since $\bar{v}$ is increasing in $\bar{\theta}$, $\bar{z}$ is decreasing in $\bar{\theta}$ if and only if $\gamma > 1$. Note that

$$
-e^{(1-\gamma)\bar{v}}\ \boldsymbol{v}_{nn} = \underbrace{(\nu+\eta)(1-\eta)e^{\bar{\theta}+\bar{n}}}_{(1)}\ \underbrace{e^{-\gamma\bar{v}}}_{(2)},
$$

where (1) captures the loss from deviating away from the optimal labor, expressed in the unit of final good and approximated at the 0^{th} order expansion:

$$\frac{\partial^2}{\partial \bar{n}^2}\left(e^{\bar{\theta}+\eta\bar{N}+(1-\eta)\bar{n}} - \frac{\psi}{1+\nu}e^{(1+\nu)\bar{n}}\right) = -(\eta+\nu)(1-\eta)e^{\bar{\theta}+\bar{n}},$$

and (2) translate the loss from final good to util.

Proof of Proposition 1

Because $\gamma > 1$, Lemma 2 implies

$$\frac{\partial}{\partial \bar{\theta}}\bar{z} < 0.$$

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

Lemma 9

The four uncertainty measures as functions of $\bar{z}$ has the following characterization:

1. $SD(\widetilde{Y}|\bar{\theta}) \approx \big(1+\Omega(\bar{z})\big)\,\sigma$ *is increasing in* $\bar{z}$,
2. $Disp(\tilde{y}_i|\bar{\theta}) \approx \Omega(\bar{z})\frac{1}{\sqrt{\bar{z}}}$,
3. $Disp(\mathbb{E}_i[\widetilde{Y}]|\bar{\theta}) \approx \big(1+\Omega(\bar{z})\big)\lambda(\bar{z})\frac{1}{\sqrt{\bar{z}}}$

are hump-shaped functions of $\bar{z}$, which equals 0 at $\bar{z}=0$, goes to 0 as $\bar{z}\to\infty$, and

> 0 *in between, and*

$$4.\ SD(\widetilde{Y}|x_i, \bar{\theta}) \approx \big(1 + \Omega(\bar{z})\big)\sqrt{1 - \lambda(\bar{z})}\ \sigma$$

is either decreasing or a hump-shaped function of $\bar{z}$*, which equals* σ *at* $\bar{z} = 0$*, and goes to* 0 *as* $\bar{z} \to \infty$*. It is a hump-shaped function if and only if* $r = \frac{1}{\eta+\nu} > \frac{1}{2}$*.*

1. The fact that conditional volatility is increasing in $\bar{z}$ follows directly from $\Omega(\bar{z})$ is increasing in $\bar{z}$.

2. Write

$$d_y(\lambda) := \Omega(z(\lambda))\frac{1}{\sqrt{z(\lambda)}}$$

where $z(\lambda)$ is the inverse of $\lambda(z) = \frac{\sigma^2}{\sigma^2+1/z}$. The following condition characterize how dispersion of individual output changes with λ:

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

where

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

Since $g_y(\lambda)$ is increasing in λ, $g_y(0) = -\infty$ and $g_y(1) = 1$, there exists $\lambda_{d_y} := \inf_{\lambda} g_y(\lambda) > s$, and $d_y(\lambda)$ has a single peak at λ_{d_y}. Moreover, as $s \to 1$, $\lambda_{d_y} \to 1$

and $z(\lambda_{d_y}) \to \infty$.

3. Write

$$f_d(\lambda) := \big(1 + \Omega(z(\lambda))\big)\lambda \frac{1}{\sqrt{z(\lambda)}}.$$

The following condition characterize how forecast dispersion about aggregate output changes with $\bar{z}$:

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

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

$$u_{f_d} := \inf_{\lambda} \{\lambda \mid g_d(\lambda) > s\}.$$

Moreover,

$$h_d(\frac{1}{2}) = 0,\ h_d(u_{f_d}) = \infty$$

and $\forall \lambda \in (\frac{1}{2}, u_{f_d})$,

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

Therefore, there exists

$$\lambda_{f_d} \coloneqq \inf_{\lambda < u_{f_d}} \{\lambda \,| h_d(\lambda) > r\},$$

and $f_d(\lambda)$ has a single peak at λ_{f_d}.

Moreover, since

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

we have $\lambda_{f_d} \to 1$ as $s \to 1$.

4. Write

$$f_u(\lambda) \coloneqq \big(1 + \Omega(z(\lambda))\big)\sqrt{1-\lambda}\,\sigma.$$

The following condition characterize how forecast uncertainty about aggregate output changes with λ:

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

where

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

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

$$u_{f_u} \coloneqq \inf_{\lambda} \{\lambda \,|g_u(\lambda) > s\}.$$

Moreover,

$$h_u(u_{f_u}) = \infty$$

and $\forall \lambda \in [0, u_{f_u})$,

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

Therefore, there exists

$$\lambda_{f_u} \coloneqq \inf_{\lambda < u_{f_u}} \{\lambda \,|h_u(\lambda) > r\},$$

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

Moreover, since

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

we have $\lambda_{f_u} \to 1$ as $s \to 1$.

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

A.2 Extensions

General attention cost

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

where, recall that

$$\boldsymbol{n}_x = \frac{r\lambda}{1 - s\lambda}, \quad \lambda = \frac{\sigma^2}{\sigma^2 + 1/\bar{z}}, \quad \boldsymbol{v}_{nn} = -(1-\eta)(1+\nu).$$

With $\bar{v} = \bar{v}(\bar{\theta}$ and $\bar{z} = \bar{z}(\bar{\theta}$, totally differentiate with respect to $\bar{\theta}$ and solve for $\bar{z}'(\bar{\theta}$, we have

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

where $\bar{v}'(\bar{\theta}) > 0$.

Suppose that $\kappa''(\cdot) > 0$, then

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

Attention choice amplifies exogenous changes in volatility σ

In this section, I consider how endogenous attention affects the economy's response to an exogenous increase in σ.

To study the effect of a change in σ, I consider a slightly different parameterization of attention

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

This is the appropriate specification because under this specification, the same attention cost will lead to the same amount of information, as measured by the ratio of variance reduction regardless of σ. (Note this is irrelevant for the analysis in the main text because when σ^2 is fixed, it is only a normalization of the cost function.)

With this specification,

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

and an increase in σ leads to an increase in attention:

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

The approximation of the four uncertainty measures are respectively:

$$
\begin{aligned}
&SD(\widetilde{Y}|\bar{\theta}) \approx \big(1+\Omega(\bar{z})\big)\sigma \\
&Disp(\tilde{y}_i) \approx \Omega(\bar{z})\frac{\sigma}{\sqrt{\bar{\bar{z}}}} \\
&Disp(\mathbb{E}[\widetilde{Y}|\bar{\theta}, x_i]) \approx \big(1+\Omega(\bar{z})\big)\lambda\frac{\sigma}{\sqrt{\bar{\bar{z}}}} \\
&SD(\widetilde{Y}|\bar{\theta}, x_i) \approx \big(1+\Omega(\bar{z})\big)\sqrt{(1-\lambda(z))}\sigma.
\end{aligned}
$$

Under the same conditions as proposition 1,

$$
\begin{aligned}
&\frac{d}{d\log\sigma}\log\big(1+\Omega(\bar{z})\big)\sigma > 1 \\
&\frac{d}{d\log\sigma}\log\frac{\sigma}{\sqrt{\bar{\bar{z}}}} > 1 \\
&\frac{d}{d\log\sigma}\log\big(1+\Omega(\bar{z})\big)\lambda\frac{\sigma}{\sqrt{\bar{\bar{z}}}} > 1 \\
&\frac{d}{d\log\sigma}\log\big(1+\Omega(\bar{z})\big)\sqrt{(1-\lambda(z))}\sigma > 1
\end{aligned}
$$

Note that without endogenous attention choice, these elasticities all equals to one. In this sense, the mechanism in proposition 1 not only provides a new source of uncertainty fluctuations, but also a mechanism that amplifies exogenous changes in volatility.

Separating income effect on attention from risk aversion

Consider the following preference

$$u_i = \frac{1}{1-\psi}\mathbb{E}[v_i^{1-\gamma}]^{\frac{1-\psi}{1-\gamma}} - \kappa(z)$$

where, as before,

$$v_i = \max\{c_i - \nu(n_i), \underline{c}\}.$$

Similar approximation as in Appendix A.1 gives:

$$\lim_{\delta\to 0} \frac{d^2}{d\delta^2}\left(\frac{1}{1-\psi}\mathbb{E}\left[-\frac{\epsilon_i^2-1}{2z(\delta)}e^{(1-\gamma)\log v_i(\delta)}\Big| z(\delta)\right]^{\frac{1-\psi}{1-\gamma}} - \delta^2\kappa'(z)\right) = 0$$
$$\iff \mathbb{E}\Bigg[-\frac{\epsilon_i^2-1}{2\bar{z}}\left((1-\gamma)\hat{v}_i^2 + \hat{\hat{v}}_i\right)e^{(1-\psi)\bar{v}}\Big|\ \bar{z}\ \Bigg] - 2\kappa'(\bar{z}) = 0$$

where

$$\hat{\hat{v}}_i = \boldsymbol{v}_n\hat{\hat{n}}_i + \boldsymbol{v}_{nn}\hat{n}_i^2 + \cdots, \quad \boldsymbol{v}_n = 0,\ \boldsymbol{v}_{nn} = -(1-\eta)(1+\nu),$$

and the omitted part is orthogonal to $\epsilon_i^2 - 1$ by lemma 7.

Substitute back, this pins down the 0^{th} order expansion of attention $\bar{z}$:

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

If $\kappa(z) = \kappa\ z$, then

$$\bar{z} = \frac{1}{1-s}\left(\sqrt{\frac{r^2 e^{(1-\psi)\bar{v}}|\boldsymbol{v}_{nn}|}{2\kappa}} - \frac{1}{\sigma^2}\right).$$

A.3 Dynamic model

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

$$h_i^t \coloneqq \{x_i^{t-1}, p_i^{*t-1}\},\ \forall t \geq 1$$

where $p_{i,t}^* = e^{\theta_t+\omega_{i,t}}N_t^\eta$ is a transformation of $p_{i,t}$ that does not alter the information content.

Let $h_i^{t'}, h_i^{t''}$ denote the period t history including idiosyncratic signal $x_{i,t}$ and information from price of good i:

$$h_i^{t'} \coloneqq \{h_i^t, x_{i,t}\}, \forall t \geq 0,$$
$$h_i^{t''} \coloneqq \{h_i^t, x_{i,t}, p_{i,t}^*\},\ \forall t \geq 0.$$

and $H_i^{t'}, H_i^{t''}$ are their collection.

Rewrite period payoff as

$$U(\theta, N, n, \omega) \coloneqq u(c(\theta, N, n, \omega), n)$$

and write the variance of noise and attention cost in a more general form

$$\sigma^2(z) \coloneqq \frac{1}{z}, \quad \kappa(z) \coloneqq \kappa\ z.$$

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

$$z_{i,t} : H_i^t \to \mathbb{R}_+, \quad n_{i,t} : H_i^{t'} \to \mathbb{R}_+.$$

Given $\{N_t\}$, agent i's discounted payoff after history h_i^t and $h_i^{t'}$is

$$v_i(n_i, z_i | h_i^t) \coloneqq \mathbb{E}\Big[\sum_{s=t}^{\infty} \beta^{s-t} \Big(U(\theta_s, N_s, n_{i,s}(h_i^{s'}), \omega_{i,s}) - \kappa(z_{i,s}(h_i^s))\Big)\Big| h_i^t, z_i\Big]$$

A strategy $\{n_t, z_t\}_{t=0}^{\infty}$ is a pure strategy perfect symmetric equilibrium only if

$$\int U(\theta_t, N_t, n_t(h_i^{t'}), \omega_{i,t}) - \kappa(z_t(h_i^t)) + \beta v_i(n, z | h_i^{t+1})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1} | h_i^t, z^t)$$
$$\geq \int U(\theta_t, N_t, n_t(h_i^{t'}), \omega_{i,t}) - \kappa(\tilde{z}) + \beta v_i(n, z | h_i^{t+1})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1} | h_i^t, \{z^{t-1}, \tilde{z}\})$$

and

$$\int U(\theta_t, N_t, n_t(h_i^{t'}), \omega_{i,t}) - \kappa(z_t(h_i^t)) + \beta v_i(n, z|h_i^{t+1})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1}|h_i^{t'}, z^t)$$
$$\geq \int U(\theta_t, N_t, \tilde{n}, \omega_{i,t}) - \kappa(z_t(h_i^t)) + \beta v_i(n, z|h_i^{t+1})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1}|h_i^{t'}, z^t)$$

for all $\tilde{n}, \tilde{z} \in \mathbb{R}_+$ after any history h_i^t, where in an symmetric equilibrium

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

The following two first order condition follows

$$\frac{\partial}{\partial \tilde{z}} \Bigg(\int U(\theta_t, N_t, n_t(h_i^{t'}), \omega_{i,t}) - \kappa(\tilde{z}) + \beta v_i(n, z|h_i^{t+1})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1}|h_i^t, \{z^{t-1}, \tilde{z}\}) \Bigg)\Bigg|_{\tilde{z}=z_{i,t}} = 0,$$

$$\frac{\partial}{\partial \tilde{n}} \left(\int U(\theta_t, N_t, \tilde{n}, \omega_{i,t})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1}|h_i^{t'}, z^t) \right)\Bigg|_{\tilde{n}=n_{i,t}} = 0.$$

Suppose that differentiation can pass through integral, then the F.O.C. for $\tilde{n}$ becomes

$$\int \frac{\partial}{\partial \tilde{n}} U(\theta_t, N_t, n_{i,t}, \omega_{i,t})\ d\Phi(\theta^t, \omega_i^t, h_i^{t+1}|h_i^{t'}, z_i^t) = 0.$$

Let $\varphi(\theta^t, \omega_i^t, h_i^{t+1} | h_i^t, \{z^{t-1}, \tilde{z}\})$ denote the density of $\Phi(\theta^t, \omega_i^t, h_i^{t+1} | h_i^t, \{z^{t-1}, \tilde{z}\})$, then

$$\frac{\partial}{\partial \tilde{z}} \varphi(\theta^t, \omega_i^t, h_i^{t+1} | h_i^t, \{z^{t-1}, \tilde{z}\}) = \left(\frac{\partial}{\partial \tilde{z}} \varphi(x_{i,t} | \theta_t, \tilde{z}) \right) \varphi(p_{i,t}^*, x_{t+1} | \theta^t, \epsilon^t) \varphi(\theta^t, \omega_i^t | h_i^t, z^{t-1}).$$

Expanding the derivative and evaluate at $\tilde{z} = z_{i,t}$

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

where $\phi(\cdot)$ denotes the density of standard normal.

Hence,

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

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

$$\sum_{s=t}^{\infty} \beta^{s-t} \int \frac{\sigma'(z_{i,t})}{\sigma(z_{i,t})} (\epsilon_{i,t}^2 - 1) \Big(U(\theta_s, N_s, n_{i,s}, \omega_{i,s}) - \kappa(z_{i,s}) \Big) \, d\Phi(\theta^s, h_i^{s+1} | h_i^t, z_i^t) = \kappa'(z_{i,t}).$$

Solving the dynamic model

The computational procedure for approximating dynamic dispersed information economies takes two steps.

Step 1: Characterization of expansions

The first step consists of constructing expansions of equilibrium processes and deriving explicit systems that characterize these expansions from equilibrium conditions. This step resembles the procedure in section 1.2.

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

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

and look for Taylor expansions of the equilibrium with respect to δ around the limiting economy where $\delta \to 0$:

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

In the expressions above, $\bar{N}, \bar{n}, \bar{z}$ are the limits at $\delta = 0$, and $\hat{N}_t, \hat{n}_{i,t}, \hat{z}_{i,t}, \ldots$ are all now stochastic processes.

To characterize these expansion sequences, I expand the equilibrium conditions as in the static model, evaluate at $\delta = 0$, and obtain the following:

Lemma 10

$$\mathbb{E}\Big[\nabla \bar{U}_n \begin{pmatrix} \theta_t \\ \hat{N}_t \\ \hat{n}_{i,t} \\ \omega_{i,t} \end{pmatrix} \Big| \bar{\mathcal{F}}'_{i,t}\Big] = 0,$$

$$\hat{N}_t = \int \hat{n}_{j,t},$$

$$\sum_{\tau=t}^{\infty} \beta^{\tau-t}\mathbb{E}\Big[-\frac{\epsilon_{i,t}^2 - 1}{2\bar{z}}\ \bar{U}_{nn}\hat{n}_{i,\tau}^2 \Big| \bar{\mathcal{F}}_{i,t}\Big] = 2\kappa$$

where

$$U_n(\theta, N, n, \omega) \coloneqq \frac{\partial}{\partial \log n} u(p(\theta, N, n, \omega)n, n).$$

$\nabla \bar{U}_n$ *denotes its gradient with respect to* $\theta, \log N, \log n,$ *and* ω*, and* $\bar{U}_{nn}$ *denotes its derivative with respect to* $\log n$*, both evaluated at* $\delta = 0$*.*

This characterization corresponds to the expanding the optimality condition for labor input, the aggregation condition, and the optimality condition for attention in Lemma 3.

The system in lemma 10 jointly characterize the 0^{th} order exapansion of attention, $\bar{z}$, and the 1^{st} order expansion for labor, $\hat{n}_{i,t}$ and $\hat{N}_t$. Note this characterization gives the identical result as the quadratic approximation of payoff commonly used in the dispersed information and rational inattention literature.[1] The difference is that, whereas quadratic approximation of payoff cannot be generalized, the perturbation method generalizes to an arbitrary higher order and can be easily applied to all existing dispersed information models to understand their higher-order implications, including the concern for uncertainty, asymmetric response, state dependency, and the distributional effect of belief heterogeneity.

For this paper, I take one step forward and solve for the first order dynamics of attention, $\hat{z}_{i,t}$, and the second order dynamics of labor input, $\hat{\hat{n}}_{i,t}$ and $\hat{N}_t$. The following lemma characterizes the second order dynamics the equilibrium:

Lemma 11

1. See Angeletos and Lian (2016) and Mackowiak et al. (2018) for recent surveys.

$$\mathbb{E}\Big[\bar{U}_{nn}\hat{\hat{n}}_{i,t}+\bar{U}_{nN}\hat{\hat{N}}_t\Big|\bar{\mathcal{F}}'_{i,t}\Big]+\mathbb{E}\Big[(\theta_t\ \hat{N}_t\ \hat{n}_{i,t}\ \omega_{i,t})\ \bar{H}(U_n)\begin{pmatrix}\theta_t\\ \hat{N}_t\\ \hat{n}_{i,t}\\ \omega_{i,t}\end{pmatrix}\Big|\bar{\mathcal{F}}'_{i,t}\Big]$$

$$+2\frac{d}{d\delta}\mathbb{E}\Big[\nabla\bar{U}_n\begin{pmatrix}\theta_t\\ \hat{N}_t\\ \hat{n}_{i,t}\\ \omega_{i,t}\end{pmatrix}\Big|\mathcal{F}'_{i,t}(\delta)\Big]\Big|_{\delta=0}=0$$

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

$$\sum_{\tau=t}^{\infty}\beta^{\tau-t}\mathbb{E}\Big[-\frac{\epsilon_{i,t}^2-1}{2\bar{z}}\ \bar{U}^{(3)}_{i,\tau}\ \Big|\bar{\mathcal{F}}_{i,t}\Big]=6\ \kappa\hat{z}_{i,t}$$

where

$$U_n(\theta,N,n,\omega):=\frac{\partial}{\partial\log n}u(c(\theta,N,n,\omega),n),$$

and $\bar{H}(U_n)$ *denotes the Hessian of* U_n *with respect to* $\theta,\log N,\log n,\omega,$ *evaluated at* $\delta=0$*, and* $\bar{U}^{(3)}_{i,\tau}$ *is the third order expansion of* $U(\theta_\tau(\delta),N_\tau(\delta),n_{i,\tau}(\delta),\omega_{i,\tau}(\delta))$ *evaluated at* $\delta=0$*.*

Step 2: Solving expansions sequences

Transforming the problem from solving the stochastic system in lemma 3 into solving the system in lemma 10 (or ones that correspond to higher order expansions) reduces the problem significantly. This is because expansion sequences have special structures we can exploit. Recall that in the static model, once we construct the expansions, we can solve them easily from the expanded the equilibrium conditions because the problem reduces to solving for coefficients $\bar{\boldsymbol{n}}_x, \bar{\boldsymbol{n}}_\delta, \bar{\boldsymbol{n}}_{xx}, \ldots$, where

$$\log n_i(\delta) = \boldsymbol{n}(x_i(\delta), \delta) = \bar{n} + (\bar{\boldsymbol{n}}_x\ \hat{x}_i + \bar{\boldsymbol{n}}_\delta)\delta + \frac{1}{2}(\hat{x}_i^\intercal\ \bar{\boldsymbol{n}}_{xx}\ \hat{x}_i + \cdots)\ \delta^2 + \cdots .$$

Similar is the case for the dynamic problem. Let

$$X_i^t(\delta) \coloneqq (x_i^t(\delta),\ p_i^{t-1}(\delta))$$

denote agent i's information in economy δ when they make labor input choice; let $\boldsymbol{n}\left(X_i^t(\delta), \delta\right)$ denote the policy function that maps signals into log equilibrium labor input.

Differentiate with respect to δ and evaluate at $\delta = 0$, the Taylor expansion of labor input is then is given by:

$$\log n_{i,t}(\delta) = \boldsymbol{n}\left(X_i^t(\delta), \delta\right) = \bar{n} + \left(\bar{\boldsymbol{n}}_X \hat{X}_i^t + \bar{\boldsymbol{n}}_\delta\right)\ \delta + \frac{1}{2}\left(\hat{X}_i^{t\intercal} \bar{\boldsymbol{n}}_{XX} \hat{X}_i^t + \cdots\right)\ \delta^2 + \cdots$$

where $\hat{X}_i^t$ denotes the first derivative of $X_{i,t}(\delta)$ with respect to δ evaluated at $\delta = 0$. As in the static model, approximating the economy around $\delta = 0$ reduces the the problem reduce from solving a non-linear policy function of signal histories into solving multi-linear maps $\bar{\boldsymbol{n}}_X, \bar{\boldsymbol{n}}_{XX}, \ldots$.

Yet, different from the static setup, infinite regress problem in the dynamic setting implies each agent's decision depends on the whole history of their signals. Although the problem reduce from solving non-linear policy function into solving multi-linear maps, these multi-linear maps are all infinite-dimension objects. On top of that, the equilibrium depends on a whole cross-section of agents with different histories of signals. This complexity reflects the difficulty that in dispersed information models, the "state variables" of the economy, in principal, can be the whole distribution over infinite-dimensional histories of signals.

In the companion paper , I exploit two features of the problem to address this difficulty. First, Gaussian shocks with a continuum of agents in the model allow me to aggregate the expansions sequence of individual variables to their aggregate counterparts. This aggregation deals with the cross-sectional heterogeneity that stems from dispersed information. Second, the coefficients $\bar{\boldsymbol{n}}_X, \bar{\boldsymbol{n}}_{XX}, \ldots$ are all symmetric multi-linear maps. This suggests the best finite-dimensional approximate state of the expansion sequence is a polynomial of linear processes driven by the fundamental shocks and noises. This motivates an effective way to summarize the infinite-dimensional histories of signal with finite-dimensional auto-regressive processes.

Exploiting these two features, I provide a simple algorithm to approximate higher-order dynamics of dispersed information economies with infinite regress problem in the companion paper . The algorithm applies to general dispersed information models and its application to this paper allows me to capture the equilibrium dynamics of attention under dispersed information and study the fluctuations of uncertainty over business cycles.

Proof of Lemma 10 and 11

Aggregation condition

1^{st} *order*

$$\hat{N}_t = \frac{1}{1-\eta}\frac{1}{\int e^{(1-\eta)\bar{n}}}\int e^{(1-\eta)\bar{n}}(1-\eta)\hat{n}_{i,t} \iff \hat{N}_t = \int \hat{n}_{i,t}.$$

2^{nd} *order*

$$\begin{aligned}\hat{\hat{N}}_t &= \frac{1}{1-\eta}\Big(\frac{-1}{\big(\int e^{(1-\eta)\bar{n}}\big)^2}\left(\int e^{(1-\eta)\bar{n}}(1-\eta)\hat{n}_{i,t}\right)^2 \\ &+ \frac{1}{\int e^{(1-\eta)\bar{n}}}\int e^{(1-\eta)\bar{n}}\left((1-\eta)^2\hat{n}_{i,t}^2 + (1-\eta)\hat{\hat{n}}_{i,t}\right)\Big) \\ \iff \hat{\hat{N}}_t &= (1-\eta)\int\left(\hat{n}_{i,t} - \int \hat{n}_{i,t}\right)^2 + \int \hat{\hat{n}}_{i,t}\end{aligned}$$

Labor optimality

1^{st} *order*

$$\begin{aligned}&\frac{d}{d\delta}\mathbb{E}[U_n(\theta_t(\delta), N_t(\delta), n_{i,t}(\delta), \omega_{i,t}(\delta))\ |\mathcal{F}'_{i,t}(\delta)]\Big|_{\delta=0} = 0 \\ \iff &\mathbb{E}[\bar{U}_{n\theta}\ \theta_t + \bar{U}_{nN}\hat{N}_t + \bar{U}_{nn}\hat{n}_{i,t} + \bar{U}_{n\omega}\omega_{i,t}|\bar{\mathcal{F}}'_{i,t}] + \frac{d}{d\delta}\mathbb{E}[\bar{U}_n|\mathcal{F}'_{i,t}(\delta)]\Big|_{\delta=0} = 0,\end{aligned}$$

where the second term equals zero because $\bar{U}_n \equiv 0$.

2^{nd} *order*

$$
\begin{aligned}
&\frac{d^2}{d\delta^2}\mathbb{E}[U_n(\theta_t(\delta), N_t(\delta), n_{i,t}(\delta), \omega_{i,t}(\delta))\ |\mathcal{F}'_{i,t}(\delta)]\Big|_{\delta=0} = 0 \\
\iff &\mathbb{E}\left[\bar{U}_{nn}\ \hat{\hat{n}}_{i,t} + \bar{U}_{nN}\hat{\hat{N}}_t + U^{(1)}_{n\theta}\theta_t + U^{(1)}_{nN}\hat{N}_t + U^{(1)}_{nn}\hat{n}_{i,t} + U^{(1)}_{n\omega}\omega_{i,t}\ \Big|\bar{\mathcal{F}}'_{i,t}\right] \\
&+ 2\frac{d}{d\delta}\mathbb{E}[U^{(1)}_n|\mathcal{F}'_{i,t}(\delta)]\Big|_{\delta=0} \\
&+ \frac{d^2}{d\delta^2}\mathbb{E}[\bar{U}_n|\mathcal{F}'_{i,t}(\delta)]\Big|_{\delta=0} \\
&= 0
\end{aligned}
$$

Expanding and rearranging gives

$$
\mathbb{E}\Big[\bar{U}_{nn}\hat{\hat{n}}_{i,t} + \bar{U}_{nN}\hat{\hat{N}}_t\Big|\bar{\mathcal{F}}'_{i,t}\Big] + \mathbb{E}\Big[(\theta_t\ \hat{N}_t\ \hat{n}_{i,t}\ \omega_{i,t})\ \bar{H}(U_n)\begin{pmatrix}\theta_t\\ \hat{N}_t\\ \hat{n}_{i,t}\\ \omega_{i,t}\end{pmatrix}\Big|\bar{\mathcal{F}}'_{i,t}\Big]
$$

$$
+2\frac{d}{d\delta}\mathbb{E}\Big[\nabla\bar{U}_n\begin{pmatrix}\theta_t\\ \hat{N}_t\\ \hat{n}_{i,t}\\ \omega_{i,t}\end{pmatrix}\Big|\mathcal{F}'_{i,t}(\delta)\Big]\Big|_{\delta=0} = 0
$$

Attention optimality

Lemma 12

$$\frac{d}{d\delta}\mathbb{E}\left[(\epsilon_{i,t}^2 - 1)U_{i,s}^{(m)}\middle|\ \mathcal{F}_{i,t}(\delta)\right] = 0, \quad \forall m = 0, 1, 2, \forall s.$$

Since $\epsilon_{i,t} \perp \mathcal{F}_{i,t}(\delta), \theta_s, \omega_{i,s}, \forall s$ and $\epsilon_{i,t} \perp \epsilon_{i,s}, \forall s \neq t$ and $\bar{U}_n = 0$, the only non-zero terms for $\mathbb{E}\left[(\epsilon_{i,t}^2 - 1)U_{i,s}^{(m)}\middle|\ \mathcal{F}_{i,t}(\delta)\right] = 0$ is when $m = 2$ and, in that case the expectation is a constant that does not change with δ:

$$\mathbb{E}\left[(\epsilon_{i,t}^2 - 1)U_{i,s}^{(2)}\middle|\ \mathcal{F}_{i,t}(\delta)\right] = \mathbb{E}\left[(\epsilon_{i,t}^2 - 1)\boldsymbol{v}_{nn}\hat{n}_{i,s}^2\middle|\ \mathcal{F}_{i,t}(\delta)\right] = \boldsymbol{v}_{nn}\frac{\partial \boldsymbol{n}}{\partial x_{i,t}}\mathbb{E}\left[(\epsilon_{i,t}^2 - 1)\epsilon_{i,t}^2\right].$$

With the lemma above, the 2^{nd} and 3^{rd} order expansions of attention optimality can be derived as follows.

2^{nd} *order*

$$\frac{d^2}{d\delta^2}\left(\frac{-1}{2}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}\Big(U_{i,s}(\delta)-\delta^2\kappa\ z_{i,s}(\delta)\Big)\Big|\ \mathcal{F}_{i,t}(\delta)\right]-\delta^2\kappa\ z_{i,t}(\delta)\right)\Bigg|_{\delta=0}\ =\ 0$$

$$\iff\ \frac{-1}{2}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}\left(U_{i,s}^{(2)}-2\ \kappa\ \bar{z}\right)\Big|\ \bar{\mathcal{F}}_{i,t}\right]$$

$$+\frac{-1}{2}\sum_{l=1}^{2}\frac{d^l}{d\delta^l}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}U_{i,s}^{(2-l)}\Big|\ \mathcal{F}_{i,t}(\delta)\right]\Bigg|_{\delta=0}-2\kappa\ \bar{z}\ =\ 0$$

$$\iff\ \frac{-1}{2}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}U_{i,s}^{(2)}\Big|\ \bar{\mathcal{F}}_{i,t}\right]-2\kappa\ \bar{z}\ =\ 0.$$

3^{rd} *order*

$$\frac{d^3}{d\delta^3}\left(\frac{-1}{2}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}\Big(U_{i,s}(\delta)-\delta^2\kappa\ z_{i,s}(\delta)\Big)\Big|\ \mathcal{F}_{i,t}(\delta)\right]-\delta^2\kappa\ z_{i,t}(\delta)\right)\Bigg|_{\delta=0}\ =\ 0$$

$$\iff\ \frac{-1}{2}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}\left(U_{i,s}^{(3)}-6\ \kappa\ \bar{z}\ \hat{z}_{i,s}\right)\Big|\ \bar{\mathcal{F}}_{i,t}\right]$$

$$+\frac{-1}{2}\sum_{l=1}^{3}\frac{d^l}{d\delta^l}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}U_{i,s}^{(3-l)}\Big|\ \mathcal{F}_{i,t}(\delta)\right]\Bigg|_{\delta=0}-6\kappa\ \bar{z}\ \hat{z}_{i,t}=\ 0$$

$$\iff\ \frac{-1}{2}\mathbb{E}\left[(\epsilon_{i,t}^2-1)\sum_{s=t}^{\infty}\beta^{s-t}U_{i,s}^{(3)}\Big|\ \bar{\mathcal{F}}_{i,t}\right]-6\kappa\ \bar{z}\ \hat{z}_{i,t}=\ 0.$$

APPENDIX B

SUPPLEMENTARY DATA

Countercyclical Attention

Attention is intrinsically unobservable. To understand how people's attention to economic events varies over the business cycle, I use internet traffic data to construct proxies for people's attention.

The main data set I used to construct a measure of people's attention on economic issues over time is the Google Trend data set (https://trends.google.com/).[1] Google Trend provides query share of any group of search terms relative to the total amount of queries on Google in any period of time since 2004. Query share of search terms is assigned into different categories using a natural-language-processing algorithm by Google.

To measure people's attention to economic issues, I calculate the monthly Google search share of 30 major U.S. media, such as CNN, Fox News, and so on. To focus on searches related to economic issues, I restrict my analysis to searches under the "Business and Industrial" category. This category contains queries such as "CNN Dow Jones," "CNN premarket," and more recently, searches such as "CNN coronavirus" and "CNN stimulus check."

1. Da et al. (2011) use Google search of stock tics to study retail investors' attention to cross-sectional stocks. Drake et al. (2012) use Google search to study investors' information demand around earnings announcements.

To the extent that these queries on Google represent the relative time and information processing capacity people assign to these issues relative to other alternative information in a period of time, this measure provides a proxy for the level of attention people assign to macroeconomic economic issues such as the coronavirus pandemic and financial crisis.

As shown in Figure B.1, the search share of major news media in the Business and Industrial category increases by 100% during the past two recessions, 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.

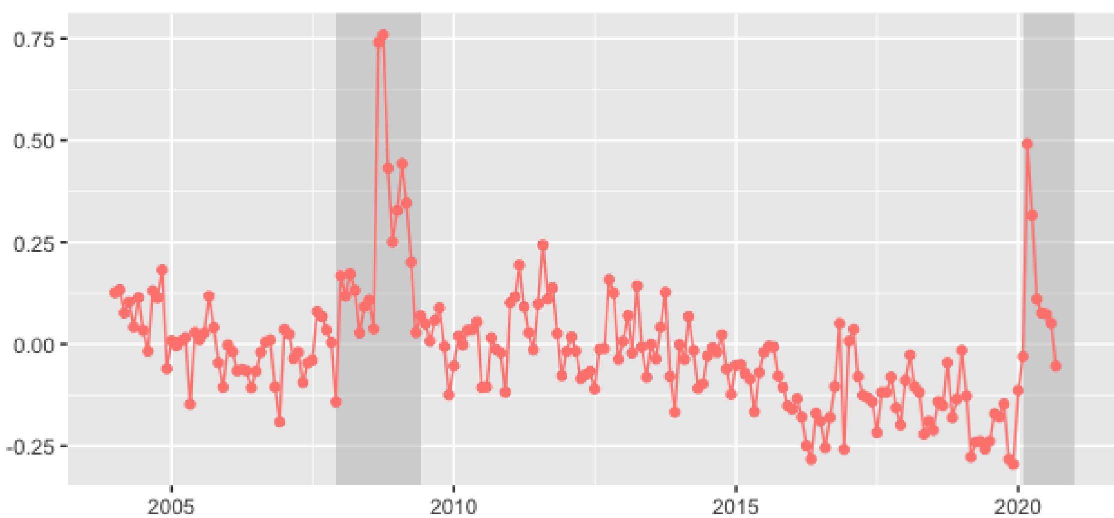


Figure B.1: Measure of attention over time: news search

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.

Alternative measures of attention

The fact that people pay higher attention to economic issues during recessions is robust to alternative proxies of attention. First, I construct an alternative measure using Google search share of a group of words classified as "economic words" in the textual-analytics literature,[2] including terms such as "bank," "unemployment," and "gold." 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. Second, I calculate the internet download traffic on SEC's EDGAR system, which contains the periodic report and announcement by all companies in the United States. The EDGAR system is one of the major sources of financial information and is often used by the empirical finance and accounting literature to reflect information demand from the corporate sector. Both measures are countercyclical and suggest the demand for economic information is higher during recessions.

2. See Da et al. (2015) for example.

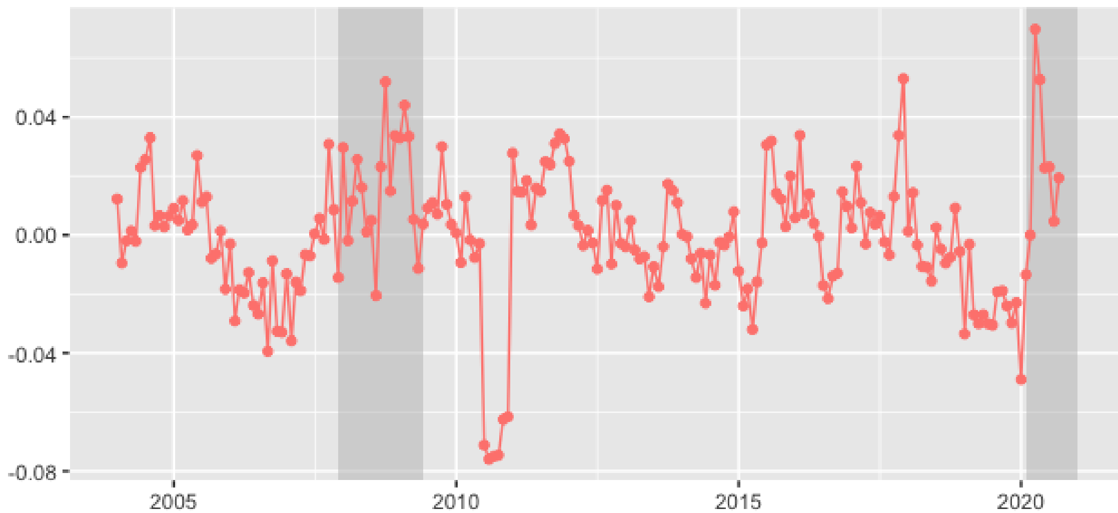


Figure B.2: Measure of attention over time: economic terms

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.

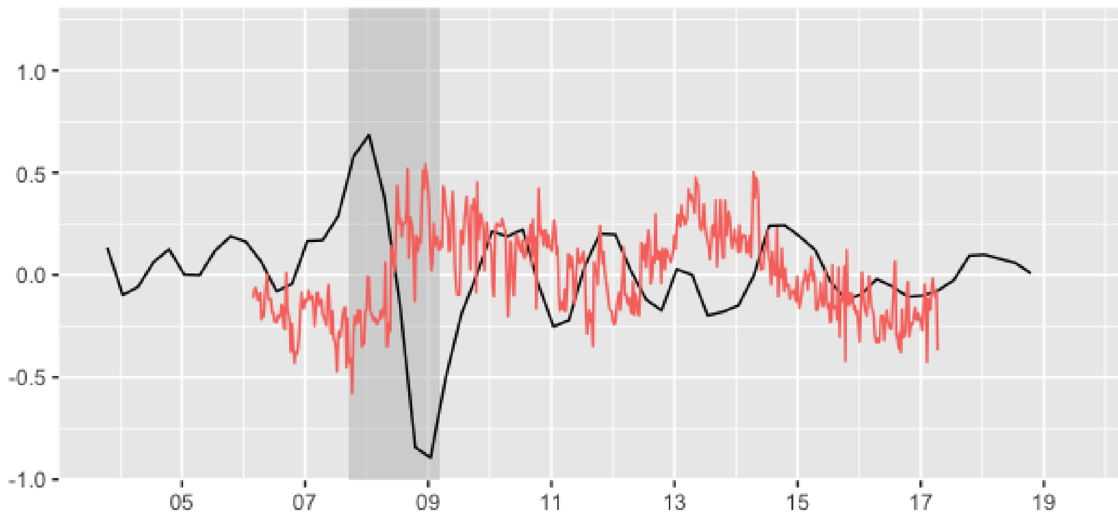


Figure B.3: Measure of attention over time: EDGAR filing

x-axis: time; y-axis: log-deviation of internet traffic on SEC's EDGAR website from its long-run average, seasonality adjusted and detrended ; gray area indicates NBER recessions.

Conditional volatility of TFP

This section demonstrate the potential problem of relying on countercylical shocks to the volatility of TFP as a source of uncertainty fluctuations.

Figure B.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 blue line comes from the unadjusted TFP series and the purple line comes from the utilization-adjusted TFP, both from Fernald (2014)'s data set. [3]

3. The exact GARCH specification does not affect the conclusion too much. For example, conditional volatility of utilization-adjusted TFP is acyclical with very small magnitude of fluctuation

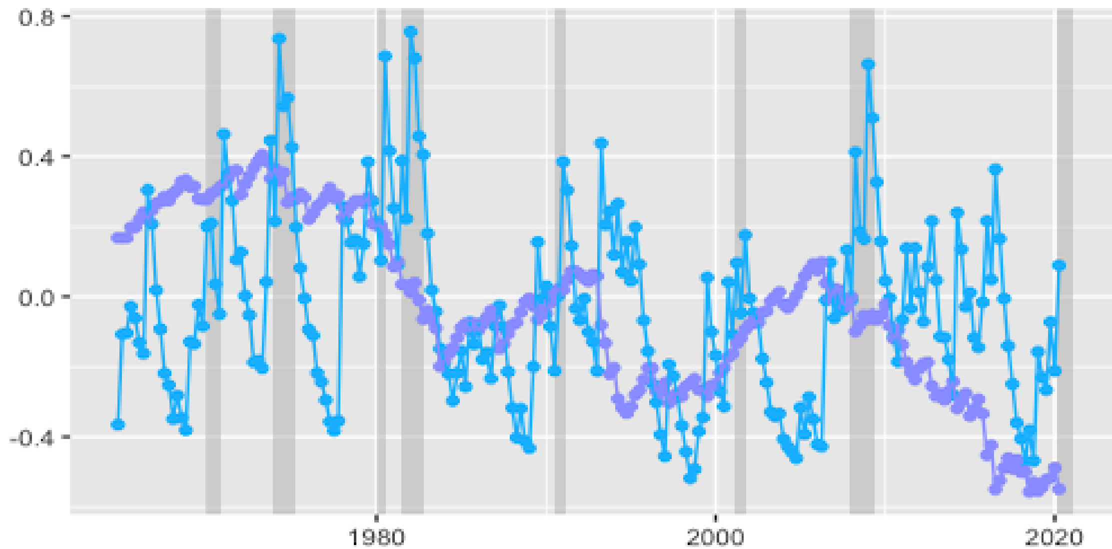


Figure B.4: Measure of aggregate TFP volatility over time

Conditional heteroskedasticity of TFP estimated with EGARCH(1,1)-ARMA(2,2). Blue line: not utilization adjusted; purple line: utilization adjusted. Both series are from Fernald (2014)'s data set. The result is robust to different GARCH and mean process 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 -.44 with filtered output. On the other hand, the utilization-adjusted series has a correlation of .16 with filtered output.

when I follow Bloom et al. (2018) and estimate with a GARCH(1,1)-ARMA(1,1) model.

BIBLIOGRAPHY

[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*. Tech. rep. National Bureau of Economic Research, 2018.

[5] George-Marios Angeletos and Zhen Huo. "Myopia and anchoring". In: *American Economic Review* 111.4 (2021), pp. 1166–1200.

[6] George-Marios Angeletos and Jennifer La'O. "Sentiments". In: *Econometrica* 81.2 (2013), pp. 739–779.

[7] George-Marios Angeletos and Jennifer La'o. "Noisy business cycles". In: *NBER Macroeconomics Annual* 24.1 (2010), pp. 319–378.

[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] Jess Benhabib et al. "Endogenous information acquisition and countercyclical uncertainty". In: *Journal of Economic Theory* 165 (2016), pp. 601–642.

[13] Anmol Bhandari et al. *Inequality, business cycles, and monetary-fiscal policy.* Tech. rep. National Bureau of Economic Research, 2018.

[14] Nicholas Bloom. "Fluctuations in uncertainty". In: *Journal of Economic Perspectives* 28.2 (2014), pp. 153–76.

[15] Nicholas Bloom. "The impact of uncertainty shocks". In: *econometrica* 77.3 (2009), pp. 623–685.

[16] Nicholas Bloom et al. "Really uncertain business cycles". In: *Econometrica* 86.3 (2018), pp. 1031–1065.

[17] Jaroslav Borovicka and Lars Peter Hansen. "Robust Preference Expansions". In: (2013).

[18] Lawrence J Christiano et al. "Risk shocks". In: *American Economic Review* 104.1 (2014), pp. 27–65.

[19] 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.

[20] Zhi Da et al. "In search of attention". In: *The Journal of Finance* 66.5 (2011), pp. 1461–1499.

[21] 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.

[22] 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.

[23] Pablo D Fajgelbaum et al. “Uncertainty traps”. In: *The Quarterly Journal of Economics* 132.4 (2017), pp. 1641–1692.

[24] John Fernald. “A quarterly, utilization-adjusted series on total factor productivity”. In: Federal Reserve Bank of San Francisco. 2014.

[25] Jesús Fernández-Villaverde et al. “Risk matters: The real effects of volatility shocks”. In: *American Economic Review* 101.6 (2011), pp. 2530–61.

[26] Joel P Flynn and Karthik Sastry. “Attention Cycles”. In: *Available at SSRN 3592107* (2020).

[27] Simon Gilchrist et al. *Uncertainty, financial frictions, and investment dynamics*. Tech. rep. National Bureau of Economic Research, 2014.

[28] 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.

[29] Zhen Huo and Marcelo Pedroni. “A single-judge solution to beauty contests”. In: *American Economic Review* 110.2 (2020), pp. 526–68.

[30] Zhen Huo and Naoki Takayama. “Rational expectations models with higher order beliefs”. In: *Manuscript, Yale Univ* (2015).

[31] 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.

[32] Kenneth L Judd. *Numerical methods in economics*. MIT press, 1998.

[33] Kyle Jurado et al. “Measuring uncertainty”. In: *American Economic Review* 105.3 (2015), pp. 1177–1216.

[34] Nicholas Kozeniauskas et al. “What are uncertainty shocks?” In: *Journal of Monetary Economics* 100 (2018), pp. 1–15.

[35] 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.

[36] Giovanni Lombardo and Harald Uhlig. “A Theory Of Pruning”. In: *International Economic Review* 59.4 (2018), pp. 1825–1836.

[37] Robert E Lucas. “Expectations and the Neutrality of Money”. In: *Journal of economic theory* 4.2 (1972), pp. 103–124.

[38] Bartosz Maćkowiak and Mirko Wiederholt. “Business cycle dynamics under rational inattention”. In: *The Review of Economic Studies* 82.4 (2015), pp. 1502–1532.

[39] Bartosz Mackowiak and Mirko Wiederholt. “Optimal sticky prices under rational inattention”. In: *American Economic Review* 99.3 (2009), pp. 769–803.

[40] Bartosz Maćkowiak et al. “Dynamic rational inattention: Analytical results”. In: *Journal of Economic Theory* 176 (2018), pp. 650–692.

[41] Bartosz Mackowiak et al. “Rational inattention: A disciplined behavioral model”. In: *CEPR Discussion Papers* 13243 (2018).

[42] 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.

[43] David P Myatt and Chris Wallace. “Endogenous information acquisition in coordination games”. In: *The Review of Economic Studies* 79.1 (2012), pp. 340–374.

[44] Kristoffer Nimark. “Dynamic higher order expectations”. In: (2017).

[45] Kristoffer P Nimark. “Man-bites-dog business cycles”. In: *American Economic Review* 104.8 (2014), pp. 2320–67.

[46] Anna Orlik and Laura Veldkamp. *Understanding uncertainty shocks and the role of black swans*. Tech. rep. National Bureau of Economic Research, 2014.

[47] Phelps. *Microeconomic foundations of employment and inflation theory*. WW Norton New York, 1970.

[48] Christopher A Sims. “Implications of rational inattention”. In: *Journal of monetary Economics* 50.3 (2003), pp. 665–690.

[49] Christopher A Sims. “Rational inattention and monetary economics”. In: *Handbook of monetary economics*. Vol. 3. Elsevier, 2010, pp. 155–181.

[50] Robert M Townsend. “Forecasting the forecasts of others”. In: *Journal of Political Economy* 91.4 (1983), pp. 546–588.

[51] Stijn Van Nieuwerburgh and Laura Veldkamp. “Learning asymmetries in real business cycles”. In: *Journal of monetary Economics* 53.4 (2006), pp. 753–772.

[52] 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.

[53] Michael Woodford. *Imperfect common knowledge and the effects of monetary policy*. Tech. rep. National Bureau of Economic Research, 2001.