Higher-Order Beliefs, Confidence, and Business Cycles*

Zhen Huo
Yale University

Naoki Takayama
Hitotsubashi University

October 7, 2022

Abstract

This paper presents a model where business cycles are driven by the confidence shock that shifts agents' beliefs about others' economic activities. Trade linkages and incomplete information admits belief-driven fluctuations even with fixed aggregate fluctuations, but also make equilibrium outcomes hinge on dynamic higher-order expectations. We provide an analytical solution that characterizes how aggregate dynamics are shaped by informational frictions and the general equilibrium consideration. When disciplined by the survey data on expectations, the model can match main business cycle moments and conditional responses identified in empirical VAR.

Keywords: Higher-order beliefs, Incomplete information, Business cycles.
JEL classifications: E20, E32, F44

*Huo: Yale University, 28 Hillhouse Ave, New Haven, CT, 06510, US, zhen.huo@yale.edu. Takayama: Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8603, Japan, ntakayama@ier.hit-u.ac.jp. Acknowledgements: We are thankful for our advisers José-Víctor Ríos-Rull, Ellen McGrattan, Jonathan Heathcote, and Elena Pastorino for their support. We are grateful for comments from Marios Angeletos, Anmol Bhandari, V.V. Chari, Patrick Kehoe, Stephen Morris, Fabrizio Perri, Thomas Sargent, and Kei-Mu Yi. We also thank Sushant Acharya, Manuel Amador, Yan Bai, Luigi Bocola, Katarina Borovickova, Jaroslav Borovicka, Fend Dong, Si Guo, Boyan Jovanovic, Kyle Herkenhoff, Guido Lorenzoni, Virgiliu Midrigan, Alessandro Peri, Edouard Schaal, Kjetil Storesletten, Gianluca Violante, Pierre-Olivier Weill and the participants of at the University of Minnesota, Federal Reserve Bank of Minneapolis, NYU Stern, New York University, Federal Reserve Bank of St. Louis, Princeton University, University of Chicago, Chicago Booth, Penn State University, University of Pennsylvania, Yale University, Duke University, Boston University, and New York University for helpful comments. We thank Flint O'Neil for his excellent research assistance. Naoki Takayama thanks the JSPS KAKENHI 21K13256 for financial support.

1. Introduction

The popular idea that business cycles are driven by waves of optimism and pessimism can be traced back to Keynes' famous animal spirits. Instead of relying on irrational spontaneous impulses, recent research has formalized this idea in a dispersed information framework (Lorenzoni, 2009; Angeletos and La'O, 2013; Benhabib, Wang, and Wen, 2015) where agents rationally respond to shocks to expectations that trigger aggregate fluctuations. These shocks are orthogonal to changes in fundamentals and could potentially serve as an important source of business cycle fluctuations. Despite the conceptual appeal, it remains a challenging task to fully analyze and quantify these types of models: with incomplete information and interdependence among agents, the equilibrium outcome hinges on both first-order and higher-order expectations, which are potentially infinite-dimensional objects. The goal of this paper is to overcome this difficulty and to provide an evaluation of the quantitative relevance of belief-driven fluctuations.

This paper makes three contributions. First, we build a framework in which the confidence shock—a shock to agents' beliefs about others productivity—is responsible for all aggregate fluctuations. We illustrate how dynamic higher-order expectations matter for equilibrium outcomes and how to bypass them to obtain an analytical solution. Second, we establish that persistent forecast errors that comove with the aggregate output are necessary for persistent aggregate fluctuations, a key property that guides the quantitative exercise and separates our theory from others'. Third, we show that the confidence shock can account for a significant amount of aggregate fluctuations when disciplined by the subjective beliefs in the survey data, and the propagation mechanism is consistent with the main business cycle shock identified in Angeletos, Collard, and Dellas (2020).

Framework with dynamic higher-order expectations. Our model economy features decentralized trading and information frictions, based on the structure specified in Angeletos and La'O (2013). A continuum of islands differing in productivity are randomly matched and trade with each other. Information frictions prevent agents from observing their trading partner's fundamentals, and only noisy signals about others' productivity are observed. With a positive noise, agents tend to overestimate their trading partners' productivity and the value of their own products, and consequently increase their own output due to trade linkages. If the noise is correlated across islands, it gives rise to economy-wide output fluctuations. We label this correlated noise as the confidence shock.

Notably, the equilibrium outcome is shaped by higher-order expectations in our model environment. With incomplete information and interdependence, agents not only care about the inference about others' productivity, but also others' inference about their own productivity and other's inference about their own inference about other's productivity, and so on. As the confidence shock induces systemic over- and under-estimation, the aggregate output in the end is a weighted average of both the first-order and higher-order *expectation errors* about the confidence shock. The weight that regulate their relative importance corresponds to the strategic complementarity or the general equilibrium

(GE) consideration, which increases in trade linkages.

With persistent confidence shock, higher-order expectations are more dampened and more sluggish than first-order ones as the former are more anchored by the common prior (Woodford, 2003; Nimark, 2008; Angeletos and Huo, 2021). It implies that that higher-order expectation errors about the confidence shock are more persistent and more responsive. These properties are crucial in shaping the dynamics of the aggregate output, but it is a daunting task to characterize all the higher-order expectations: their laws of motion become increasingly complex as the order increases, and the entire history of signals is relevant for inference.

Equilibrium characterization. Despite these complications, we show that the equilibrium outcome admits a surprisingly simple closed-form solution. The aggregate output follows an AR(1) process with an endogenously determined persistence. The persistence is increasing in informational frictions, as it takes longer to resolve both first-order and higher-order uncertainty about the confidence shock. The persistence is also increasing in the strength of GE consideration, as more weight is shifted towards higher-order expectation errors. The effects of incomplete information therefore loom more prominent and the role of higher-order beliefs is intensified.

The equilibrium outcome is also equivalent to a particular first-order expectation error where the expectation is conditional on modified signals with discounted precision. On the technical side, this observation highlights that the weighted sum of infinite higher-order expectation errors collapse to a tractable first-order expectation error at the fixed point, and agents in equilibrium only need to keep track of a small number of state variables. On the applied side, it underscores the propagation of the confidence shock is through the systemic errors when agents form expectations.

Connection with evidence on expectations. We further connect the equilibrium properties with the survey evidence on expectations. The behavior of the forecast errors about the aggregate outcome inherits that of the expectation errors about the confidence shock, while only the former are directly observed. Empirically, the aggregate forecast errors are persistent and positively correlated with outcomes. These patterns cannot be easily rationalized by economies with perfect or static information, but arise naturally in our model economy. We establish that persistent aggregate forecast errors are necessary for persistent aggregate fluctuations, which is a defining feature of our model. Meanwhile, when the degree of informational frictions varies, the persistence of the forecast error and the outcome comove with each other, which is a property that helps guide our quantitative exercise .

The joint distribution of the outcome and the forecast error also helps separate different approaches in forming expectations. With our approach, agents remain rational at the individual level, and the aggregate forecasts about the output under-react due to noises, consistent with the documented pattern in Coibion and Gorodnichenko (2015). Had we adopted the heterogeneous-prior approach (Angeletos and La'O, 2013; Angeletos, Collard, and Dellas, 2018), the forecasts would overreact at both individual and aggregate level. Despite these differences, the two approaches complement each other. Our theoretical results provide guidance on how to discipline the shock processes with the heterogeneous-prior approach by connecting with the forecast errors in the survey data. This strategy is easy to implement while still capturing the essential implications from the rational expectations benchmark, which can be a fruitful way going forward when analyzing aggregate implications of shocks to expectations.

Quantification. Our final contribution is to quantify the role of confidence shock in accounting for business cycle fluctuations. We extend the model to allow endogenous capital accumulation, and we estimate the shock process and the informational frictions by matching the forecast errors in the Survey of Professional Forecasters.

With only confidence shocks, the model is able to capture a significant fraction of business cycle fluctuations, with main aggregate variables (including labor wedge) comoving with output in the right direction. From an individual agent's perspective, the waves of optimism and pessimism about others' output induced by the confidence shock is isomorphic to variations in their own TFP. Put it differently, agents respond to these economy-wide shifts of beliefs as if there are aggregate TFP shocks, but these as-if TFP shocks never materialize. Therefore, the propagation mechanism and the success in matching business cycle moments is similar to that in a standard neoclassical growth model, albeit that the magnitude and the persistence of the as-if TFP shocks are endogenously determined and the shocks only exist in expectation.

The responses of main aggregate variables to the confidence shock resemble the responses to the main business cycle (MBC) shock identified in Angeletos, Collard, and Dellas (2020), even though we do not directly target these moments. The MBC shock is constructed to maximize the business-cycle variation in a macroeconomic variable and it explains the bulk of volatility and comovement in major macroeconomic variables. A key property of the MBC shock is that it is almost orthogonal to the TFP shock, and this holds true in our model economy by construction. The confidence shock can be interpreted as a demand shock, but it induces business cycle comovements as a supply shock, which is what the MBC shock asks for. In addition, we estimate the conditional response of the forecast error of unemployment rate to the MBC shock, and it is consistent with our model's prediction. Taking stock, we consider the confidence shock in our model as a promising micro foundation for the MBC shock.

Related literature. Our paper is most closely related to Angeletos and La'O (2013) and Angeletos, Collard, and Dellas (2018). On the theoretical side, our work differs from them by providing an analytical equilibrium characterization with dynamic information and intertemporal higher-order expectations. We show how the persistence of the equilibrium outcome is endogenously determined by the GE consideration and information frictions with rational expectations, instead of being fixed by the exogenous shock process when the heterogeneous-prior approach is taken. On the applied side, we document that the forecast error is the key to discipline the model. In addition to aggregate real variables, we move forward to show the model can simultaneously match the response of the aggregate forecasts to the MBC shock identified in Angeletos, Collard, and Dellas (2020), which brings additional validation for the confidence shock as an important source for aggregate fluctuations.

Our paper complements the literature on belief-driven business cycles. A large number of papers introduce exogenous noises that confound the aggregate fundamentals (Lorenzoni, 2009; Angeletos and La'O, 2010; Barsky and Sims, 2012; Nimark, 2014; Chahrour and Jurado, 2018). These models typically require the existence of uncertainty about aggregate fundamentals in the first place to make room for aggregate noises to be relevant, while only uncertainty at idiosyncratic level is required in our model economy. This property is similar to the literature where the noises are endogenously determined in the equilibrium (Benhabib, Wang, and Wen, 2015; Acharya, Benhabib, and Huo, 2021). When abstracting from the details of the information structure, Bergemann and Morris (2013), Bergemann, Heumann, and Morris (2015), and Chahrour and Ulbricht (2022) characterize the bound for the aggregate volatility in correlated equilibria, and Hébert and La'O (2020) provide the information criterion for the existence of noise-driven fluctuations. More broadly, our work also complements the literature that studies the business cycles driven by changes in ambiguity (Ilut and Schneider, 2014; Bhandari, Borovička, and Ho, 2019), though we focus on different facets of the survey data.

In terms of the shock propagation mechanism, our paper is related to the large literature that studies how incomplete information modifies the responses to fundamental shocks (Taub, 1989; Mankiw and Reis, 2002; Woodford, 2003; Maćkowiak and Wiederholt, 2009; Nimark, 2008; Amador and Weill, 2012; Venkateswaran, 2014; Angeletos and Lian, 2018; Angeletos and Huo, 2021). A subtle difference is that the equilibrium properties in this line of work often boil down to the properties of higher-order expectations, while what matter in our model economy are the properties of higher-order expectation errors. In addition, the presence of capital accumulation makes the beauty-contest game in our model both forward and backward looking, which adds additional challenge when characterizing the equilibrium.

Finally, our paper is related the fast growing literature on the survey evidence on expectation formation (Coibion and Gorodnichenko, 2015; Bordalo, Gennaioli, Ma, and Shleifer, 2020; Fuhrer, 2018; Kohlhas and Walther, 2021; Coibion, Gorodnichenko, Kumar, and Ryngaert, 2021; Angeletos, Huo, and Sastry, 2021). The prediction in our model is consistent with the generic feature documented in these papers that beliefs are dispersed and aggregate forecasts under-react to new information. When extending to allow extrapolative expectations, our model can also match the estimated sign-switching pattern of the forecast error.

The rest of the paper is organized as follows. Section 2 sets up a simple economy with an analytical solution and describes how the equilibrium outcome depends on higher-order expectations. Section 3 discusses the connection between the output and its forecast error, and how the properties of forecast errors can be used to distinguish alternative theories. Section 4 evaluates the quantitative performance of the full model when disciplined with survey data on expectations. Section 5 concludes.

2. An Analytical Model

In this section, we present a simple island model to introduce the confidence shock which triggers aggregate fluctuations. This model builds on Angeletos and La'O (2013), and we allow the signals to be persistent over time. This is a natural extension to make the model empirically relevant, but it also gives rise to an infinite number of persistent higher-order expectations. We illustrate how these higher-order expectations shape the equilibrium outcome and how to bypass them to obtain an analytical characterization.

2.1 Model Setup

The economy consists of a continuum of islands indexed by $i \in [0, 1]$. The total factor productivity on island i is a_i, which is drawn from a normal distribution $\mathcal{N}(0, \sigma_a^2)$ but fixed over time. Each island is populated by a continuum of identical households. In each household, there is a producer and a shopper. The producer decides how much to produce. The shopper then receives the output from the producer and makes transaction and consumption plans.

Every period, island i is randomly matched with another island. Households value both local and foreign goods, and they trade with the island they are matched with. There is no centralized market in the economy and all the trading is decentralized. Let $m(i, t)$ denote the index of island i's trading partner in period t. With a slight abuse of notation, we will use $j \equiv m(i, t)$ to denote the index of island i's contemporary trading partner. It should be clear that island j refers to a different island at each period.

The distribution of productivity across islands is fixed over time, but island i's specific trading partner changes every period. We assume that the production plan has to be made at the beginning of a period without the perfect knowledge of their trading partner's productivity level. Even though agents on each island understand that there is no change in aggregate fundamentals, they still face uncertainty due to the decentralized trading arrangement and communication frictions. The need to infer their trading partner's output and the lack of common knowledge leaves room for confidence shocks and higher-order expectations to play substantial roles.

Timing and Information. Each period has two stages: production and trade. At the beginning of the first stage when production occurs, island i is randomly matched with another island. Once the match is drawn, producers on island i receive two signals. The first signal x_{it}^1 is on their trading partner's productivity, but is corrupted by an economy-wide common noise ξ_t

$$x_{it}^1 = a_j + \xi_t, \tag{2.1}$$

where $a_j \sim \mathcal{N}(0, \sigma_a^2)$. Crucially, we assume that the common noise ξ_t follows a persistent process

$$\xi_t = \rho \xi_{t-1} + \eta_t, \tag{2.2}$$

where $\rho \in (0,1)$ and $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$. A positive (negative) realization of ξ_t induces all agents in the economy to feel optimistic (pessimistic) about their trading partner's productivity. Therefore, we label this common noise shock as a confidence shock.

The second signal x_{it}^2 provides private information on the confidence shock

$$x_{it}^2 = \xi_t + u_{it}, \tag{2.3}$$

where $u_{it} \sim \mathcal{N}(0, \sigma_u^2)$ is a private noise. The variance of u_{it} effectively regulates the severity of the informational friction: with $\sigma_u^2 = 0$, the economy is at its frictionless benchmark. Producers can observe ξ_t perfectly and figure out their trading partner's productivity without error. With $\sigma_u^2 > 0$, producers face first-order uncertainty in their trading partner's production capacity. Moreover, producers are unsure about how their trading partners on island j perceive the production decision on island i, and so on, which correspond to higher-order uncertainty.

The producers' information set on island i at time t includes its own productivity and all the signals received up to time t

$$\Omega_{it} = \left\{ a_i,\ x_{it}^1, x_{it-1}^1, x_{it-2}^1, \ldots,\ \ x_{it}^2, x_{it-1}^2, x_{it-2}^2, \ldots \right\}. \tag{2.4}$$

To fix notation, we use $\mathbb{E}_{it}[\cdot]$ to denote the expectation conditional on i's information up to period t, i.e., $\mathbb{E}_{it}[\cdot] = \mathbb{E}[\cdot|\Omega_{it}]$. Since trading histories and idiosyncratic noises differ across islands, producers on different islands inherit heterogeneous information sets. After observing the signals, the producers decide the output level Y_{it}, which completes the first stage of a period.

The second stage is when trade occurs. Shoppers on island i receive output from their producers and trade with shoppers from island j in a competitive goods market. In this stage, shoppers can observe the other island's output and productivity. To prevent information from being fully revealed, we assume that the current shoppers are replaced by new shoppers in the following period. Effectively, shoppers cannot communicate with producers after the transaction stage.

Remark. The assumption of shoppers' turnover is only a means to implement the idea that the communication between producers and shoppers is not perfect. Supposing that we allow imperfect communication between producers and shoppers, producers will receive another noisy signal on a_j and ξ_t, but this is conceptually equivalent to setting the variance of u_{it} to a smaller value. In the end, what matters is how much producers can learn, but not exactly how they learn.

Shoppers' Problem. Let us proceed backward from the second stage. In the trading stage, goods markets are competitive and the prices for local goods and foreign goods are P_{it} and P_{jt} respectively.

The shoppers on island i receive the output Y_{it} produced in the first stage and they solve the following static problem maximizing the amount of composite consumption goods enjoyed locally:

$$\max_{C_{ii,t}, C_{ij,t}} \left(\frac{C_{ii,t}}{\omega}\right)^{\omega} \left(\frac{C_{ij,t}}{1-\omega}\right)^{1-\omega}, \qquad \text{subject to} \qquad P_{it} C_{ii,t} + P_{jt} C_{ij,t} = P_{it} Y_{it},$$

where $C_{ii,t}$ is local consumption goods, $C_{ij,t}$ is foreign consumption goods, and $\omega \in (0,1)$ determines the degree of home bias.

In equilibrium, local and foreign consumption are equal to a fixed fraction of local and foreign output, thanks to the Cobb-Douglas preference. It is straightforward to show that the equilibrium allocation satisfies

$$C_{ii,t} = \omega Y_{it}, \qquad C_{ij,t} = (1-\omega) Y_{jt},$$

and the terms of trade is given by

$$\frac{P_{it}}{P_{jt}} = \frac{Y_{jt}}{Y_{it}}. \tag{2.5}$$

Note that the terms of trade is increasing in the foreign output, which connects the production decision between the two matched islands.

Producers' Problem. Now we turn to the first stage. From the perspective of producers on island i, the value of one additional unit of local output can be expressed as a function of the terms of trade

$$\mathcal{M}_{it} \equiv \left(\frac{C_{ij,t}}{C_{ii,t}} \frac{\omega}{1-\omega}\right)^{1-\omega} = \left(\frac{P_{it}}{P_{jt}}\right)^{1-\omega}. \tag{2.6}$$

As a result, the optimal production scale depends on their trading partners' output. When informational frictions are absent, the productivities on both islands become common knowledge, and the outcomes are uniquely determined by the fundamentals. When information is incomplete, the output level on island i instead depends on the *expected* output level of their trading partner.

In this section, we abstract from the endogenous capital accumulation and assume that the only variable input is labor. The producers on island i choose output Y_{it} and labor N_{it} to maximize the expected utility in the current period:

$$\max_{Y_{it}, N_{it}} \mathbb{E}_{it} \left[\left(\frac{P_{it}}{P_{jt}}\right)^{1-\omega} Y_{it} - N_{it}^{1+\gamma} \right], \qquad \text{subject to} \qquad Y_{it} = \exp(a_i)\, N_{it}^{\theta},$$

where $\gamma > 0$ is the inverse of the Frisch elasticity and $\theta \in (0,1)$ is the output elasticity with respect to labor. Importantly, when expected Y_{jt} increases, the terms of trade improves, which encourages producers on island i to produce more output. This logic becomes transparent in the following first-order condition for the output level

$$Y_{it} = \left(\frac{\theta}{1+\gamma}\right)^{\frac{\theta}{1+\gamma-\omega\theta}} \exp\left(\frac{1+\gamma}{1+\gamma-\omega\theta}\, a_i\right) \mathbb{E}_{it}\left[Y_{jt}^{1-\omega}\right]^{\frac{\theta}{1+\gamma-\omega\theta}}, \tag{2.7}$$

where we have invoked condition (2.5) for the terms of trade and the fact that all producers within the same island are identical. The output is increasing in both producers' own fundamental a_i and the expected output of their trading partner.

Beauty-contest game. Condition (2.7) can also be interpreted as the best response function in a two-player game. To see this more clearly, we work with its linearzied version. Let y_{it} denote the log deviation of island i's output from the steady-state value of the aggregate output, condition (2.7) becomes

$$y_{it} = \varphi a_i + \alpha \mathbb{E}_{it}[y_{jt}], \tag{2.8}$$

where $$\varphi = \frac{1+\gamma}{1+\gamma-\omega\theta}, \qquad \alpha = \frac{\theta(1-\omega)}{1+\gamma-\omega\theta}.$$

This is a beauty-contest game closely related to those explored in Morris and Shin (2002), Woodford (2003), Angeletos and La'O (2010), and Bergemann and Morris (2013). Here, $\varphi > 1$ captures the partial equilibrium (PE) consideration holding their trading partner's output constant. On the other hand, $\alpha \in (0, 1)$ captures the general equilibrium (GE) consideration or the strategic complementarity, which is increasing in θ and decreasing in ω and γ. Intuitively, the dependence on their trading partner is stronger when the foreign output is more important in local households' consumption basket, and when labor is more responsive and plays a bigger role in production.

2.2 Higher-Order Expectations on the Confidence Shock

To understand the properties of the equilibrium outcome, consider momentarily the frictionless benchmark where the underlying shocks can be perfectly observed.

Proposition 2.1. *When $\sigma_u^2 = 0$, the output on island i is*

$$y_{it}^* = \varphi a_i + \alpha y_{jt} = \frac{\varphi}{1-\alpha^2}\left(a_i + \alpha a_j\right),$$

and the aggregate output y_t remains constant

$$y_t = \int y_{it} = 0.$$

A direct implication of Proposition 2.1 is that idiosyncratic outputs are pined down by fundamentals only. The confidence shock ξ_t plays no role in determining the aggregate outcome.

Next, consider the case with informational frictions. To infer the output on island j, producers on island i needs to infer the productivity on island j. But this is not the end. As the same logic applies to island j, to infer island j's output, it is also necessary to take into account island j's expectations about island i's productivity, which is a second-order expectation. Due to the lack of common knowledge, the decision ultimately depends on all the relevant higher-order expectations. These expectations can be mapped to the expectations about the confidence shock, which leads to the following result.

Proposition 2.2. *The output on island i can be expressed as*

$$y_{it} = y_{it}^* + \frac{\varphi}{1+\alpha} \sum_{k=1}^{\infty} \alpha^k (\xi_t - \mathbb{E}_{it}^k[\xi_t]), \tag{2.9}$$

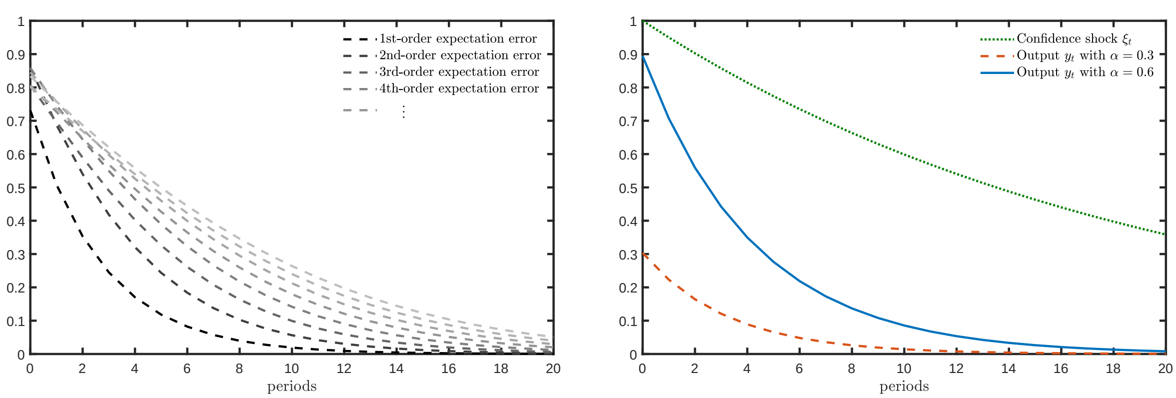
where the higher-order expectations are defined recursively as

$$\mathbb{E}_{it}^1[\xi_t] = \mathbb{E}_{it}[\xi_t], \quad \mathbb{E}_{it}^2[\xi_t] = \mathbb{E}_{it}[\mathbb{E}_{jt}[\xi_t]], \quad \textit{and} \quad \mathbb{E}_{it}^k[\xi_t] = \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}^{k-2}[\xi_t], \quad \textit{for } k > 2.$$

The aggregate output can be expressed as

$$y_t = \frac{\varphi}{1+\alpha} \sum_{k=1}^{\infty} \alpha^k \left(\xi_t - \int \mathbb{E}_{it}^k[\xi_t] \right). \tag{2.10}$$

Condition (2.9) makes it clear that relative to the frictionless benchmark, the expectation errors about the confidence shock cause fluctuations in aggregate output. If ξ_t is underestimated, then producers tend to overestimate their trading partners' productivities. Through trade linkages, all islands increase their own output because they expect a higher output from their trading partners, and therefore an economy-wide boom occurs. Furthermore, because of the GE consideration, not only first-order uncertainty but also higher-order uncertainty matters.



(a) average expectation errors (b) aggregate output

Figure 1: Response to the Confidence Shock

Note: the parameters used to generate this figure are: $\varphi = 1.2$, $\rho = 0.95$, $\sigma_a = \sigma_u = 5$.

At the aggregate level, the responses to idiosyncratic productivities wash out. However, the aggregate output still inherits the correlated responses to the confidence shock. As emphasized in Woodford (2003), Nimark (2008), and Angeletos and Huo (2021), higher-order expectations are typically more dampened in terms of magnitude and more sluggish than first-order expectations. In our setting, what matter for the aggregate outcome are the expectation errors. Figure 1a plots the impulse responses of $\xi_t - \int \mathbb{E}_{it}^k[\xi_t]$ for different k to an innovation to the confidence shock. As the order increases, the expectation errors becomes more persistent and more responsive. In the end, the output is a weighted average of the first-order and higher-order expectation errors, where their relative importance is determined by the strategic complementarity α.

First-order expectation error. Before solving for the equilibrium outcome, it is useful to characterize the following first-order expectation error for comparison.

Lemma 1. *The aggregate first-order expectation error is*

$$\xi_t - \overline{\mathbb{E}}_t[\xi_t] = \frac{\lambda}{\rho}\frac{1}{1-\lambda L}\eta_t, \tag{2.11}$$

where $\tau_a = \frac{\sigma_\eta^2}{\sigma_a^2}$, $\tau_u = \frac{\sigma_\eta^2}{\sigma_u^2}$, *and* $\lambda \in (0, \rho)$ *is given by*

$$\lambda = \frac{1}{2}\left[\left(\frac{1}{\rho} + \rho + \frac{\tau_a + \tau_u}{\rho}\right) - \sqrt{\left(\frac{1}{\rho} + \rho + \frac{\tau_a + \tau_u}{\rho}\right)^2 - 4}\right]. \tag{2.12}$$

Here, λ captures the persistence of the expectation error, which is *independent* of the GE consideration α. As shown in Figure 1a, the aggregate output remains positive even after the first-order expectation error has vanished. That is, the confidence shock generates aggregate fluctuations when agents can accurately predict it. To understand this result, note that an individual agent may be aware of the confidence shock, while still have doubts about whether other agents are aware of this fact at the same time. When the GE consideration is strong, it allows higher-order expectation errors to have a long-lasting effect on agents beliefs about others. This argument hints that the persistence of output is higher than λ, which we confirm in the next subsection.

2.3 Equilibrium Characterization

Though Proposition 2.2 provides a useful perspective to understand the effects of the confidence shock, it is impossible to compute and to keep track of all the higher-order expectations. In what follows, we bypass the higher-order expectations and proceed to characterize the equilibrium outcome directly. The challenge is that ex ante, it is not clear whether there exists any sufficient statistics to summarize the history. To overcome this difficulty, we apply the method in Huo and Takayama (2021) and obtain a finite-state representation of the equilibrium outcome.

Proposition 2.3. *The equilibrium outcome is*

$$y_t = \alpha\chi\frac{\vartheta}{\rho}\frac{1}{1-\vartheta L}\eta_t, \tag{2.13}$$

where $\tau_a = \frac{\sigma_\eta^2}{\sigma_a^2}$, $\tau_u = \frac{\sigma_\eta^2}{\sigma_u^2}$, ϑ *and* χ *are given by*

$$\vartheta = \frac{1}{2}\left[\left(\frac{1}{\rho} + \rho + \frac{(1-\alpha)(\tau_a+\tau_u)}{\rho}\right) - \sqrt{\left(\frac{1}{\rho} + \rho + \frac{(1-\alpha)(\tau_a+\tau_u)}{\rho}\right)^2 - 4}\right], \tag{2.14}$$

$$\chi = \frac{\varphi\rho(\tau_a+\tau_u)}{\rho(\tau_a+\tau_u) - \alpha^2(\rho\tau_u + \vartheta\tau_a)}. \tag{2.15}$$

Even though the higher-order expectations are with different persistence, the eventual equilibrium outcome follows a simple AR(1) process. The weighted average of all higher-order expectation errors turns out to be much simpler than each single one of them—a particularly convenient feature of the fixed point. Figure 1b visualizes the response of the aggregate output to the confidence shock. In the beginning, agents underestimate the confidence shock on average and consequently, they overestimate their trading partners' productivity and output. Due to strategic complementarity, their best response is to increase their own output, resulting in an increase in aggregate output. The confusion will not be resolved immediately. Agents gradually learn the true state of the economy, and during this process, the output remains above its steady state.

Condition (2.13) can be decomposed into three parts. The confidence shock plays a role only if agents care about others' output, namely when the complementarity α is positive. The constant χ then measures to which extent agents' trading partners' output responds to their own productivity a_j. The last term, $\frac{\vartheta}{\rho}\frac{1}{1-\vartheta L}$, represents the effects of over (under) estimation of the productivity, which resembles the form of the first-order expectation error (2.11) by replacing λ with the endogenous persistence ϑ. To better appreciate the last point, the following corollary provides an alternative representation of the equilibrium outcome.

Corollary 1. *The equilibrium policy rule can be expressed as*

$$y_t = \alpha\chi(\xi_t - \widetilde{\mathbb{E}}_t[\xi_t]). \tag{2.16}$$

The average expectation $\widetilde{\mathbb{E}}_t[\xi_t]$ *is conditional on a modified signal process where the variance of* a_{jt} *and* u_{it} *are amplified to* $\sigma_a^2(1-\alpha)^{-1}$ *and* $\sigma_u^2(1-\alpha)^{-1}$*, respectively.*

Corollary 1 states that the nature of the equilibrium outcome is effectively a first-order expectation error about the confidence shock based on a particular average expectation. The GE footprint lies in the modified precision: with a higher α, agents behave as if the signals are nosier, which makes the effects of the confidence shock loom more prominent. This representation is reminiscent of the single-agent solution in Huo and Pedroni (2020). In Appendix A.5, we further illustrate that this observation extends to a more general information structure.

Persistence and volatility of output. The persistence of aggregate output ϑ provided in equation (2.14) is different than the exogenous persistence of the confidence shock. As mentioned earlier, this endogenous persistence encapsulates the effects of higher-order expectations and their iterations with the general equilibrium considerations. Its properties are summarized in the following proposition.

Proposition 2.4. *The endogenous persistence $\vartheta \in (\lambda, \rho)$ is increasing in α and σ_u.*

Proposition 2.4 first states that ϑ is bounded from above by the persistence of the confidence shock ρ. Intuitively, the aggregate output deviates from its steady state only when there is an average expectation error about the confidence shock. As agents learn over time, the error has to be less persistent compared with the confidence shock itself. At the same time, ϑ is bounded from below by λ from condition (2.11), the persistence of the first-order expectation error. As higher-order expectation errors are more persistent, the output has to be more persistent as well.

Secondly, holding the degree of the informational friction fixed, the endogenous persistence is higher with a stronger strategic complementarity. Recall that higher-order expectation errors are more persistent than first-order expectation errors. A larger α implies that agents have a stronger GE consideration and put more emphasis on higher-order uncertainty. Therefore, the equilibrium outcome resembles the patterns of higher-order expectation errors.

Thirdly, as expected, with more severe informational friction, agents learn slower and face more uncertainty. The expectation errors become larger, and the effects of the confidence shock are amplified.

Similar to the persistence, the volatility of output is increasing in the uncertainty about the confidence shock. As higher-order expectation errors are more responsive than first-order ones, the volatility of output is also higher with a stronger GE consideration. This argument is formalized in the following proposition.

Proposition 2.5. *The variance of output is increasing in α and σ_u.*

2.4 Endogenous Information

So far, we have assumed that the signal process is exogenously determined and independent of agents' actions. In this subsection, we show that allowing endogenous information aggregation is not likely to modify the model's main implications but will significantly reduce its tractability.

Specifically, instead of observing a noisy signal about the confidence shock, we allow agents to observe a noisy endogenous signal each period:

$$x^1_{it} = a_j + \xi_t, \quad \text{and} \quad x^2_{it} = y_t + \varepsilon_{it}. \tag{2.17}$$

The first signal is the same as before, while the second signal contains the aggregate output that is the result of the agents' forecasting problem. Since the aggregate output is driven by the confidence shock, the second signal provides relevant information, though its informativeness is endogenously determined in the equilibrium. In addition to the fixed-point problem in the best response, the endogeneity of information imposes an additional fixed-point requirement that the perceived law of motion of the signal process has to be consistent with its implied actual outcomes. Following the strategy in Huo and Takayama (2021), we show that the endogenous outcome does not permit a finite-state process.

Proposition 2.6. *With endogenous signal process (2.17), the aggregate output does not admit a finite ARMA representation.*

Though the equilibrium outcome no longer permits a closed-form solution, it should be clear that the endogenous-information equilibrium can be viewed as a particular exogenous-information equilibrium. To see this, note that for each individual producer, they do not care about how the signal process is determined and simply take its law of motion as given. In order to quantify the role of the confidence shock, what really matters is whether the aggregate outcome is significantly altered with endogenous information aggregation.

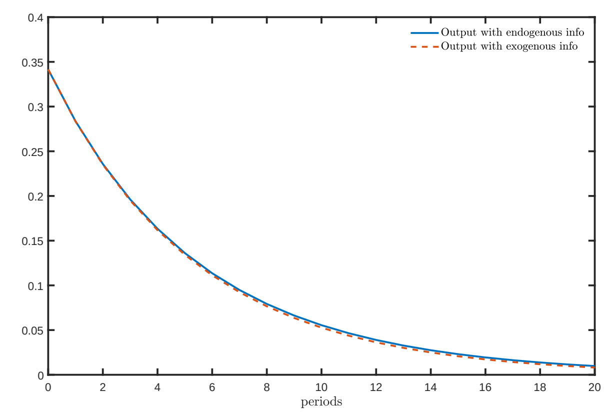


Figure 2: Output Response with Endogenous Information

Notes: The parameters used to generate this figure are: $\varphi = 1.2$, $\alpha = 0.3$, $\rho = 0.95$, $\sigma_\eta = 1$, $\sigma_a = \sigma_\varepsilon = 5$. To match the response on impact in the economy with endogenous information, the required level of noise in the economy with exogenous information is $\sigma_u = 15.8$.

It turns out that for the environment we have considered, the dynamics of the aggregate output with endogenous information can be well approximated by the equilibrium process with exogenous information (2.13). Figure 2 compares the responses of outputs to the confidence shock with endogenous and exogenous information. We choose the variance of the private noise u_{it} in the economy with exogenous information so that the response on impact is identical to that in the endogenous-information economy. Though our proof indicates these two processes cannot be identical, they trace each other very closely and are difficult to be distinguished. We have also verified this similarity holds for a large range of parameter values.

Since our purpose is to understand the effects of incomplete information in propagating business cycle fluctuations, whether the source of information is endogenous or not is less crucial, while the benefit of a transparent analytical solution is significant. That being said, for other purposes, such as to study how information gets aggregated through prices or other endogenous indicators, or to analyze the signaling effects of policies, endogenizing information can be essential.

3. Aggregate Fluctuations and Forecast Errors

In this section, we discuss how the equilibrium outcomes in our model are related to observed moments on expectations. Particularly, we show that aggregate fluctuations have an intimate relationship with the forecast error of output, a property that helps guide our quantification exercise in the next section. The joint pattern between output and its forecast error also helps illustrate how our results complement and differ from alternative approaches to modeling expectation formation.

3.1 Persistent Forecast Errors

As illustrated in Corollary 2.4, the aggregate output is proportional to a modified expectation error about the confidence shock. In this subsection, we further establish that this insight to a large extent can be carried over to the forecast error of aggregate output, which has more direct empirical counterparts. The following lemma characterizes the forecast error.

Lemma 2. *Let $\{\zeta_k\}_{k=1}^{\infty}$ be the impulse response function of the average one-step-ahead forecast error of output*

$$\zeta_k \equiv \frac{\partial(y_{t+k} - \overline{\mathbb{E}}_{t+k-1}[y_{t+k}])}{\partial \eta_t} = \frac{\vartheta \chi}{\rho}\left(\lambda^k - (1-\alpha)\vartheta^k\right),$$

and ζ_k is positive on impact, reaches zero after a finite number of periods, and converges to zero eventually.

Lemma 2 states that the forecast error is persistent given that the confidence shock is not transitory ($\rho > 0$) and noises are present. Figure 3a plots the output and the forecast error. Overall, the forecast errors track the outcomes well and the two series are positively correlated. The dark-color lines correspond to IRFs with more noisy signals, and the light-color lines correspond to those with less noisy signals. It is clear that more persistent forecast errors are associated with more persistent output. When the confusion lasts longer, the length of the boom extends as well. This intuition is summarized in the following proposition.

Proposition 3.1. 1. *Persistent aggregate fluctuations exist only if the forecast error $\left(\{\zeta_k\}_{k=1}^{\infty}\right)$ is persistent.*

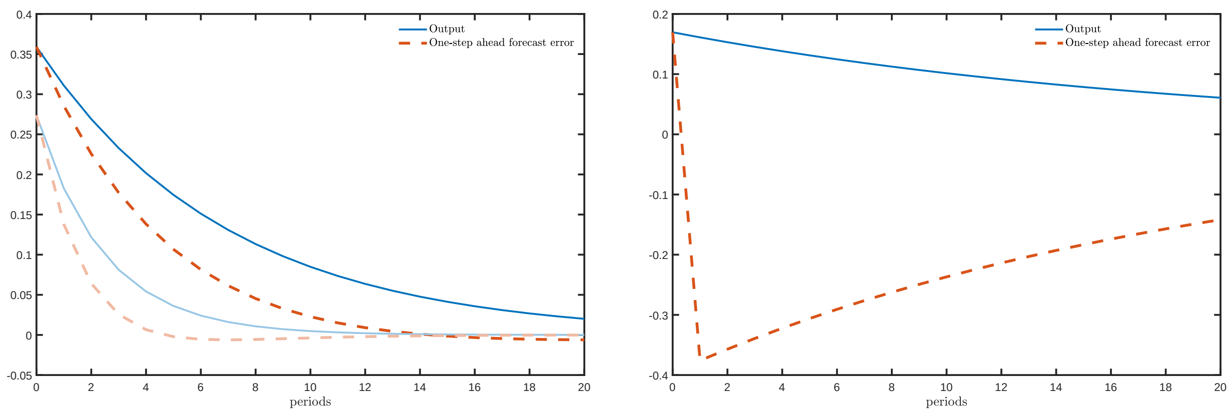
2. *Denote the time takes for the forecast error to reach zero as T*

$$T = \left[\frac{\log(1-\alpha)^{-1}}{\log\vartheta - \log\lambda}\right].$$

As informational frictions vary, the persistence and volatility of the output is positively correlated with T.

Part 1 of Proposition 3.1 highlights the key property of our model that aggregate fluctuations are results of persistent forecast errors about others' activities, rather than persistent aggregate fundamentals. This is in contrast with standard DSGE models where exogenous fundamentals such as TFP or monetary policy shocks drive business cycles and forecast errors disappear once innovations are observed. This is also different from models with "noisy business cycles" (Lorenzoni, 2009; Angeletos and La'O, 2010) where persistent exogenous fundamentals are still prerequisites for belief shocks to play a role in shaping aggregate fluctuations.

The time index T can be viewed as a proxy for the persistence of the forecast error. Part 2 of Proposition 3.1 establishes that as information frictions become more severe, the persistence of both the output and the forecast error increase. The forecast error therefore can be treated as a device that helps gauge the magnitude of informational frictions and guide the parameterization in our quantitative model.



(a) rational expectations

(b) heterogeneous prior

Figure 3: Forecast Error of Output

Note: the parameters used to generate this figure are: $\varphi = 1.2$, $\alpha = 0.3$, $\rho = 0.95$, $\sigma_\eta = 1$, $\sigma_a = \sigma_u = 5$.

These results indicate that in order to discipline our theory, it is important to utilize the survey data on expectations. Particularly, the forecast error can serve as an informative statistics when determining the process of the confidence shock and the informational friction. In Section 4.1, we implement this strategy by connecting our model with the data on forecast errors and quantifying the confidence-driven business cycle fluctuations.

3.2 Comparison with Bounded Rationality

In our model solution, agents are fully rational and they utilize the entire history of signals to form beliefs. In contrast, Angeletos and La'O (2013) and Angeletos, Collard, and Dellas (2018) propose a heterogeneous-prior formulation which accommodates higher-order doubt and is computationally more convenient as there is no need to record previous signals. In a nutshell, agents can observe both the confidence shock ξ_t and productivity a_{jt} perfectly, but they believe their trading partners observe a_i with a bias ξ_t, that is,

$$\mathbb{E}_{it}[a_{jt}] = a_{jt}, \qquad \mathbb{E}_{it}[\mathbb{E}_{jt}[a_i]] = a_i + \xi_t. \tag{3.1}$$

Conceptually, the confidence shock entering equation (3.1) modifies agents' higher-order expectations about their trading partners' productivity, which induces waves of optimism and pessimism as in our baseline specification. These two approaches, however, have rather different predictions in terms of forecast errors of aggregate output.

Proposition 3.2. *With heterogeneous priors, the aggregate output follows*

$$y_t = \frac{\varphi\alpha^2}{(1-\alpha)(1-\alpha^2)}\xi_t. \tag{3.2}$$

The IRF of the one-step ahead forecast error of output, $\{\zeta_k\}_{k=1}^{\infty}$, is given by

$$\zeta_k = -\rho^k \frac{\varphi\alpha}{1-\alpha^2} < 0, \tag{3.3}$$

which is negatively correlated with the IRF of output.

In terms of the outcome, both the process (3.2) and our equilibrium process (2.13) are driven by the innovation to the confidence shock and take an AR(1) form. If one only observes aggregate outcomes, these two processes cannot be easily distinguished.

Despite this similarity, there remain at least two main differences between these two approaches. First, the equilibrium outcome under the rational expectations approach is endogenous to both the informational frictions and the GE consideration. Therefore, the persistence and volatility of output is amplified with a stronger trading linkage or more noisy signals. With heterogeneous prior, however, both the persistence and volatility are independent of the information frictions, and the persistence is solely determined by the property of the exogenous shock.

Secondly, these two approaches have different implications on agents' expectations both at the macro level and the micro level. In our model with rational expectations, in response to a positive confidence shock, agents will overestimate their individual trading partner's output but underestimate the aggregate output. As a result, the aggregate forecast error is positively correlated with the output. In contrast, with heterogeneous prior, agents overestimate both their trading partner's and the aggregate output, leaving a negative correlation between the forecast error and the output. Figure 3b illustrates this point by showing the output and the output forecast error move in the opposite direction in response to the confidence shock.

As discussed in Huo and Takayama (2021), through the lens of the correlation between forecast error and forecast revision (Coibion and Gorodnichenko, 2015; Bordalo, Gennaioli, Ma, and Shleifer, 2020), the rational expectations benchmark features an under-reaction of the forecasts only at the aggregate level, while the forecasts under the heterogeneous-prior approach over-react at both the individual and the aggregate level. The documented empirical evidence from the survey of expectations indicates that the aforementioned correlation is higher at the aggregate level, which is more consistent with the pattern under the rational expectations framework.

Having recognizing these differences, these two approaches complement each other. The main takeaway from our theoretical results is that the persistence of the output resembles that of the forecast error. Following this theoretical guidance, for quantitative purposes, one may discipline the shock process in models with heterogeneous prior by the forecast error in the data. At the aggregate level, this could serve as a useful shortcut to capture the main effects of shocks to expectations while exploiting the essential implications from the rational expectations benchmark.

4. Quantification

In this section, we extend the model with endogenous capital stock accumulation to evaluate its quantitative performance. The key parameters are disciplined by the survey data on expectations. We show that the model can capture important business cycle properties. Notably, the responses to the confidence shock resemble those in the empirical VAR identified with the main business cycle shock (Angeletos, Collard, and Dellas, 2020).

4.1 Model

In our baseline dynamic model, we maintain the same assumption on the information structure and the shock processes as in Section 2, and alternative specifications are explored in Subsection 4.5. We focus on how the environment and agents' best response are modified when introducing the intertemporal decision of investment.

Producers in the first stage choose not only the scale of production, but also the amount of composite investment goods I_{it} for capital accumulation. It follows that in the second stage, shoppers need to purchase both consumption goods and investment goods. The shoppers' problem on island i becomes

$$\max_{C_{ii,t},C_{ij,t},I_{ii,t},I_{ij,t}} \left(\frac{C_{ii,t}}{\omega}\right)^{\omega}\left(\frac{C_{ij,t}}{1-\omega}\right)^{1-\omega},$$

subject to

$$P_{it}(C_{ii,t}+I_{ii,t})+P_{jt}(C_{ij,t}+I_{ij,t})=P_{it}Y_{it},$$
$$\left(\frac{I_{ii,t}}{\omega}\right)^{\omega}\left(\frac{I_{ij,t}}{1-\omega}\right)^{1-\omega}=I_{it}.$$

The goods market clearing condition continues to imply that the terms of trade is determined by the relative output, $\frac{P_{it}}{P_{jt}} = \frac{Y_{jt}}{Y_{it}}$, and the optimal consumption plan satisfies

$$C_{ii,t} + \omega I_{i,t} = \omega Y_{it}, \qquad C_{ij,t} + (1-\omega) I_{i,t} = \omega Y_{jt}.$$

Therefore, one additional unit of local output increases the amount of composite consumption goods enjoyed locally by $\mathcal{M}_{it} = \left(\frac{P_{it}}{P_{jt}}\right)^{1-\omega}$ units, and one additional unit of investment I_{it} decreases the composite consumption goods by one unit.

From producers' perspective, their problem is to choose a state contingent plan for Y_{it}, K_{it+1} and N_{it} to maximize the expected utility

$$\max_{Y_{it}, N_{it}, K_{it+1}, I_{it}} \mathbb{E}_{i0} \sum_{t=0}^{\infty} \beta^t \frac{\left(\mathcal{M}_{it} Y_{it} - I_{it} - N_{it}^{1+\gamma}\right)^{1-\sigma}}{1-\sigma},$$

subject to

$$Y_{it} = \exp(a_i)\, K_{it}^{1-\theta} N_{it}^{\theta}, \qquad \text{and} \qquad K_{it+1} = (1-\delta) K_{it} + I_{it} - \Xi(I_{it}, K_{it}).$$

For the preference, we adopt the GHH utility function (Greenwood, Hercowitz, and Huffman, 1988). Similar to Jaimovich and Rebelo (2009) where weakened income effects are necessary for a boom in response to goods news about future productivities. This assumption helps encourage the labor supply to increase in response to a positive confidence shock in our context. We also assume that the capital accumulation is subject to standard quadratic adjustment costs $\Xi(I_{it}, K_{it})$ with the following functional form

$$\Xi(I_{it}, K_{it}) = \frac{\Xi}{2} \left(\frac{I_{it}}{K_{it}} - \delta\right)^2 K_{it}.$$

The first order condition with respect to Y_{it} is

$$\mathbb{E}_{it}\left[\left(\mathcal{M}_{it} Y_{it} - I_{it} - N_{it}^{1+\gamma}\right)^{-\sigma} \left(\mathcal{M}_{it} \theta \frac{Y_{it}}{N_{it}} - (1+\gamma) N_{it}^{\gamma}\right)\right] = 0. \tag{4.1}$$

It is useful to recognize that the average of the marginal value of output $\mathcal{M}_{it}$ can be interpreted as the labor wedge. Linearizing condition (4.1) gives

$$\mathbb{E}_{it}[\log \mathcal{M}_{it}] = -(\log \mathrm{MPL}_{it} - \log \mathrm{MRS}_{it}).$$

With full information, the realization of the marginal value of local output $\mathcal{M}_{it}$ is disturbed only by the randomness of matches between islands, so its economy-wide average remains constant. However, with incomplete information, the expected value of $\mathcal{M}_{it}$ varies with the confidence shock in all islands, which maps to shifts of the measured labor wedge at the aggregate level.

Turn to the intertemporal decision, the optimal capital accumulation satisfies

$$\frac{\mathbb{E}_{it}\left[\left(\mathcal{M}_{it}Y_{it} - I_{it} - N_{it}^{1+\gamma}\right)^{-\sigma}\right]}{1 - \Xi_i(K_{it}, I_{it})} = \beta\mathbb{E}_{it}\Bigg[\left(\mathcal{M}_{it}Y_{it} - I_{it} - N_{it}^{1+\gamma}\right)^{-\sigma} \left(\mathcal{M}_{i,t+1}(1-\theta)\frac{Y_{it+1}}{K_{it+1}} + \frac{1-\delta-\Xi_k(K_{it+1}, I_{it+1})}{1-\Xi_i(K_{it+1}, I_{it+1})}\right)\Bigg]. \tag{4.2}$$

Note that to infer the terms of trade, producers now have to form beliefs about their trading partners' productivity level and capital stock. Even after the confidence shock has become effectively common knowledge, the output level could still be higher than its steady-state level simply due to previously accumulated additional capital stock. The confidence shock can therefore have a more persistent impact on the economy relative to the static model.

If one interprets $\mathcal{M}_{it}$ as a TFP shock, then these two first-order conditions are isomorphic to those in neoclassical stochastic growth models, except that the perceived TFP shocks are only in expectation. The confidence shock induce systemic shifts in beliefs about $\mathcal{M}_{it}$. The economy behaves as if there are changes in aggregate TFP, but these changes never actually take place. This property will prove crucial when accounting for the identified impulse responses to the main business cycle shock.

4.2 Dynamic Beauty Contest

In this subsection, we formulate the dynamic model as a game. The first-order conditions (4.1) and (4.2) effectively summarize producers' decisions, and the log-linearized system of these two equations can be expressed as

$$\begin{bmatrix} y_{it} \\ k_{it+1} \end{bmatrix} = \begin{bmatrix} \varphi \\ 0 \end{bmatrix} a_i + \mathbb{E}_{it} \underbrace{\begin{bmatrix} 0 & \frac{(1+\gamma)(1-\omega)L}{1+\gamma-\omega\theta} \\ \frac{\beta(1-\theta\varkappa_2)L^{-1}-\varkappa_1(1-\omega)}{\varkappa_1\varkappa_3} & \frac{(\varkappa_1+\beta\Xi\varkappa_2)(\beta L+L^{-1})}{\beta\varkappa_3} \end{bmatrix}}_{\mathbf{\Psi}(L)} \begin{bmatrix} y_{it} \\ k_{it+1} \end{bmatrix} + \mathbb{E}_{it} \underbrace{\begin{bmatrix} \frac{\theta(1-\omega)}{1+\gamma-\omega\theta} \\ \frac{(1-\omega)\left(\varkappa_1(1-L^{-1})+\beta(1-\theta)\varkappa_2 L^{-1}\right)}{\varkappa_1\varkappa_3} \end{bmatrix}}_{\mathbf{\Gamma}(L)} y_{jt} \tag{4.3}$$

where $\{\varkappa_1, \varkappa_2, \varkappa_3\}$ are composites of the model parameters

$$\varkappa_1 = \frac{\beta(1-\theta)}{\delta\beta+1-\beta}, \quad \varkappa_2 = 1 - \delta\varkappa_1 - \frac{\theta}{1+\gamma}, \qquad \varkappa_3 \equiv \frac{\beta(1-\theta)\varkappa_2}{\varkappa_1} + \left(1+\frac{1}{\beta}\right)\varkappa_1 + (1+\beta)\Xi\varkappa_2.$$

This system is effectively a multivariate beauty-contest game. Particularly, the polynomial matrix $\mathbf{\Psi}(L)$ represents the partial equilibrium consideration, capturing agents' response holding other islands' outputs fixed. The matrix $\mathbf{\Gamma}(L)$ instead represents the general equilibrium consideration which describes how agents should respond to other islands' output changes. Compared with the standard beauty-contest model in the style of Morris and Shin (2002), several complications arise: first, the optimal response involves both forward-looking and backward-looking decisions due to the presence of endogenous capital accumulation. Secondly, the coordination motive can no longer be summarized by a single parameter. Producers have to think about others' outputs in the current period and in the future, which in turn leads to inferences about others' capital stock and so on. The relevant types of higher-order expectations become much more involved and are intertemporal in nature, as discussed in Angeletos and Huo (2021).

Even with these additional complications, the equilibrium dynamics still admit a tractable finite-state representation as shown in the following proposition.

Proposition 4.1. *The aggregate output and capital dynamics are given by*

$$y_t = \mu_y \frac{1 - r_y L}{(1-\vartheta_1 L)(1-\vartheta_2 L)}\, \eta_t$$
$$k_{t+1} = \mu_k \frac{1 - r_k L}{(1-\vartheta_1 L)(1-\vartheta_2 L)(1-\lambda L)}\, \eta_t$$

where ϑ_1 and ϑ_2 are the reciprocals of the outside roots of the determinant of the following polynomial matrix

$$\mathbf{T}(L) = \mathbf{I} - \mathbf{\Psi}(L) - \frac{L(1-\lambda\rho)(\rho-\lambda)}{\rho(1-\lambda L)(L-\lambda)} \begin{bmatrix} \mathbf{\Gamma}(L) & \mathbf{0} \end{bmatrix}, \tag{4.4}$$

and $\{\mu_y, \mu_k, r_y, r_k\}$ specified in Appendix A.8 are scalars that depend on ϑ_1 and ϑ_2.

The dynamics of the equilibrium outcomes are jointly shaped by the PE consideration, the GE consideration, and the informational friction. The precise way of how these factors interact with each other is encoded in the determinant of $\mathbf{T}(L)$. Notably, it is the product of the term related to informational frictions and the term related to the GE consideration that enters the equation, which correspond to the elements that make higher-order expectations relevant. When informational frictions vanish, the persistence of first-order expectation, λ, approaches to 0, and $\mathbf{T}(L)$ reduces to $\mathbf{I} - \left(\mathbf{\Psi}(L) + \begin{bmatrix} \mathbf{\Gamma}(L) & \mathbf{0} \end{bmatrix}\right)$. In this case, only the total sum of PE and GE consideration matters, while its decomposition becomes irrelevant. On the other hand, when there is no need for trade ($\omega = 1$), the GE consideration vanishes and there is no need to forecast others' productivity. As a result, the degree of informational friction becomes irrelevant.

4.3 Parameterization

To quantify the model, we calibrate part of the parameters that are common in the business cycle literature and estimate the remaining parameters that are closely related to the informational frictions to match the forecast errors in the survey data on expectations.

The model period is a quarter. We set the discount rate β to 0.99, which implies that the annual rate of return is 4%. We set the risk aversion σ to 1. The three parameters $\{\gamma, \theta, \omega\}$ are directly related to the degree of strategic complementarity. The Frisch elasticity is set to be $\frac{1}{\gamma} = 0.67$, which lies between the common micro and macro estimates. The output elasticity to labor θ is set to be $\theta = 0.66$, mapping to the average labor share in the data. We set ω to be 0.45, so that $1 - \omega$ equals to the average intermediate goods expenditure share across sectors.[1] The magnitude of the capital adjustment costs Ξ is related to the dynamic property of the model. We calibrate it so that the model implied volatility of investment relative to that of output is the same as its counterpart in the HP-filtered U.S. data.

Table 1: Calibrated Parameters

Parameter	Description	Value
β	Discount rate	0.99
σ	Risk aversion	1.00
ω	Home bias	0.45
$\frac{1}{\gamma}$	Frisch elasticity	0.67
θ	Labor share	0.66
σ_a	Std of island specific productivity	0.25
Ξ	capital adjustment cost	35

Turning to the parameters related to the informational friction and the confidence shock process. There are four such parameters to be determined, $\{\sigma_a, \sigma_u, \sigma_\eta, \rho\}$. Among them, σ_a and σ_u have similar implications on the aggregate outcomes. We exogenously set $\sigma_a = 0.25$, and estimate the rest of the three parameters.[2] As discussed extensively in Section 3, the variations of aggregate outcomes results from agents' forecast errors and they co-move with each other. To discipline the model, we therefore estimate the remaining three parameters so that the properties of the forecast errors resemble those in the data.

We rely on the Survey of Professional Forecasters (SPF) as our main dataset for estimation. We use the forecast error of the civilian unemployment rate as our model counterpart of output forecast error. The unemployment rate is the main business cycle indicator, and it is a stationary variable that does not require any additional detrending treatment. The latter point is particularly important as filtered forecasts tend to use future information that is not available at the time when forecasts are made.[3] In our baseline specification, we match the time series of 3-step-ahead median forecast error $(x_{t+3} - \mathbb{E}_t[x_{t+3}])$. The median forecast instead of the mean is used as the aggregate forecast, which helps alleviate concerns about outliers and/or data-entry errors from driving the results.[4] We rescale the series of the forecast error by the relative volatility between HP-filtered GDP and unemployment rate, so that their magnitudes are comparable. Figure A4 in Appendix B plots the forecast errors used in our estimation.

[1]We use sectors at a level slightly more aggregated than the 2-digit ISIC revision 3, following Huo, Levchenko, and Pandalai-Nayar (2019).

[2]Given that σ_a is not too small for the confidence shock to play a role, the quantitative results do not hinge on the value of σ_a.

[3]For example, the GDP forecasts are in terms of level, which requires additional filtering to obtain a stationary variable comparable to HP-filtered GDP fluctuations.

[4]This could be influential given that there are about 40 forecasters each period.

Table 2: Estimated Parameters

	Prior			Posterior		
	Distribution	Mean	St.d	Mode	Mean	90% HPD
σ_u	Inv Gamma	0.10	0.20	0.200	0.169	[0.130, 0.208]
σ_η	Inv Gamma	0.05	0.20	0.038	0.038	[0.034, 0.043]
ρ	Beta	0.60	0.20	0.953	0.943	[0.895, 0.991]

Table 2 shows the choice of prior distributions, the estimated posterior mode, the posterior mean, and also the 90% highest posterior density (HPD) interval of the posterior distribution. Throughout, we use the parameters at their posterior modes to generate the quantitative results, though they are quite similar to the posterior means.

4.4 Results

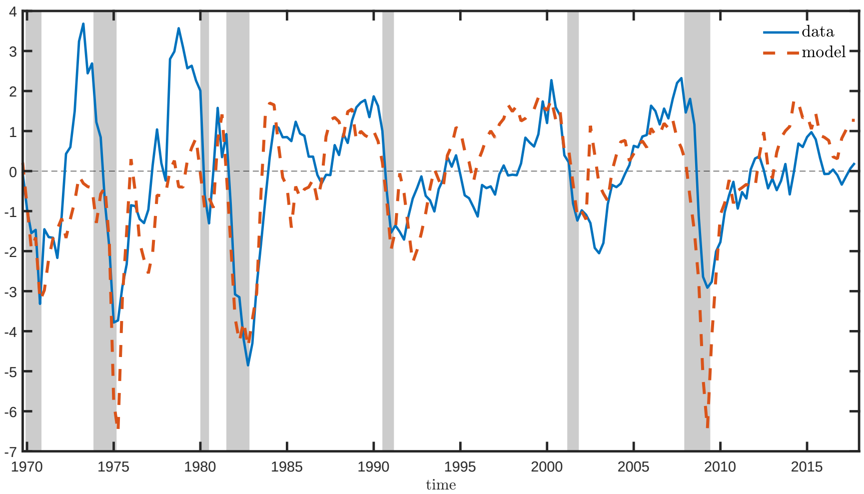
We offer two perspectives when evaluating our model's quantitative performance: the first one is to compare the traditional unconditional business cycle moments with the data, and the second one is to compare with the conditional response to the main business cycle shock identified in Angeletos, Collard, and Dellas (2020).

Baseline model with confidence shock. Figure 4 plots the HP-filtered GDP series and the output series in the model driven by the estimated confidence shock. Though the shocks are estimated to match the forecast error, the output series in the model tracks its data counterpart reasonably well and capture important business cycle fluctuations. As illustrated in Section 3, in response to a positive confidence shock, agents in the model economy overestimate their trading partners' productivities, but underestimate the confidence shock and the aggregate output. As learning takes time, such underestimation is persistent. In the survey data, the forecasters are on average under-react to news (Coibion and Gorodnichenko, 2015), and forecast errors are persistent and positively correlated with the output. These properties help explain why our model is able to match the output process.

Table 3 presents the unconditional business cycle moments for the main variables. The standard deviations of output and investment are about two thirds of their data counterparts. The relative volatility between consumption and output is smaller than that in the data, while the relative volatility of labor is higher than that in the data. Note that in our baseline model, the aggregate TFP is completely muted. Without TFP movements, labor has to be more responsive to generate the same amount of output changes. Producers in the model also overestimate the value of their products, leaving less consumption than desired ex post. In Subsection 4.5, we show that the fit of the unconditional moments can be improved if the model is extended to allow endogenous TFP movements.

The aggregate variables display sizable persistence, which is driven by the fact that forecast errors are fairly persistent. It is useful to note that in DSGE models with noisy observations of fundamen-

Figure 4: Model v.s. Data: Output Fluctuations



Notes: The blue solid line corresponds to the HP-filtered GDP series in the data. The red dashed line corresponds to our model implied aggregate output when simulating with estimated confidence shocks.

tals (Angeletos and Huo, 2021), informational frictions brings additional sluggishness to the long-lasting movement of fundamentals themselves. In our model, there is no such persistent aggregate fundamental in the first place, and all the persistent effects originate from the slow process that the confidence shock approaches to common knowledge.

Table 3: Business Cycle Statistics

	Standard deviation		Corr w/ output		Auto-correlation	
	data	model	data	model	data	model
Y	1.48	0.94	1.00	1.00	0.85	0.63
C	5.63	3.21	0.93	0.99	0.86	0.62
I	0.81	0.19	0.82	0.79	0.88	0.81
N	1.43	2.80	0.84	0.98	0.86	0.62
$\mathcal{M}$	3.00	6.10	-0.72	-0.97	0.84	0.62

Notes: The sample for main aggregate variables is from 1955 q1 to 2018 q4. All variables are logged and HP-filtered. $\mathcal{M}$ corresponds to the measured labor wedge, which equals the negative of $\overline{\log \mathcal{M}_{it}}$.

Though the confidence shock does not shift the supply side of the economy, all the main variables display the right comovements with output as in standard RBC models. This is because agents rationally treat the effects of the confidence shock as if there are aggregate TFP changes. In addition, such shifts of beliefs manifest themselves as a counter-cyclical labor wedge, which is consistent with Chari, Kehoe, and McGrattan (2007) who emphasis the importance of labor wedge in accounting for business cycle fluctuations.

Comparison with MBC shock. Angeletos, Collard, and Dellas (2020) identify a single shock that accounts for most of the short-run fluctuations of main real variables, which is labeled as the main business cycle shock. This shock is almost orthogonal to measured Solow residuals and inflation, which suggests that models with "demand-driven cycles without a strict reliance on nominal rigidity hold promise."

Our model provides one particular micro-foundation to rationalize this empirical pattern. Importantly, agents are all rational and the confidence shock to beliefs is disciplined by the survey data on expectations. In our baseline specification, the aggregate TFP remains unchanged, which is consistent with the response to the MBC shock. The solid blue lines in Figure 5a to Figure 5c reproduce the responses of output, investment, and consumption to the main business cycle shock identified by maximizing the unemployment volatility. The dashed red lines correspond to the responses of variables to the confidence shock in our model economy. The structural responses in our model resemble those from the empirical VAR, though we do not target these moments directly. The responses to the MBC shock ask for a propagation mechanism that looks like the responses to an aggregate TFP shock without actual changes in TFP, which is exactly how our model economy works. The confidence shock is perceived as a TFP shock by individual agents, but it manifests itself as a demand shock to econometricians at the aggregate level.

Different from models with perfect information, our model economy speaks jointly to the aggregate variables and their forecasts. To examine the properties of expectations in the data,[5] we estimate the impulse response of the forecast error to the MBC shock via an instrumental-variables ARMA(p, q) representation

$$\text{error}_t = \sum_{k=1}^{p} \beta_k \, \text{error}_{t-k}^{\text{IV}} + \sum_{k=0}^{q} \alpha_k \varepsilon_{t-k} + \text{residuals}. \tag{4.5}$$

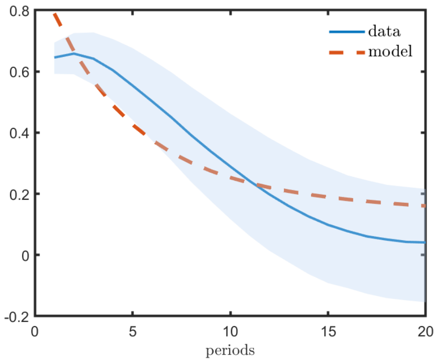
where error_t stands for the rescaled 3-step-ahead forecast error of unemployment rate, and ε_t is the main business cycle shock.[6]

Figure 5d plots the estimated IRF of the forecast error and the theoretical counterpart implied by the model. The two IRFs both are positive on impact and they share a similar magnitude and persistence. These patterns imply that agents under-react to the shocks, and the confusion is long-lasting, consistent with our model's prediction. Despite the overall similarity, the response of forecast error in the data displays a sign-switching pattern, suggesting that the forecasts also feature a delayed over-reaction as documented in Angeletos, Huo, and Sastry (2021). In Subsection 4.5, we show that by introducing over-extrapolation of the shock process helps account for this pattern, and the main implications of the baseline model remain valid.

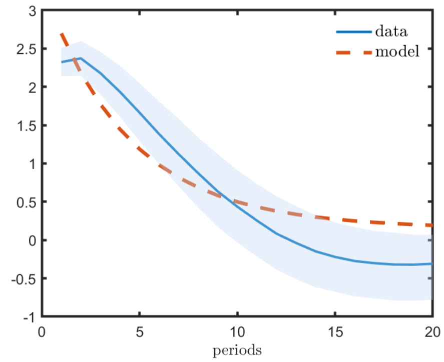
The response of the forecast error clearly rules out models with full information rational expec-

[5] In the VAR system specified by Angeletos, Collard, and Dellas (2020), the forecasts of variables are not included.

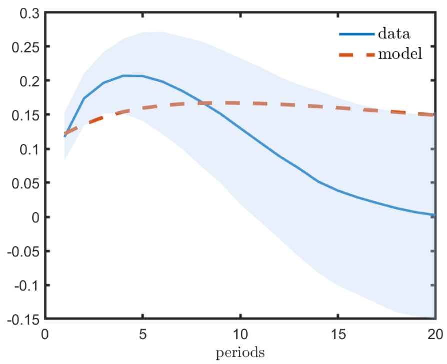
[6] The estimated IRF is similar to the result of a local projection method of Jordà (2005) .



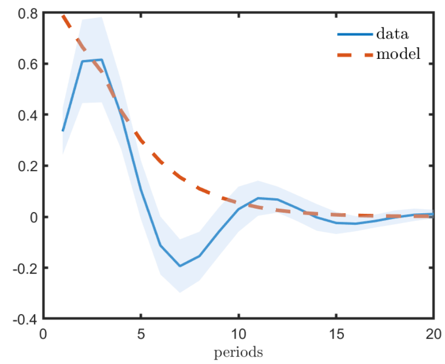
(a) output



(b) investment



(c) consumption



(d) output forecast error

Figure 5: Comparison with Impulse Responses to Main Business Cycle Shock

Notes: The blue lines correspond to the responses of aggregate variables in response to the MBC shock that maximizes unemployment volatility in Angeletos, Collard, and Dellas (2020). The red dashed lines correspond to responses to the confidence shock in our model economy.

tations (FIRE) in accounting for the main business cycle shock. The forecast errors in these types of models disappear immediately after a shock realizes, which is inconsistent with the persistent nature of the forecast error. The response of aggregate forecast error favors models where agents under-react to aggregate shocks (at least on average), which are potentially consistent with models with sticky information (Mankiw and Reis, 2002), rational inattention (Sims, 2003), dispersed information (Lucas, 1972), cognitive discounting (Gabaix, 2017), level-k thinking (Farhi and Werning, 2017), and so on. The properties of the forecast errors at individual level are important in distinguishing alternative theories (Bordalo, Gennaioli, Ma, and Shleifer, 2020) but may not have direct consequence for the aggregate variables, as discussed in Angeletos, Huo, and Sastry (2021).

Taking stock, the confidence shock in our baseline model reproduces the key features of the MBC shock. Furthermore, our theory simultaneously matches important empirical patterns of forecasts. The response of forecast errors provides guidance on what type of theories are most relevant in the business cycle context, and it calls for a deviation from FIRE models.

Role of higher-order expectations. The quantitative results so far has shown that the confidence shock to expectations can generate a significant amount of aggregate fluctuations. But what are the forces that drive the quantitative bite? The shifts of the confusion about their trading partners' productivity (the first-order expectations), or the confusion about the forecast of others (the higher-order expectations)?

To answer this question, we construct a particular auxiliary economy where agents ignore any higher-order expectations about each others' productivity. That is, only $\mathbb{E}_{it}[a_{jt}]$ is taken into account, but terms such as $\mathbb{E}_{it}[\mathbb{E}_{jt}[a_i]]$, $\mathbb{E}_{it}[\mathbb{E}_{jt}[\mathbb{E}_{it}a_{jt}]]]$ and so on will be ignored. In this auxiliary economy, the counterpart of the best-response system (4.3) becomes[7]

$$\begin{bmatrix} y_{it} \\ k_{it+1} \end{bmatrix} = \begin{bmatrix} \varphi \\ 0 \end{bmatrix} a_i + \varphi \mathbf{\Gamma}(L)_{+} \mathbb{E}_{it}[a_{jt}] + \mathbf{\Psi}(L) \begin{bmatrix} y_{it} \\ k_{it+1} \end{bmatrix}.$$

Note that this response functions only depends on agents' own fundamental, the first-order expectation about others' fundamental, and their own past or future choices (the PE part of the original system). Therefore, it isolates the impact of the confidence shock only through the first-order expectations about others' productivities.

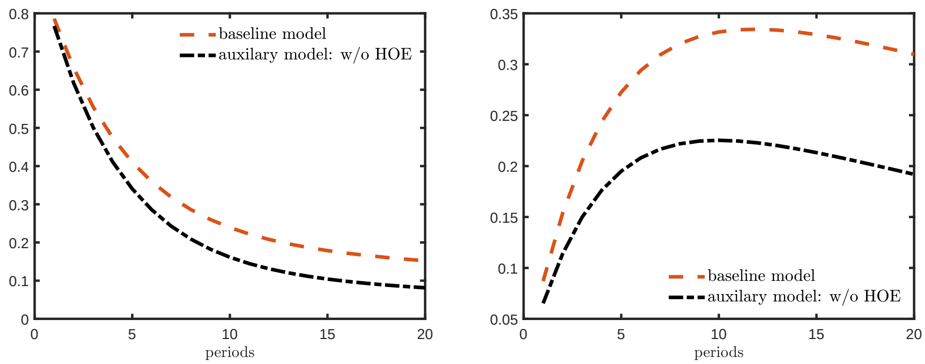
Figure 6 compares the baseline model (red dashed lines) with the auxiliary model (black broken lines). With only first-order expectations, the responses of output and capital are dampened and less persistent. This result confirms the intuition developed in subsection 2.2 that higher-order expectations co-move with first-order expectations but are more persistent. The response of output on impact is not significantly affected when ignoring the higher-order terms, but the difference from the baseline model is magnified later on. This is mainly due to the fact that the implied degree of complementarity $\alpha = \frac{\theta(1-\omega)}{1+\gamma-\omega\theta}$ that shows up in the first row of $\mathbf{\Gamma}(L)$ is about 0.16 under our calibration, which limits the role of the GE consideration, and that the income effect is absent with the GHH preference. In contrast, the response of capital is significantly weakened, showing that higher-order expectations are important in shaping the slow-moving factor's dynamics. The more transitory nature of the capital response without higher-order expectations contributes to the faster convergence of output to the steady state.

4.5 Alternative Specifications

In this subsection, alternative modeling and estimation strategies are explored. Our main results survive in these exercises.

Persistent trading partner. In our baseline specification, we impose that the matching follows an i.i.d. process, that is, the productivity of an island's trading partner is uncorrelated over time. This assumption helps simplify the analytical results, but one may have the concern about whether a more

[7]The term $\mathbf{\Gamma}(L)_{+}$ is to apply the annihilation operator on the polynomial matrix, which equal to $\begin{bmatrix} \frac{\theta(1-\omega)}{1+\gamma-\omega\theta} & \frac{1-\omega}{\varkappa_3} \end{bmatrix}$.



(a) output (b) capital

Figure 6: Role of Higher-Order Expectations

Notes: The red-dashed lines correspond to the baseline model. The black broken lines correspond to the maxillary model where only first-order expectations about trading partners' productivity are taken into account.

realistic persistent matching quality eliminates the role of the confidence shock as agents expect their trading partners to be similar to those in the past.

To address this concern, we explore the following alternative matching process. Recall that the index of island i's trading partner in period t is $m(i,t)$, and assume that the productivity of $m(i,t)$ follows an AR(1) process

$$a_{m(i,t)} = \rho_a a_{m(i,t-1)} + \nu_{it}, \tag{4.6}$$

where ρ_a determines the persistence and ν_{it} is the innovation to the matching quality. With $\rho_a = 0$, it reduces to the baseline i.i.d process.

We set $\rho_a = 0.7$, which lies in the middle of various empirical estimates of firm-level productivity. We impose $\frac{\sigma_\nu^2}{1-\zeta^2} = \sigma_a^2$, that is, the dispersion of the productivity distribution remains the same as the baseline model but the matching process is more stable from an individual island's perspective. The responses of aggregate variables are presented in Figure A2 in Appendix B.1. Compared with the baseline model, adding persistent matching only modestly lower the persistence of the aggregate variables. Therefore, our results do not hinge on the particular assumption on the matching process.

Endogenous TFP process. A salient feature of our baseline model is that there is no *exogenous* aggregate TFP changes. As a result, only the primary inputs labor and capital are responsible for output movements. In this part, we explore the extension where the measured TFP varies *endogenously* with the aggregate demand. Particularly, we follow the work by Bai, Ríos-Rull, and Storesletten (2011) and Michaillat and Saez (2015) where goods market frictions enable output to be demand determined without nominal rigidity and the aggregate productivity is endogenous. The basic idea is that shoppers have to search for goods before they can consume them, and goods have to be found before they can be sold. Matching frictions in the goods market may prevent produced goods from being actually sold. The probability that goods can be sold is determined by the amount of search effort exerted by shoppers. As a result, the search effort creates a wedge between potential output and actual output, which corresponds to the measured Solow residuals. Crucially, the amount of search effort exerted by shoppers is correlated with the level of production, which results in shifts of the Solow residuals over business cycles.

We describe the details of the model in Appendix B.2. In a nutshell, the measured Solow residual and actual output in this economy can be expressed as

$$z_t = \frac{\kappa}{1-\kappa} y_t, \qquad \text{output} = z_t + y_t,$$

where κ is the goods market matching elasticity and y_t now becomes the potential output. We set $\kappa = 0.3$ so that the roughly one third of the output variation is due to measured Solow residuals and re-estimate the model to match the forecast error. Table A1 compares the unconditional business cycle moments with the baseline model. Relative to the baseline model, the volatility of consumption is higher and the volatility of labor is lower. As output is driven by both primary inputs and endogenous TFP, it no longer requires labor to do all the heavy lifting to match the output volatility. The endogenous TFP also provides an additional boost for the consumption, bringing it closer to the data.

We do not view the goods market friction as the only reason for endogenous TFP changes, which could also come from variable capital utilization (Basu, 1996), labor hoarding (Burnside, Eichenbaum, and Rebelo, 1993), and many other sources. This extension simply shows that once augmented with elements that generates TFP comovement, the confidence shock that acts as a demand shock is capable of generating business cycle moments similar to those in canonical RBC-type models.

Over-extrapolation. As shown in Figure 5d, the response of forecast error to the MBC shock displays an overshooting pattern. Our baseline model shares this sign-switching pattern qualitatively (see Lemma 2), but is unable to match it quantitatively. To account for this pattern, we follow the approach in Angeletos, Huo, and Sastry (2021) and assume that agents perceive that the confidence shock follows an AR(1) process with persistence $\hat{\rho}$ where $\hat{\rho} > \rho$. This over-extrapolation formulation is consistent with the empirical evidence documented in Greenwood and Shleifer (2014) and Gennaioli, Ma, and Shleifer (2016).

We modify our baseline parameterization by setting $\hat{\rho} = 0.97$ and $\rho = 0.84$. Figure A3 shows that IRF of the forecast error of output now displays a significant over-shooting pattern. Under this specification, after a confidence shock, the aggregate forecasts under-react due to the presence of noises. Over time, as learning takes place, the effect of over-extrapolation starts to dominate, resulting in a delayed over-reaction of the forecasts. The response of output inherits the properties of the forecast error, which reaches zero faster with a slight overshooting pattern. The over-extrapolation, however, does not change the nature of the confidence shock, which induce persistent fluctuations of main aggregate variables that co-move with each other.

Matching forecasts with different horizons. In our model, the magnitude and the persistence of aggregate variables in response to the confidence shock are positively related to those of the forecast errors. In the SPF, the forecast errors become more persistent as the forecasting horizon increases. In our baseline estimation, we have followed the convention in the literature to match the three-step ahead forecast error. We also experiment with alternative specification by matching the average of one-step ahead to three-step ahead forecast errors. Figure A2 in Appendix B.1 plots the IRFs of output and forecast error. As expected, the output response is smaller and less persistent with shorter forecasting horizons, but quantitatively similar to our baseline results.

5. Conclusion

We formalize the idea that waves of optimism and pessimism drive business cycles in a rational expectations framework. With trade linkages and informational frictions, the equilibrium outcome is a weighted average of first-order and higher-order expectation errors about the confidence shock. With dynamic information, such higher-order expectations become increasingly complex. We prove that in equilibrium, such complexity disappears and the aggregate outcome follows a simple AR(1) process. The persistence and volatility of output are endogenously determined, both of which are increasing in the informational frictions and the general equilibrium consideration.

In our quantitative evaluation, we discipline the model by survey data on expectations. In terms of unconditional moments, the confidence shock can account for a large fraction of aggregate fluctuations. The performance of our model is comparable to a textbook RBC model, while the as-if TFP shock in our model is endogenously determined by forecast errors. Our model can also be viewed as a micro foundation for the MBC shock identified in Angeletos, Collard, and Dellas (2020). The responses of aggregate variables to the confidence shock capture the key features of the response of the MBC shock. In addition, the responses of the aggregate forecasts to the confidence shock and to the MBC shock are also consistent with each other.

Left outside this paper was a more comprehensive investigation of the trading structure. Recent research has incorporated production networks with informational frictions (Chahrour, Nimark, and Pitschner, 2021; La'O and Tahbaz-Salehi, 2020; Bui, Huo, Levchenko, and Pandalai-Nayar, 2022), but the implications of the interaction between network games and shocks to beliefs remain largely unexplored. Another direction we leave for future research is to connect the inferred confidence shocks with more directly measured shocks to beliefs such as fake news (Bybee, Kelly, Manela, and Xiu, 2021). The framework and the technique developed in this paper could be extended to these studies.

References

ACHARYA, S., J. BENHABIB, AND Z. HUO (2021): "The anatomy of sentiment-driven fluctuations," *Journal of Economic Theory*, 195, 105280.

AMADOR, M., AND P.-O. WEILL (2012): "Learning from private and public observations of others actions," *Journal of Economic Theory*, 147(3), 910–940.

ANGELETOS, G.-M., F. COLLARD, AND H. DELLAS (2018): "Quantifying Confidence," *Econometrica*, 86(5), 1689–1726.

——— (2020): "Business-cycle anatomy," *American Economic Review*, 110(10), 3030–70.

ANGELETOS, G.-M., AND Z. HUO (2021): "Myopia and Anchoring," *American Economic Review*, 111(4), 1166–1200.

ANGELETOS, G.-M., Z. HUO, AND K. SASTRY (2021): "Imperfect Macroeconomic Expectations: Evidence and Theory," in *NBER Macroeconomics Annual 2020*, vol. 35, chap. 1.

ANGELETOS, G.-M., AND J. LA'O (2010): "Noisy Business Cycles," in *NBER Macroeconomics Annual 2009, Volume 24*, pp. 319–378. University of Chicago Press.

——— (2013): "Sentiments," *Econometrica*, 81(2), 739–779.

ANGELETOS, G.-M., AND C. LIAN (2018): "Forward guidance without common knowledge," *American Economic Review*, 108(9), 2477–2512.

BAI, Y., J.-V. RÍOS-RULL, AND K. STORESLETTEN (2011): "Demand Shocks as Productivity Shocks," manuscript, Federal Reserve Bank of Minneapolis.

BARSKY, R. B., AND E. R. SIMS (2012): "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence," *American Economic Review*, 102(4), 1343–77.

BASU, S. (1996): "Procyclical Productivity: Increasing Returns or Cyclical Utilization?," *Quarterly Journal of Economics*, 111(3), 719–751.

BENHABIB, J., P. WANG, AND Y. WEN (2015): "Sentiments and Aggregate Demand Fluctuations," *Econometrica*, 83, 549–585.

BERGEMANN, D., T. HEUMANN, AND S. MORRIS (2015): "Information and volatility," *Journal of Economic Theory*, 158, 427–465.

BERGEMANN, D., AND S. MORRIS (2013): "Robust Predictions in Games With Incomplete Information," *Econometrica*, 81(4), 1251–1308.

Bhandari, A., J. Borovička, and P. Ho (2019): "Survey Data and Subjective Beliefs in Business Cycle Models," University of Minnesota, New York University and Federal Reserve Bank of Richmond.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110(9), 2748–82.

Bui, H., Z. Huo, A. A. Levchenko, and N. Pandalai-Nayar (2022): "Information Frictions and News Media in Global Value Chains," Discussion paper, National Bureau of Economic Research.

Burnside, C., M. Eichenbaum, and S. Rebelo (1993): "Labor hoarding and the business cycle," *Journal of Political Economy*, 101(2), 245–273.

Bybee, L., B. T. Kelly, A. Manela, and D. Xiu (2021): "Business news and business cycles," Discussion paper, National Bureau of Economic Research.

Chahrour, R., and K. Jurado (2018): "News or noise? The missing link," *American Economic Review*, 108(7), 1702–36.

Chahrour, R., K. Nimark, and S. Pitschner (2021): "Sectoral media focus and aggregate fluctuations," *American Economic Review*, 111(12), 3872–3922.

Chahrour, R., and R. Ulbricht (2022): "Robust predictions for DSGE models with incomplete information," Forthcoming, *American Economic Journal: Macroeconomics*.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2007): "Business Cycle Accounting," *Econometrica*, 75, 781–836.

Coibion, O., and Y. Gorodnichenko (2015): "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts," *American Economic Review*, 105(8), 2644–78.

Coibion, O., Y. Gorodnichenko, S. Kumar, and J. Ryngaert (2021): "Do You Know that I Know that You Know…? Higher-Order Beliefs in Survey Data," *The Quarterly Journal of Economics*, 136(3), 1387–1446.

Farhi, E., and I. Werning (2017): "Monetary Policy, Bounded Rationality, and Incomplete Markets," *NBER Working Paper No. 23281*.

Fuhrer, J. C. (2018): "Intrinsic expectations persistence: evidence from professional and household survey expectations," Working Paper.

Gabaix, X. (2017): "A Behavioral New Keynesian Model," *Harvard mimeo*.

Gennaioli, N., Y. Ma, and A. Shleifer (2016): "Expectations and investment," *NBER Macroeconomics Annual*, 30(1), 379–431.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.

Greenwood, R., and A. Shleifer (2014): "Expectations of returns and expected returns," *The Review of Financial Studies*, 27(3), 714–746.

Hébert, B. M., and J. La'O (2020): "Information Acquisition, Efficiency, and Non-Fundamental Volatility," Discussion paper, National Bureau of Economic Research.

Huo, Z., A. A. Levchenko, and N. Pandalai-Nayar (2019): "International comovement in the global production network," Discussion paper, National Bureau of Economic Research.

Huo, Z., and M. Pedroni (2020): "A Single-Judge Solution to Beauty Contests," *American Economic Review*, 110(2), 526–68.

Huo, Z., and N. Takayama (2021): "Rational Expectations Models with Higher Order Beliefs," Manuscript, Yale University.

Ilut, C. L., and M. Schneider (2014): "Ambiguous Business Cycles," *American Economic Review*, 104(8), 2368–99.

Jaimovich, N., and S. Rebelo (2009): "Can News about the Future Drive the Business Cycle?," *American Economic Review*, 99(4), 1097–1118.

Jordà, Ò. (2005): "Estimation and inference of impulse responses by local projections," *American economic review*, 95(1), 161–182.

Kohlhas, A., and A. Walther (2021): "Asymmetric attention," Forthcoming, *American Economic Review*.

La'O, J., and A. Tahbaz-Salehi (2020): "Optimal monetary policy in production networks," Discussion paper, National Bureau of Economic Research.

Lorenzoni, G. (2009): "A Theory of Demand Shocks," *American Economic Review*, 99(5), 2050–84.

Lucas, R. E. J. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4(2), 103–124.

Maćkowiak, B., and M. Wiederholt (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3).

Mankiw, N. G., and R. Reis (2002): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, pp. 1295–1328.

Michaillat, P., and E. Saez (2015): "Aggregate demand, idle time, and unemployment," *The Quarterly Journal of Economics*, 130(2), 507–569.

Morris, S., and H. S. Shin (2002): “Social Value of Public Information,” *American Economic Review*, 92(5), 1521–1534.

Nimark, K. (2008): “Monetary policy with signal extraction from the bond market,” *Journal of Monetary Economics*, 55(8), 1389–1400.

Nimark, K. P. (2014): “Man-bites-dog business cycles,” *American Economic Review*, 104(8), 2320–67.

Sims, C. A. (2003): “Implications of Rational Inattention,” *Journal of Monetary Economics*, 50(3), 665–690.

Taub, B. (1989): “Aggregate fluctuations as an information transmission mechanism,” *Journal of Economic Dynamics and Control*, 13(1), 113–150.

Trejos, A., and R. Wright (1995): “Search, Bargaining, Money, and Prices,” *Journal of Political Economy*, 103(1), 118–141.

Venkateswaran, V. (2014): “Heterogeneous Information and Labor Market Fluctuations,” Working paper.

Woodford, M. (2003): “Imperfect Common Knowledge and the Effects of Monetary Policy,” *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, p. 25.

Online Appendix

A. Proof of Theorems and Propositions

A.1 Proof of Proposition 2.2

Let j denote $m(i,t)$. With the optimal output rule (2.8), successive substitution leads to

$$\begin{aligned}
y_{it} &= \varphi a_i + \alpha \mathbb{E}_{it}\left[y_{jt}\right] \\
&= \varphi a_i + \alpha \mathbb{E}_{it}\left[\varphi a_j + \alpha \mathbb{E}_{jt}\left[y_{it}\right]\right] \\
&= \varphi a_i + \varphi\alpha \mathbb{E}_{it}\left[a_j\right] + \alpha^2 \mathbb{E}_{it}\mathbb{E}_{jt}[y_{it}] \\
&= \varphi a_i + \varphi\alpha \mathbb{E}_{it}\left[a_j\right] + \alpha^2 \mathbb{E}_{it}\mathbb{E}_{jt}[\varphi a_i + \alpha \mathbb{E}_{it}\left[y_{jt}\right]] \\
&= \varphi a_i + \varphi\alpha^2 \mathbb{E}_{it}\mathbb{E}_{jt}[a_i] + \varphi\alpha \mathbb{E}_{it}\left[a_j\right] + \alpha^3 \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}[y_{jt}] \\
&= \varphi a_i + \varphi\alpha^2 \mathbb{E}_{it}\mathbb{E}_{jt}[a_i] + \varphi\alpha \mathbb{E}_{it}\left[a_j\right] + \varphi\alpha^3 \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}[a_j] + \alpha^4 \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}\mathbb{E}_{jt}[y_{it}] \\
&\ \ \vdots \\
&= \varphi \sum_{k=0}^{\infty} \alpha^{2k} \mathbb{E}_{it}^{2k}[a_i] + \varphi \sum_{k=0}^{\infty} \alpha^{2k+1} \mathbb{E}_{it}^{2k+1}[a_j].
\end{aligned}$$

Given that $\alpha \in (0,1)$ and the modulus of the expectation is bounded from above, the summation in the last line is well defined. The expectation operator $\mathbb{E}_{it}^k[\cdot]$ stands for higher order beliefs and is defined as

$$\begin{aligned}
&\mathbb{E}_{it}^0[a_i] = a_i \\
&\mathbb{E}_{it}^1[a_j] = \mathbb{E}_{it}[a_j] \\
&\mathbb{E}_{it}^k[a_i] = \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}^{k-2}[a_i], \text{ for } k = 2,4,6\ldots \\
&\mathbb{E}_{it}^k[a_j] = \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}^{k-2}[a_j], \text{ for } k = 3,5,7\ldots
\end{aligned}$$

We can derive $\mathbb{E}_{it}^k[a_i]$ or $\mathbb{E}_{it}^k[a_j]$ in the following recursive way

$$\begin{aligned}
&\mathbb{E}_{it}[a_j] = x_{it}^1 - \mathbb{E}_{it}[\xi_t] \\
&\mathbb{E}_{jt}^2[a_i] = \mathbb{E}_{it}[x_{jt}^1 - E_{jt}[\xi_t]] = a_i + \mathbb{E}_{it}[\xi_t] - \mathbb{E}_{it}\mathbb{E}_{jt}[\xi_t] \\
&\mathbb{E}_{it}^3[a_j] = \mathbb{E}_{it}[a_j + \mathbb{E}_{jt}[\xi_t] - \mathbb{E}_{jt}\mathbb{E}_{it}[\xi_t]] = \mathbb{E}_{it}[a_j] + \mathbb{E}_{it}\mathbb{E}_{jt}[\xi_t] - \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}[\xi_t] \\
&\mathbb{E}_{it}^4[a_i] = \mathbb{E}_{it}[\mathbb{E}_{jt}[a_i] + \mathbb{E}_{jt}\mathbb{E}_{it}[\xi_t] - \mathbb{E}_{jt}\mathbb{E}_{it}\mathbb{E}_{jt}[\xi_t]] = \mathbb{E}_{it}\mathbb{E}_{jt}[a_i] + \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}[\xi_t] - \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}\mathbb{E}_{jt}[\xi_t].
\end{aligned}$$

More compactly,

$$\mathbb{E}_{it}^k[a_i] = a_i - \sum_{n=1}^{k}(-1)^n \mathbb{E}_{it}^n[\xi_t], \text{ for } k = 0,2,4,6\ldots \qquad \mathbb{E}_{it}^k[a_j] = x_{it}^1 + \sum_{n=1}^{k}(-1)^n \mathbb{E}_{it}^n[\xi_t], \text{ for } k = 1,3,5,7\ldots$$

The the output in island i is

$$
\begin{aligned}
y_{it} &= \varphi \sum_{k=0}^{\infty} \alpha^{2k} \mathbb{E}_{it}^{2k}[a_i] + \varphi \sum_{k=0}^{\infty} \alpha^{2k+1} \mathbb{E}_{it}^{2k+1}[a_j] \\
&= \frac{\varphi}{1-\alpha_1^2} a_i + \frac{\varphi\alpha}{1-\alpha_1^2} x_{it}^1 - \frac{\varphi}{1+\alpha} \sum_{k=1}^{\infty} \alpha^k \mathbb{E}_{it}^k[\xi_t] \\
&= \frac{\varphi}{1-\alpha_1^2} a_i + \frac{\varphi\alpha}{1-\alpha_1^2} a_j + \frac{\varphi}{1+\alpha} \sum_{k=1}^{\infty} \alpha^k (\xi_t - \mathbb{E}_{it}^k[\xi_t]).
\end{aligned}
$$

In aggregate,

$$
y_t = \frac{\varphi}{1+\alpha} \sum_{k=1}^{\infty} \alpha^k \left(\xi_t - \int \mathbb{E}_{it}^k[\xi_t] \right).
$$

A.2 Proof of Proposition 2.3

The signal process can be written as

$$
\boldsymbol{x}_{it} = \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix} = \begin{bmatrix} \sigma_a & 0 & \frac{\sigma_\eta}{1-\rho L} \\ 0 & \sigma_u & \frac{\sigma_\eta}{1-\rho L} \end{bmatrix} \begin{bmatrix} \widehat{a}_{m(i,t)} \\ \widehat{u}_{it} \\ \widehat{\eta}_t \end{bmatrix} = \mathbf{M}(L) \boldsymbol{s}_{it},
$$

where we have normalized the shock process to be with unit variance. As shown in Huo and Takayama (2021), the fundamental representation that satisfies $\mathbf{B}(L)\mathbf{V}\mathbf{B}(L^{-1})^{-1}) = \mathbf{M}(L)\mathbf{M}'(L^{-1})$ can be written as

$$
\mathbf{B}(L)^{-1} = \frac{1}{1-\lambda L} \begin{bmatrix} 1 - \frac{\tau_a\rho + \lambda\tau_u}{\tau_a+\tau_u} L & \frac{\tau_u(\lambda-\rho)}{\tau_a+\tau_u} L \\ \frac{\tau_a(\lambda-\rho)}{\tau_a+\tau_u} L & 1 - \frac{\tau_u\rho+\lambda\tau_a}{\tau_a+\tau_u} L \end{bmatrix}, \qquad \mathbf{V}^{-1} = \frac{\tau_u\tau_a}{\rho(\tau_a+\tau_u)} \begin{bmatrix} \frac{\tau_u\rho+\lambda\tau_a}{\tau_u} & \lambda-\rho \\ \lambda-\rho & \frac{\tau_a\rho+\lambda\tau_u}{\tau_a} \end{bmatrix}.
$$

where $\tau_a = \frac{\sigma_\eta^2}{\sigma_a^2}$, $\tau_u = \frac{\sigma_\eta^2}{\sigma_u^2}$ and λ is given by

$$
\lambda = \frac{1}{2}\left[\left(\frac{1}{\rho} + \rho + \frac{\tau_u+\tau_a}{\rho} \right) - \sqrt{\left(\frac{1}{\rho} + \rho + \frac{\tau_u+\tau_a}{\rho} \right)^2 - 4} \right].
$$

Denote the equilibrium policy rule as $y_{it} = \chi a_i + h_1(L) x_{it}^1 + h_2(L) x_{it}^2$, and it has to satisfy the best response $y_{it} = \varphi a_i + \gamma \mathbb{E}_{it}[y_{m(i,t)t}]$. To predict $y_{m(i,t)t}$, it is equivalent to forecast

$$
y_{m(i,t)t} = \chi a_{m(i,t)} + h_1(L)\left(a_{m(m(i,t),t)} + \frac{1}{1-\rho L}\eta_t \right) + h_2(L)\left(u_{m(i,t)t} + \frac{1}{1-\rho L}\eta_t \right).
$$

Note that $\mathbb{E}_{it}[a_{m(m(i,t),\tau)}] = a_i$ for $\tau = t$ and $\mathbb{E}_{it}[a_{m(m(i,t),\tau)}] = 0$ for $\tau \neq t$. Also, $\mathbb{E}_{it}[u_{m(i,t)\tau}] = 0$ for all τ. It remains to specify the forecast about $a_{m(i,t)}$ and $\frac{h_1(L)+h_2(L)}{1-\rho L}\eta_t$. By the Wiener-Hopf prediction formula, the optimal forecast about ξ_t and aggregate

$y_t = \frac{h_1(L)+h_2(L)}{1-\rho L}\eta_t$ are given by

$$\mathbb{E}_{it}\left[\xi_t\right] = \frac{\lambda}{\rho}\frac{1}{(1-\lambda L)(L-\lambda)}\begin{bmatrix}\tau_a\left(L-\lambda\frac{1-\rho L}{1-\rho\lambda}\right) & \tau_u\left(L-\lambda\frac{1-\rho L}{1-\rho\lambda}\right)\end{bmatrix}\begin{bmatrix}x_{it}^1\\ x_{it}^2\end{bmatrix},$$

$$\mathbb{E}_{it}\left[\frac{h_1(L)+h_2(L)}{1-\rho L}\eta_t\right] = \frac{1}{1-\lambda L}\begin{bmatrix}\frac{\lambda\tau_a}{\rho(L-\lambda)}\left(L[h_1(L)+h_2(L)]-\lambda[h_1(\lambda)+h_2(\lambda)]\frac{1-\rho L}{1-\rho\lambda}\right)\\ \frac{\lambda\tau_u}{\rho(L-\lambda)}\left(L[h_1(L)+h_2(L)]-\lambda[h_1(\lambda)+h_2(\lambda)]\frac{1-\rho L}{1-\rho\lambda}\right)\end{bmatrix}'\begin{bmatrix}x_{it}^1\\ x_{it}^2\end{bmatrix}.$$

It follows that

$$\mathbb{E}_{it}[a_{m(i,t)}] = x_{it}^1 - \mathbb{E}_{it}[\xi_t] = \frac{1}{1-\lambda L}\begin{bmatrix}\frac{\tau_u\rho+\tau_a\lambda}{\rho(\tau_a+\tau_u)}-\lambda L\\ \frac{\tau_u(\lambda-\rho)}{\rho(\tau_u+\tau_a)}\end{bmatrix}'\begin{bmatrix}x_{it}^1\\ x_{it}^2\end{bmatrix}.$$

Using the equilibrium condition, the following system has to be true

$$\begin{aligned}&\chi a_i + h_1(L)x_{it}^1 + h_2(L)x_{it}^2 = \varphi a_i\\ &+\alpha\chi\frac{1}{1-\lambda L}\begin{bmatrix}\frac{\tau_u\rho+\tau_a\lambda}{\rho(\tau_a+\tau_u)}-\lambda L\\ \frac{\tau_u(\lambda-\rho)}{\rho(\tau_u+\tau_a)}\end{bmatrix}'\begin{bmatrix}x_{it}^1\\ x_{it}^2\end{bmatrix} + \alpha h_1(0)a_i + \alpha\frac{1}{1-\lambda L}\begin{bmatrix}\frac{\lambda\tau_a}{\rho(L-\lambda)}\left(L[h_1(L)+h_2(L)]-\lambda[h_1(\lambda)+h_2(\lambda)]\frac{1-\rho L}{1-\rho\lambda}\right)\\ \frac{\lambda\tau_u}{\rho(L-\lambda)}\left(L[h_1(L)+h_2(L)]-\lambda[h_1(\lambda)+h_2(\lambda)]\frac{1-\rho L}{1-\rho\lambda}\right)\end{bmatrix}'\begin{bmatrix}x_{it}^1\\ x_{it}^2\end{bmatrix}.\end{aligned}$$

Taking χ as given, $h_1(L)$ and $h_2(L)$ have to satisfy

$$C(L)\begin{bmatrix}h_1(L)\\ h_2(L)\end{bmatrix} = d[L, h_1(\lambda)+h_2(\lambda)],$$

where

$$C(L) = \begin{bmatrix}1-\alpha\frac{\lambda\tau_a}{\rho}\frac{L}{(1-\lambda L)(L-\lambda)} & -\alpha\frac{\lambda\tau_a}{\rho}\frac{L}{(1-\lambda L)(L-\lambda)}\\ -\alpha\frac{\lambda\tau_u}{\rho}\frac{L}{(1-\lambda L)(L-\lambda)} & 1-\alpha\frac{\lambda\tau_u}{\rho}\frac{L}{(1-\lambda L)(L-\lambda)}\end{bmatrix}, \quad d(z) = \begin{bmatrix}\chi\alpha\frac{\frac{\tau_u\rho+\tau_a\lambda}{\rho(\tau_a+\tau_u)}-\lambda L}{1-\lambda L} - \alpha\frac{\lambda^2\tau_a}{\rho}\frac{1}{1-\rho\lambda}\frac{1-\rho L}{(1-\lambda L)(L-\lambda)}[h_1(\lambda)+h_2(\lambda)]\\ \chi\alpha\frac{\tau_u(\lambda-\rho)}{\rho(\tau_a+\tau_u)}\frac{1}{1-\lambda L} - \alpha\frac{\lambda^2\tau_u}{\rho}\frac{1}{1-\rho\lambda}\frac{1-\rho L}{(1-\lambda L)(L-\lambda)}[h_1(\lambda)+h_2(\lambda)]\end{bmatrix}.$$

By the Cramer's rule, we have

$$h_1(L) = \frac{\det\begin{bmatrix}d(L) & C_2(L)\end{bmatrix}}{\det C(L)}.$$

The determinant of $C(L)$ is $\det C(z) = \frac{\frac{\lambda}{\vartheta}(z-\vartheta)(1-\vartheta z)}{(1-\lambda z)(z-\lambda)}$ where ϑ is

$$\vartheta = \frac{1}{2}\left[\left(\frac{1}{\rho}+\rho+\frac{(1-\alpha)(\tau_u+\tau_a)}{\rho}\right) - \sqrt{\left(\frac{1}{\rho}+\rho+\frac{(1-\alpha)(\tau_u+\tau_a)}{\rho}\right)^2 - 4}\right].$$

To make sure that $h_1(L)$ does not have poles in the unit circle, $h_1(\lambda)+h_2(\lambda)$ has to satisfy $\det\begin{bmatrix}d(L) & C_2(L)\end{bmatrix}|_{L=\vartheta} = 0$ to remove the pole at ϑ. This leads to

$$h_1(\lambda)+h_2(\lambda) = \frac{\chi(\vartheta-\lambda)(1-\rho\lambda)}{\lambda(\tau_u+\tau_a)}.$$

The policy function $h_1(L)$ and $h_2(L)$ then can be solved in a straightforward way

$$h_1(L) = \frac{\alpha\chi\left(\frac{\rho\tau_u+\vartheta\tau_a}{(\tau_a+\tau_u)\rho} - \vartheta L\right)}{1-\vartheta L}, \qquad h_2(L) = -\frac{\alpha\chi\frac{\tau_u(\rho-\theta)}{\rho(\tau_a+\tau_u)}}{1-\vartheta L}.$$

The aggregate output follows

$$y_t = (h_1(L)+h_2(L))\eta_t = \alpha\chi\frac{\vartheta}{\rho}\frac{1}{1-\vartheta L}\eta_t.$$

Finally, χ can be obtained by solving the following linear equation

$$\chi = \varphi + \alpha h_1(0) = \varphi + \alpha_1^2\chi\frac{\rho\tau_u+\vartheta\tau_a}{(\tau_a+\tau_u)\rho} = \frac{\varphi\rho(\tau_a+\tau_u)}{\rho(\tau_a+\tau_u)-\alpha^2(\rho\tau_u+\vartheta\tau_a)}.$$

A.3 Proof of Corollary 1

By the Wiener-Hopf prepdiction formula, the forecast about ξ_t is given by

$$\mathbb{E}_{it}[\xi_t] = \frac{\lambda}{\rho}\frac{1}{(1-\lambda L)(L-\lambda)}\begin{bmatrix}\tau_a\left(L-\lambda\frac{1-\rho L}{1-\rho\lambda}\right) & \tau_u\left(L-\lambda\frac{1-\rho L}{1-\rho\lambda}\right)\end{bmatrix}\begin{bmatrix}x_{it}^1\\ x_{it}^2\end{bmatrix},$$

and the aggregate forecast is

$$\overline{\mathbb{E}}_t[\xi_t] = \left(1-\frac{\lambda}{\rho}\right)\frac{1}{(1-\lambda L)(1-\rho L)}\eta_t.$$

The aggregate forecast error follows

$$\xi_t - \overline{\mathbb{E}}_t[\xi_t] = \frac{\lambda}{\rho}\frac{1}{1-\lambda L}\eta_t.$$

Consider an auxiliary signal process with enhanced variance of the noise $\hat{\sigma}_u^2$ and dispersion of productivity $\hat{\sigma}_a^2$, such that $\hat{\tau}_a = \frac{\sigma_\eta^2}{\hat{\sigma}_a^2} = (1-\alpha)\tau_a$ and $\hat{\tau}_u = \frac{\sigma_\eta^2}{\hat{\sigma}_u^2} = (1-\alpha)\tau_u$. Under this alternative signal process, the forecast error about the confidence shock is simply replacing λ by ϑ

$$\xi_t - \widetilde{\mathbb{E}}_t[\xi_t] = \frac{\vartheta}{\rho}\frac{1}{1-\vartheta L}\eta_t.$$

Note that the aggregate output is $y_t = \alpha\chi\frac{\vartheta}{\rho}\frac{1}{1-\vartheta L}\eta_t = \alpha\chi(\xi_t - \widetilde{\mathbb{E}}_t[\xi_t])$, which completes the proof.

A.4 Proof of Lemma 2.4 and Proposition 2.5

Define $f(a)$ as the inside root of the following function

$$g(x) \equiv x + \frac{1}{x} - a,$$

where $a > 2$ and the inside root refers to the root inside the unit circle. It is easy to show that $f(a)$ is decreasing in a.

Note that ϑ satisfies the following condition

$$\vartheta + \frac{1}{\vartheta} = \rho + \frac{1}{\rho} + \frac{(1-\alpha)(\tau_a + \tau_u)}{\rho}.$$

It follows that $\vartheta = f\left(\rho + \frac{1}{\rho} + \frac{(1-\alpha)(\tau_a+\tau_u)}{\rho}\right)$. Since the term $\frac{(1-\alpha)(\tau_a+\tau_u)}{\rho}$ is increasing in τ_u and decreasing in α, ϑ is decreasing in τ_u and increasing in α.

Meanwhile, λ satisfies

$$\lambda + \frac{1}{\lambda} = \rho + \frac{1}{\rho} + \frac{\tau_a + \tau_u}{\rho},$$

or $\lambda = f\left(\rho + \frac{1}{\rho} + \frac{\tau_a+\tau_u}{\rho}\right)$. It follows that $\vartheta > \lambda$.

The variance of the output is

$$\mathrm{V}(y_t) = \alpha^2 \chi^2 \frac{\vartheta^2}{\rho^2} \frac{1}{1-\vartheta^2} \sigma_\eta^2.$$

Recognizing that $\vartheta < \rho$, it is straightforward to show χ is increasing in α and decreasing in τ_u. Meanwhile, we have shown that ϑ is increasing in α and decreasing in τ_u. Since $\mathrm{V}(y_t)$ is increasing in both α and ϑ, $\mathrm{V}(y_t)$ shares the same comparative statics with respect to α and τ_u.

A.5 Proof of Proposition 2.6

Part 1. We first show that the the equilibrium outcome is equivalent to a modified first-order expectation error about the confidence shock. Denote the policy function as

$$y_{it} = h_a a_i + h_1(L) x_{it}^1 + h_2(L) x_{it}^2.$$

The best response can be expressed as

$$y_{it} = \varphi a_i + \alpha \mathbb{E}_{it}[y_{jt}] = \varphi a_i + \alpha h_a \mathbb{E}_{it}[a_{jt}] + \alpha \mathbb{E}_{it}[y_t].$$

Therefore, the aggregate output follows

$$y_t = \alpha h_a(\xi_t - \overline{\mathbb{E}}_t[\xi_t]) + \alpha \mathbb{E}_{it}[y_t].$$

By the equivalence result from Huo and Pedroni (2020), the aggregate output can be expressed as

$$y_t = \alpha h_a(\xi_t - \widetilde{\overline{\mathbb{E}}}_t[\xi_t]),$$

where the expectation $\widetilde{\mathbb{E}}$ is based on a modified signal process where the precision of private noises is discounted by α.

Based on this observation, we show that for any σ_a and σ_ε, the forecast error about ξ_t cannot be consistent with the perceived law of motion of output.

Part 2. Let $\phi(L) \equiv \frac{a(L)}{b(L)}$ denote the perceived law motion of the aggregate outcome, where $a(L)$ and $b(L)$ are finite polynomials, $b(L)$ does not contain any inside root, and $a(L)$ and $b(L)$ do not contain any common root. That is,

$$a_t = \phi(L)\eta_t = \frac{a(L)}{b(L)}\eta_t.$$

Assume that the fundamental follows an AR(1) process $\xi_t = \frac{1}{1-\rho L}\eta_t$. The signal process is summarized by

$$\boldsymbol{x}_{it} = \mathbf{M}(L)\begin{bmatrix}\eta_t \\ \hat{a}_{jt} \\ \hat{\varepsilon}_{it}\end{bmatrix}, \quad \text{where} \quad \mathbf{M}(L) \equiv \begin{bmatrix} \frac{1}{1-\rho L} & \sigma_a & 0 \\ \frac{a(L)}{b(L)} & 0 & \sigma_\varepsilon \end{bmatrix},$$

where $\hat{a}_{jt}$ and $\hat{\varepsilon}_{it}$ are normalized shocks.

Step 1: fundamental representation. We want to first conduct a spectral factorization of the signal process

$$\begin{aligned}
&\mathbf{M}(z)\mathbf{M}'(z^{-1}) = \begin{bmatrix} \frac{1+\sigma_a^2(1-\rho z)(1-\rho z^{-1})}{(1-\rho z)(1-\rho z^{-1})} & \frac{a(z^{-1})}{(1-\rho z)b(z^{-1})} \\ \frac{a(z)}{(1-\rho z^{-1})b(z)} & \frac{a(z)a(z^{-1})}{b(z)b(z^{-1})} + \sigma_\varepsilon^2 \end{bmatrix}, \\
&= \begin{bmatrix} 1 & 0 \\ \frac{(1-\rho z)a(z)}{[1+\sigma_a^2(1-\rho z)(1-\rho z^{-1})]b(z)} & 1 \end{bmatrix} \begin{bmatrix} \frac{1+\sigma_a^2(1-\rho z)(1-\rho z^{-1})}{(1-\rho z)(1-\rho z^{-1})} & 0 \\ 0 & \frac{\sigma_a^2(1-\rho z)(1-\rho z^{-1})a(z)a(z^{-1})+\sigma_\varepsilon^2[1+\sigma_a^2(1-\rho z)(1-\rho z^{-1})]b(z)b(z^{-1})}{[1+\sigma_a^2(1-\rho z)(1-\rho z^{-1})]b(z)b(z^{-1})} \end{bmatrix} \\
&\times \begin{bmatrix} 1 & \frac{(1-\rho z^{-1})a(z^{-1})}{[1+\sigma_a^2 p(z)p(z^{-1})]b(z^{-1})} \\ 0 & 1 \end{bmatrix}, \\
&= \begin{bmatrix} 1 & 0 \\ \frac{(1-\rho z)a(z)}{d^2(1-\lambda z)(1-\lambda z^{-1})b(z)} & 1 \end{bmatrix} \begin{bmatrix} \frac{d^2(1-\lambda z)(1-\lambda z^{-1})}{(1-\rho z)(1-\rho z^{-1})} & 0 \\ 0 & \frac{\delta(z)\delta(z^{-1})}{d^2(1-\lambda z)(1-\lambda z^{-1})b(z)b(z^{-1})} \end{bmatrix} \begin{bmatrix} 1 & \frac{(1-\rho z^{-1})a(z^{-1})}{d^2(1-\lambda z)(1-\lambda z^{-1})b(z^{-1})} \\ 0 & 1 \end{bmatrix},
\end{aligned}$$

where $\lambda = \frac{\rho+\frac{1}{\rho}+\frac{1}{\sigma_a^2\rho}-\sqrt{\left(\rho+\frac{1}{\rho}+\frac{1}{\sigma_a^2\rho}\right)^2-4}}{2}$, $|\lambda| < 1$ and $d = \sqrt{\frac{\sigma_a^2\rho}{\lambda}}$ satisfying

$$1 + \sigma_a^2(1-\rho z)(1-\rho z^{-1}) = \frac{\sigma_a^2\rho}{\lambda}(1-\lambda z)(1-\lambda z^{-1}),$$

and we can choose $\delta(z)$ so that all the roots of $\delta(z)$ are outside the unit circle and

$$\delta(z)\delta(z^{-1}) = \sigma_a^2(1-\rho z)(1-\rho z^{-1})a(z)a(z^{-1}) + \sigma_\varepsilon^2[1+\sigma_a^2(1-\rho z)(1-\rho z^{-1})]b(L)b(z^{-1}).$$

For later use, we establish the following property of $\delta(z)$:

$$\text{If } a(\mu) = 0 \text{ and } b(\mu)b(\mu^{-1}) = 0, \text{ then } \mu \text{ cannot be a root of } \delta(z). \tag{A.1}$$

To see this, not that if $a(\mu) = 0$, it can only be the case that $b(\mu^{-1}) = 0$ as $a(L)$ and $b(L)$ do not contain a common root by assumption. Therefore, μ is inside the unit circle. Also by assumption, $\delta(L)$ only contains roots outside the unit circle, which makes $\delta(\mu) = 0$ impossible.

Continuing the factorization,

$$
\begin{aligned}
&\mathbf{M}(z)\mathbf{M}'(z^{-1}) \\
&= \begin{bmatrix} 1 & 0 \\ \frac{(1-\rho z)a(z)z}{d^2(1-\lambda z)(z-\lambda)b(z)} & 1 \end{bmatrix} \begin{bmatrix} \frac{d(z-\lambda)}{1-\rho z} & 0 \\ 0 & \frac{\delta(z)}{d(1-\lambda z)b(z)} \end{bmatrix} \begin{bmatrix} \frac{d(z^{-1}-\lambda)}{1-\rho z^{-1}} & 0 \\ 0 & \frac{\delta(z^{-1})}{d(1-\lambda z^{-1})b(z^{-1})} \end{bmatrix} \begin{bmatrix} 1 & \frac{(1-\rho z^{-1})a(z^{-1})z^{-1}}{d^2(z^{-1}-\lambda)(1-\lambda z^{-1})b(z^{-1})} \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{d(z-\lambda)}{1-\rho z} & 0 \\ \frac{a(z)z}{d(1-\lambda z)b(z)} & \frac{\delta(z)}{d(1-\lambda z)b(z)} \end{bmatrix} \begin{bmatrix} \frac{d(z^{-1}-\lambda)}{1-\rho z^{-1}} & \frac{a(z^{-1})z^{-1}}{d(1-\lambda z^{-1})b(z^{-1})} \\ 0 & \frac{\delta(z^{-1})}{d(1-\lambda z^{-1})b(z^{-1})} \end{bmatrix} \\
&\equiv \tilde{\mathbf{\Gamma}}(z)\tilde{\mathbf{\Gamma}}'(z^{-1}).
\end{aligned}
$$

Note that $\tilde{\mathbf{\Gamma}}(z)$ is analytic inside the unit circle, but $\det \tilde{\mathbf{\Gamma}}(z) = \frac{(z-\lambda)\delta(z)}{(1-\rho z)(1-\lambda z)b(z)}$ which contains an inside root at $z = \lambda$. To obtain the fundamental representation, we need to use the Blaschke matrix to remove this inside root.

Define $\mathbf{\Gamma}(z) = \tilde{\mathbf{\Gamma}}(z)\mathbf{B}(z)\mathbf{V}$ where $\mathbf{B}(z)$ is the Blaschke matrix and $\mathbf{V}$ is an orthogonal rotation matrix

$$
\mathbf{V} = \begin{bmatrix} \frac{\delta(\lambda)}{\sqrt{a^2(\lambda)\lambda^2+\delta^2(\lambda)}} & \frac{a(\lambda)\lambda}{\sqrt{a^2(\lambda)\lambda^2+\delta^2(\lambda)}} \\ \frac{-a(\lambda)\lambda}{\sqrt{a^2(\lambda)\lambda^2+\delta^2(\lambda)}} & \frac{\delta(\lambda)}{\sqrt{a^2(\lambda)\lambda^2+\delta^2(\lambda)}} \end{bmatrix}, \quad \mathbf{B}(z) = \begin{bmatrix} \frac{1-\lambda z}{z-\lambda} & 0 \\ 0 & 1 \end{bmatrix}.
$$

It follows that

$$
\mathbf{\Gamma}(z) = \frac{1}{\sqrt{a^2(\lambda)\lambda^2 + \delta^2(\lambda)}} \begin{bmatrix} \frac{d(1-\lambda z)\delta(\lambda)}{1-\rho z} & \frac{d(z-\lambda)a(\lambda)\lambda}{1-\rho z} \\ \frac{\delta(\lambda)a(z)z-a(\lambda)\lambda\delta(z)}{d(z-\lambda)b(z)} & \frac{a(\lambda)\lambda a(z)z+\delta(\lambda)\delta(z)}{d(1-\lambda z)b(z)} \end{bmatrix}.
$$

The matrix $\mathbf{\Gamma}(z)$ is a fundamental representation of the origianl signal process as $\det \mathbf{\Gamma}(z) = \frac{\delta(z)}{(1-\rho z)b(z)}$ which does not have zeros inside the unit circle.

The average forecast about ξ_t is given by the Wiener-Hopf prediction formula

$$
\begin{aligned}
&\overline{\mathbb{E}}_t\left[\frac{1}{1-\rho L}\eta_t\right] = \left[\begin{bmatrix} \frac{1}{1-\rho L} & 0 & 0 \end{bmatrix}\mathbf{M}'(L^{-1})\mathbf{\Gamma}'(L^{-1})^{-1}\right]_+ \mathbf{\Gamma}(L)^{-1}\begin{bmatrix} \frac{1}{1-\rho L} \\ \frac{a(L)}{b(L)} \end{bmatrix}\eta_t, \\
&= \left[\begin{bmatrix} \frac{a(\lambda)\lambda a(L^{-1})L^{-1}+\delta(\lambda)\delta(L^{-1})-a(L^{-1})d^2(L^{-1}-\lambda)(1-\lambda L^{-1})a(\lambda)\lambda}{d(1-\rho L)(1-\lambda L^{-1})\delta(L^{-1})} \\ \frac{-\delta(\lambda)a(L^{-1})L^{-1}+a(\lambda)\lambda\delta(L^{-1})+a(L^{-1})d^2(L^{-1}-\lambda)(1-\lambda L^{-1})\delta(\lambda)}{d(1-\rho L)(L^{-1}-\lambda)\delta(L^{-1})} \end{bmatrix}^T\right]_+ \begin{bmatrix} \frac{a(\lambda)\lambda a(L)L+\delta(\lambda)\delta(L)-d^2(1-\lambda L)(L-\lambda)a(\lambda)\lambda a(L)}{d[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda L)\delta(L)} \\ \frac{-\delta(\lambda)a(L)L+a(\lambda)\lambda\delta(L)+d^2(1-\lambda L)(L-\lambda)\delta(\lambda)a(L)}{d[a^2(\lambda)\lambda^2+\delta^2(\lambda)](L-\lambda)\delta(L)} \end{bmatrix}\eta_t.
\end{aligned}
$$

To conduct the annihilation operation, first recall that $1 + \sigma_a^2(1-\rho z)(1-\rho z^{-1}) = d^2(1-\lambda z)(1-\lambda z^{-1})$. We have

$$
\begin{aligned}
&\left[\frac{a(\lambda)\lambda a(L^{-1})L^{-1} + \delta(\lambda)\delta(L^{-1}) - a(L^{-1})d^2(L^{-1}-\lambda)(1-\lambda L^{-1})a(\lambda)\lambda}{d(1-\rho L)(1-\lambda L^{-1})\delta(L^{-1})}\right]_+ \\
&= \left[\frac{\delta(\lambda)}{d(1-\rho L)(1-\lambda L^{-1})} - \frac{\sigma_a^2(1-\rho L^{-1})a(\lambda)\lambda a(L^{-1})L^{-1}}{d(1-\lambda L^{-1})\delta(L^{-1})}\right]_+ \\
&= \frac{\delta(\lambda)}{d(1-\rho L)(1-\lambda\rho)}.
\end{aligned}
$$

The last line follows from the fact that the term $\frac{\sigma_a^2(1-\rho L^{-1})a(\lambda)\lambda a(L^{-1})L^{-1}}{d(1-\lambda L^{-1})\delta(L^{-1})}$ only contains L with negative powers.

Next note that

$$\begin{aligned}
&\left[\frac{-\delta(\lambda)a(L^{-1})L^{-1}+a(\lambda)\lambda\delta(L^{-1})+a(L^{-1})d^2(L^{-1}-\lambda)(1-\lambda L^{-1})\delta(\lambda)}{d(1-\rho L)(L^{-1}-\lambda)\delta(L^{-1})}\right]_+\\
=&\left[\frac{a(\lambda)\lambda\delta(L^{-1})L+\sigma_a^2(1-\rho L)(1-\rho L^{-1})\delta(\lambda)a(L^{-1})}{d(1-\rho L)(1-\lambda L)\delta(L^{-1})}\right]_+\\
=&\frac{\kappa_1}{1-\rho L}+\frac{\kappa_2}{1-\lambda L},
\end{aligned}$$

where κ_1 and κ_2 are some constants. The last line follows from the partial fraction decomposition and the fact that all the roots of $\delta(L^{-1})$ are inside the unit circle and the related terms are removed by the annihilation operator. Note that we can solve for κ_1 and κ_2 by the Heaviside expansion theorem. Also note that the numerator $-\delta(\lambda)a(L^{-1})L^{-1}+a(\lambda)\lambda\delta(L^{-1})+a(L^{-1})d^2(L^{-1}-\lambda)(1-\lambda L^{-1})\delta(\lambda)$ equals zero at $L=\lambda^{-1}$, which implies that $\kappa_2=0$. Finally, we can obtain κ_1 as

$$\kappa_1=\left.\frac{a(\lambda)\lambda\delta(L^{-1})L+\sigma_a^2(1-\rho L)(1-\rho L^{-1})\delta(\lambda)a(L^{-1})}{d(1-\lambda L)\delta(L^{-1})}\right|_{L=\rho^{-1}}=\frac{a(\lambda)\lambda}{d(\rho-\lambda)}.$$

It follows the average forecast can be expressed as

$$\begin{aligned}
\overline{\mathbb{E}}_t\left[\frac{1}{1-\rho L}\eta_t\right]=&\begin{bmatrix}\frac{\delta(\lambda)}{d(1-\rho L)(1-\lambda\rho)} & \frac{a(\lambda)\lambda}{d(1-\rho L)(\rho-\lambda)}\end{bmatrix}\begin{bmatrix}\frac{a(\lambda)\lambda a(L)L+\delta(\lambda)\delta(L)-d^2(1-\lambda L)(L-\lambda)a(\lambda)\lambda a(L)}{d[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda L)\delta(L)}\\ \frac{-\delta(\lambda)a(L)L+a(\lambda)\lambda\delta(L)+d^2(1-\lambda L)(L-\lambda)\delta(\lambda)a(L)}{d[a^2(\lambda)\lambda^2+\delta^2(\lambda)](L-\lambda)\delta(L)}\end{bmatrix}\eta_t\\
=&\frac{\delta^2(\lambda)(\rho-\lambda)(L-\lambda)\delta(L)+(\rho-\lambda)(L-\lambda)a(\lambda)\lambda\delta(\lambda)[1-d^2(1-\lambda L)(1-\lambda L^{-1})]a(L)L}{d^2[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda\rho)(\rho-\lambda)(1-\lambda L)(L-\lambda)(1-\rho L)\delta(L)}\\
&+\frac{a^2(\lambda)\lambda^2(1-\lambda\rho)(1-\lambda L)\delta(L)-(1-\lambda\rho)(1-\lambda L)a(\lambda)\lambda\delta(\lambda)[1-d^2(1-\lambda L)(1-\lambda L^{-1})]a(L)L}{d^2[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda\rho)(\rho-\lambda)(1-\lambda L)(L-\lambda)(1-\rho L)\delta(L)}\eta_t,
\end{aligned}$$

and hence the average forecast error $\psi(L)\eta_t\equiv-\left(\overline{\mathbb{E}}_t\left[\frac{1}{1-\rho L}\eta_t\right]-\frac{1}{1-\rho L}\eta_t\right)$ is given by

$$\begin{aligned}
\psi(L)=&-\frac{\begin{bmatrix}[\delta^2(\lambda)(\rho-\lambda)(L-\lambda)+a^2(\lambda)\lambda^2(1-\lambda\rho)(1-\lambda L)]\delta(L)\\ -d^2[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda\rho)(\rho-\lambda)(1-\lambda L)(L-\lambda)\delta(L)\\ +[(\rho-\lambda)(L-\lambda)-(1-\lambda\rho)(1-\lambda L)]a(\lambda)\lambda\delta(\lambda)[1-d^2(1-\lambda L)(1-\lambda L^{-1})]a(L)L\end{bmatrix}}{d^2[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda\rho)(\rho-\lambda)(1-\lambda L)(L-\lambda)(1-\rho L)\delta(L)}\\
=&-\frac{\begin{bmatrix}\begin{pmatrix}\{\delta^2(\lambda)(\rho-\lambda)-a^2(\lambda)\lambda^2(1-\lambda\rho)\lambda-[a^2(\lambda)\lambda^4+\delta^2(\lambda)]d^2(1-\lambda\rho)(\rho-\lambda)\}L\\ -\{\delta^2(\lambda)(\rho-\lambda)\lambda-a^2(\lambda)\lambda^2(1-\lambda\rho)-[a^2(\lambda)\lambda^2+\delta^2(\lambda)]d^2(1-\lambda\rho)(\rho-\lambda)\lambda\}\end{pmatrix}\delta(L)\\ +(1-\lambda^2)a(\lambda)\lambda\delta(\lambda)\sigma_a^2(1-\rho L)(L-\rho)a(L)\end{bmatrix}}{d^2[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda\rho)(\rho-\lambda)(1-\lambda L)(L-\lambda)\delta(L)},\\
=&\frac{\mu_1(L-\mu_3)\delta(L)+\mu_2\sigma_a^2(1-\rho L)(L-\rho)a(L)}{(1-\lambda L)(L-\lambda)\delta(L)},
\end{aligned}$$

where we have use the the definition of λ to reach the second equality and μ_1, μ_2 and μ_3 are constants.

It is straightforward to verify that λ and λ^{-1} are roots of the numerator, which implies the term $(1-\lambda L)(L-\lambda)$ in the denominator can be removed. In addition, μ_3 is different from either λ or λ^{-1} as

$$\mu_1(\lambda-\mu_3)=-\frac{a^2(\lambda)\lambda^2(1-\lambda\rho)(1-\lambda^2)[1+d^2(\rho-\lambda)\lambda]}{d^2[a^2(\lambda)\lambda^2+\delta^2(\lambda)](1-\lambda\rho)(\rho-\lambda)}\neq 0,$$

and

$$\mu_1(\lambda^{-1} - \mu_3) = -\frac{\delta^2(\lambda)(\rho - \lambda)(1 - \lambda^2)\lambda^{-1}\sigma_a^2[\rho\lambda - 1]}{d^2[a^2(\lambda)\lambda^2 + \delta^2(\lambda)](1 - \lambda\rho)(\rho - \lambda)} \neq 0.$$

These properties of $\psi(L)$ will be used in the next step of the proof.

Step 3: Verifying Equilibrium Requirement. So, for $\psi(L)$ to match with $\phi(L)$, the following three conditions must be satisfied:

1. $\delta(L)$ contains all the roots of $b(L)$ if $b(L)$ is not a constant.
2. $(L - \mu_3)\delta(L)$ contains all the roots of $a(L)$ if $a(L)$ is not a constant.
3. The difference of the order of the numerator and the denominator of $\psi(L)$ is consistent with that of $a(L)$ and $b(L)$.

Now, recall that

$$\delta(L)\delta(L^{-1}) = \sigma_a^2(1 - \rho L)(1 - \rho L^{-1})a(L)a(L^{-1}) + \sigma_\varepsilon^2[1 + \sigma_a^2(1 - \rho L)(1 - \rho L^{-1})]b(L)b(L^{-1}).$$

Let n_λ, n_a, n_b be the orders of $\delta(L), a(L), b(L)$ respectively. Note that condition 1 implies $n_\lambda \geq n_b$. We will now show that those conditions lead to a contradiction for every possible combination of n_a and n_b.

1. $n_a \geq 3$.

   When $(L - \mu_3)\delta(L)$ to contain all the roots of $a(L)$, it has to be the case that $\delta(L)$ contains at least two of the roots of $a(L)$. Then, at least one of two roots has to be a root of $b(L)b(L^{-1})$. By property (A.1), this is not possible.

2. $n_a = 2$

   (a) $n_b \geq 3$

   For $\delta(L)$ to contain all the roots of $b(L)$, all the roots of $a(L^{-1})$ must be the roots of $b(L)$ since $a(L)$ and $b(L)$ do not have common roots. It follows that all the roots of $a(L)$ are inside the unit circle since $b(L)$ does not have any inside roots. But when $(L - \mu_3)\delta(L)$ to contain all the roots of $a(L)$, $\delta(L)$ has to contain at least one inside root of $a(L)$. Contradiction.

   (b) $n_b = 0, 1, 2$

      i. $n_\lambda = n_b$

      In order for $(L - \mu_3)\delta(L)$ to contain all the roots of $a(L)$, $\delta(L)$ has to contain at least one of the roots of $a(L)$. If $n_b > 0$, in order for $\delta(L)$ to contain all the roots of $b(L)$, $\delta(L) = d_\lambda b(L)$ must hold for some constant d_λ. Therefore, $a(L)$ and $b(L)$ have at least one common root, which is a contradiction. If $n_b = 0$, $\delta(L)$ is a constant, which cannot have a root of $a(L)$.

      ii. $n_\lambda > n_b$

      Note that $3 \geq n_\lambda > n_b$. By inspecting $\psi(L)$, the order of the numerator is at most 4, and the order of the denominator is $2 + n_\lambda$, so their difference is at least $n_\lambda - 2$. But we have $n_\lambda - 2 > n_b - n_a$ by $n_a = 2$. Therefore, condition 3 on the difference of the order between n_a and n_b can never be satisfied.

3. $n_a = 1$

   (a) $n_b \geq 2$

   In order for $\delta(L)$ to contain all the roots of $b(L)$, the roots of $b(L)$ have to coincide with the roots of $(1 - \rho L)a(L^{-1})$, since $a(L)$ and $b(L)$ do not have common roots. By property (A.1), the root of $a(L)$ cannot be the root of $\delta(L)$, and the only possibility is that $a(L) = d_a(L-\mu_3)$ for some constant d_a. Therefore, $b(L) = d_b(1-\rho L)(1-\mu_3 L)$ for some constant d_b, and $\delta(L)$ satisfies

$$\begin{aligned}\delta(L)\delta(L^{-1}) =& (1-\rho L)(1-\rho L^{-1})(1-\mu_3 L)(1-\mu_3 L^{-1})(\sigma_a^2 d_a^2 + \sigma_\varepsilon^2 d_b^2 + \sigma_\varepsilon^2\sigma_a^2 d_b^2(1-\rho L)(1-\rho L^{-1}))\\ \equiv& (1-\rho L)(1-\rho L^{-1})(1-\mu_3 L)(1-\mu_3 L^{-1})d_\lambda^2(1-\mu_4 L)(1-\mu_4 L^{-1}),\end{aligned}$$

where μ_4 is inside the unit circle and $d_\lambda > 0$ is some constant. Therefore, $\psi(L)$ can be expressed as

$$\psi(L) = \frac{[\mu_1\delta(L) + d_a\mu_2\sigma_a^2(1-\rho L)(L-\rho)](L-\mu_3)}{(1-\lambda L)(L-\lambda)\delta(L)} = \frac{[\mu_1 d_\lambda(1-\mu_3 L)(1-\mu_4 L) + d_a\mu_2\sigma_a^2(L-\rho)](L-\mu_3)}{(1-\lambda L)(L-\lambda)d_\lambda(1-\mu_3 L)(1-\mu_4 L)}.$$

For the denominator to match with $b(L)$, it has to be the case that $\mu_4 = \rho$. However, this is impossible as it requires $\sigma_a^2 d_a^2 + \sigma_\varepsilon^2 d_b^2 = 0$ and $\sigma_\varepsilon^2\sigma_a^2 d_b^2 \neq 0$.

(b) $n_b = 0, 1$

i. $n_\lambda = n_b$

For $(L-\mu_3)\delta(L)$ to contain the root of $a(L)$, $a(L) = d_a(L-\mu_3)$ must hold for some constant d_a. This is because $a(L)$ and $b(L)$ do not have any common root when $n_b = 1$ and $\delta(L)$ is a constant when $n_b = 0$. Then $\psi(L)$ can be expressed as

$$\psi(L) = \frac{[\mu_1\delta(L) + d_a\mu_2\sigma_a^2(1-\rho L)(L-\rho)](L-\mu_3)}{(1-\lambda L)(L-\lambda)\delta(L)} = \frac{[\mu_1\delta(L) - d_a\mu_2 L + d_a\mu_2 d^2(1-\lambda L)(L-\lambda)](L-\mu_3)}{(1-\lambda L)(L-\lambda)\delta(L)}.$$

Due to the fact that $\mu_1\delta(L) - d_a\mu_2 L$ is merely linear and $\mu_3 \neq \lambda$ or λ^{-1}, λ and λ^{-1} cannot both be the roots of the numerator. Contradiction.

ii. $n_\lambda > n_b$

By inspecting $\psi(L)$, the order of the numerator is at most 3 and that of the denominator is at least $3 + n_b$. The difference between the denominator and the numerator is at least n_b, which is always larger than $n_b - n_a = n_b - 1$. Therefore, $\psi(L)$ cannot match with $\phi(L)$.

4. $n_a = 0$

(a) $n_b \geq 1$

In this case, $a(L) = d_a$ for some constant d_a. For $\delta(L)$ to contain all the roots of $b(L)$, $b(L) = d_b(1-\rho L)$ has to hold for some constant d_b and $\delta(L)$ satisfies

$$\delta(L)\delta(L^{-1}) = (1-\rho L)(1-\rho L^{-1})(\sigma_a^2 d_a^2 + \sigma_\varepsilon^2 d_b^2 + \sigma_\varepsilon^2\sigma_a^2 d_b^2(1-\rho L)(1-\rho L^{-1})) \equiv (1-\rho L)(1-\rho L^{-1})d_\lambda^2(1-\mu_4 L)(1-\mu_4 L^{-1}),$$

where μ_4 is inside the unit circle and d_λ is some constant. It follows that

$$\psi(L) = \frac{\mu_1(L-\mu_3)\delta(L) + \mu_2\sigma_a^2(1-\rho L)(L-\rho)d_a}{(1-\lambda L)(L-\lambda)\delta(L)} = \frac{d_\lambda\mu_1(L-\mu_3)(1-\mu_4 L) + \mu_2\sigma_a^2(L-\rho)d_a}{(1-\lambda L)(L-\lambda)d_\lambda(1-\mu_4 L)}.$$

For the denominator to match with $b(L)$, it has to be the case that $\mu_4 = \rho$. However, this is impossible as it requires $\sigma_a^2 d_a^2 + \sigma_\varepsilon^2 d_b^2 = 0$ and $\sigma_\varepsilon^2\sigma_a^2 d_b^2 \neq 0$.

(b) $n_b = 0$

In this case, $n_\lambda = 1$. Then, by inspecting $\psi(L)$, the order of the numerator is at most 2 and that of the denominator is 3. But $n_a = n_b$ is assumed in this case, which leads to a contradiction.

Thus, $\psi(L)$ never matches with $\phi(L)$.

A.6 Proof of Lemma 2 and Proposition 3.1

As derived in the proof of Proposition 2.3, the nowcast about the aggregate output is

$$\mathbb{E}_{it}\left[\alpha\chi\frac{\vartheta}{\rho}\frac{1}{1-\vartheta L}\eta_t\right]=\alpha\chi\frac{\vartheta}{\rho}\frac{1}{1-\lambda L}\left[\frac{\lambda\tau_a}{\rho(L-\lambda)}\left(\frac{L}{1-\vartheta L}-\frac{\lambda}{1-\lambda\vartheta}\frac{1-\rho L}{1-\rho\lambda}\right)\quad\frac{\lambda\tau_u}{\rho(L-\lambda)}\left(\frac{L}{1-\vartheta L}-\frac{\lambda}{1-\lambda\vartheta}\frac{1-\rho L}{1-\rho\lambda}\right)\right]\begin{bmatrix}x_{it}^1\\x_{it}^2\end{bmatrix}.$$

The one-step ahead aggregate forecast is

$$\overline{\mathbb{E}}_t[y_{t+1}]=\vartheta\overline{\mathbb{E}}_t[y_{t+1}]=\vartheta\alpha\chi\frac{\vartheta}{\rho}\left(1-\frac{\lambda}{\rho}\right)\frac{1-\lambda\rho}{1-\lambda\vartheta}\frac{1}{(1-\lambda L)(1-\vartheta L)}\eta_t,$$

and the forecast error is

$$\begin{aligned}y_{t+1}-\overline{\mathbb{E}}_t[y_{t+1}]&=\alpha\chi\frac{\vartheta}{\rho}\eta_{t+1}+\alpha\chi\frac{\vartheta^2}{\rho}\left(\frac{1}{1-\vartheta L}-\left(1-\frac{\lambda}{\rho}\right)\frac{1-\lambda\rho}{1-\lambda\vartheta}\frac{1}{(1-\lambda L)(1-\vartheta L)}\right)\eta_t\\&=\alpha\chi\frac{\vartheta}{\rho}\eta_{t+1}+\chi\frac{\vartheta}{\rho}\left(\frac{\lambda}{1-\lambda L}-\frac{(1-\alpha)\vartheta}{1-\vartheta L}\right)\eta_t,\end{aligned}$$

where we have used the property that $\lambda+\frac{1}{\lambda}=\vartheta+\frac{1}{\vartheta}+\frac{\alpha(\tau_a+\tau_u)}{\rho}$. It follows that the IRF of the forecast error is

$$\zeta_0=\alpha\chi\frac{\vartheta}{\rho},\quad\text{and for } k>0,\quad \zeta_k=\chi\frac{\vartheta}{\rho}\left(\lambda^k-(1-\alpha)\vartheta^k\right).$$

When $k=1$, $\zeta_1=\alpha\chi\frac{\vartheta^2}{\rho}\left(1-\left(1-\frac{\lambda}{\rho}\right)\frac{1-\lambda\rho}{1-\lambda\vartheta}\right)$. Since $\lambda\in(\vartheta,\rho)$, it follows that $\zeta_1>0$. The fact that $\lambda<\vartheta$ also implies that ζ_k eventually becomes negative as the effect of ϑ dominates.

To show part (1) of Proposition 3.1, note that the forecast error is i.i.d ($zeta_k=0$ for $k\geq 1$) only if $\rho=0$ or $\tau_u+\tau_a=\infty$. In either case, $\vartheta=0$ and the output process cannot be persistent.

To show part (2) of Proposition 3.1, consider the following function $g(t)$

$$g(t)\equiv\lambda^t-(1-\alpha)\vartheta^t.$$

When the forecast error reaches zero, we have $g(T)=0$ where T is given by

$$T=\frac{\log(1-\alpha)^{-1}}{\log\frac{\vartheta}{\lambda}}.$$

Define $\tau\equiv\tau_a+\tau_u$. We now prove that $\frac{\vartheta}{\lambda}$ is increasing in τ, which will imply T is decreasing in τ.

$$\frac{\partial\frac{\vartheta}{\lambda}}{\partial\tau}=\frac{\left(\sqrt{\left(\rho+\frac{1+(1-\alpha)\tau}{\rho}\right)^2-4}-(1-\alpha)\sqrt{\left(\frac{1+\tau}{\rho}+\rho\right)^2-4}\right)\left(\frac{1+\tau}{\rho}+\rho-\sqrt{\left(\frac{1+\tau}{\rho}+\rho\right)^2-4}\right)\left(\frac{1+(1-\alpha)\tau}{\rho}+\rho-\sqrt{\left(\rho+\frac{1+(1-\alpha)\tau}{\rho}\right)^2-4}\right)}{\rho\left(\left(\frac{1}{\rho}+\rho+\frac{\tau}{\rho}\right)-\sqrt{\left(\frac{1}{\rho}+\rho+\frac{\tau}{\rho}\right)^2-4}\right)^2\sqrt{\left(\frac{1}{\rho}+\rho+\frac{\tau}{\rho}\right)^2-4}\sqrt{\left(\frac{1}{\rho}+\rho+\frac{(1-\alpha)\tau}{\rho}\right)^2-4}},$$

and hence it suffices to show that

$$\sqrt{\left(\rho+\frac{1+(1-\alpha)\tau}{\rho}\right)^2-4}>(1-\alpha)\sqrt{\left(\rho+\frac{1+\tau}{\rho}\right)^2-4}.$$

This is true because

$$\left(\rho+\frac{1+(1-\alpha)\tau}{\rho}\right)^2-4-(1-\alpha)^2\left(\left(\rho+\frac{1+\tau}{\rho}\right)^2-4\right)=2\left(\frac{1}{\rho}+\rho\right)\frac{(1-\alpha)}{\rho}\alpha\tau+(1-(1-\alpha)^2)\left(\left(\frac{1}{\rho}+\rho\right)^2-4\right)>0.$$

A.7 Proof of Proposition 3.2

With heterogeneous prior, agents can observe their the confidence shock perfectly, but perceive others' signals are biased by the amount of ξ_t and ξ_t is commonly observed. To solve for the equilibrium outcome, we proceed with a guess-and-verify approach. Guess that producers in island i's policy rule is

$$y_{it}=f_1a_i+f_2a_{jt}+f_3\xi_t,$$

then agent i believes that her trading partner's output is

$$y_{jt}=f_1a_{jt}+f_2(a_i+\xi_t)+f_3\xi_t.$$

In equilibrium,

$$y_{it}=\varphi a_i+\alpha\mathbb{E}_{it}[y_{jt}],$$

which leads to

$$y_{it}=\frac{\varphi}{1-\alpha^2}a_i+\frac{\alpha\varphi}{1-\alpha^2}a_{m(i,t)}+\frac{\varphi\alpha^2}{(1-\alpha^2)(1-\alpha)}\xi_t,\quad\text{and}\quad y_t=\frac{\varphi\alpha^2}{(1-\alpha^2)(1-\alpha)}\xi_t.$$

In expectation, agents' believe the aggregate output follows

$$\mathbb{E}_{it}[y_t]=(f_2+f_3)\xi_t=\frac{\varphi\alpha^2}{(1-\alpha^2)(1-\alpha)}\xi_t=\frac{1}{\alpha}y_t,$$

which implies the forecast error is

$$y_{t+1}-\mathbb{E}_{it}[y_{t+1}]=\frac{\varphi\alpha^2}{(1-\alpha^2)(1-\alpha)}\left(\eta_{t+1}+\rho\xi_t-\frac{1}{\alpha}\rho\xi_t\right)=\frac{\varphi\alpha^2}{(1-\alpha^2)(1-\alpha)}\eta_{t+1}-\frac{\varphi\alpha}{1-\alpha^2}\rho\xi_t.$$

That is, agents over-estimate the aggregate output in response to a positive confidence shock.

A.8 Proof of Proposition 4.1

Denote the policy function for output and capital as

$$\begin{aligned} y_{it} &= h_a a_i + h_1(L) x_{it}^1 + h_2(L) x_{it}^2, \\ k_{i,t+1} &= g_a a_i + g_1(L) x_{it}^1 + g_2(L) x_{it}^2. \end{aligned}$$

It follows that the forecast about individual trading partner can be expressed as the forecast about the aggregate output

$$\begin{aligned} \mathbb{E}_{it}[y_{jt}] &= h_a \mathbb{E}_{it}[a_{jt}] + \mathbb{E}_{it}[y_t] = h_a(x_{it}^1 - \mathbb{E}_{it}[\xi_t]) + \mathbb{E}_{it}[y_t], \\ \mathbb{E}_{it}[y_{j,t+1}] &= \mathbb{E}_{it}[y_{t+1}], \end{aligned}$$

where the second equation is due to that matching is i.i.d. Given the scalar h_a, the best response becomes

$$\begin{bmatrix} y_{it} \\ k_{it+1} \end{bmatrix} = \begin{bmatrix} \varphi \\ 0 \end{bmatrix} a_i + \mathbf{\Psi}(L) \begin{bmatrix} y_{it} \\ k_{it+1} \end{bmatrix} + \mathbf{\Gamma}(L) y_t + [\mathbf{\Gamma}(L)]_+ h_a (x_{it}^1 - \mathbb{E}_{it}[\xi_t]).$$

The scalars h_a and g_a need to satisfy the following system

$$h_a = \varphi + \mathbf{\Gamma}_1(1) h_1(0) + \mathbf{\Psi}_{12}(1) g_a, \qquad g_a = \mathbf{\Gamma}_2(1) h_1(0) + \mathbf{\Psi}_{21}(1) h_a^y + \mathbf{\Psi}_{22}(1) g_a. \tag{A.2}$$

By the Wiener-Hopf prediction formula, the relevant forecasts can be written as

$$\begin{aligned} \mathbb{E}_{it}[y_t] &= \begin{bmatrix} \frac{\lambda\tau_\varepsilon}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} (h_1(L) + h_2(L)) - \frac{\lambda^2 \tau_\varepsilon}{\rho} \frac{1}{1-\rho\lambda} \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} (h_1(\lambda) + h_2(\lambda)) \\ \frac{\lambda\tau_u}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} (h_1(L) + h_2(L)) - \frac{\lambda^2 \tau_u}{\rho} \frac{1}{1-\rho\lambda} \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} (h_1(\lambda) + h_2(\lambda)) \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix} \\ \mathbb{E}_{it}[y_{t+1}] &= \begin{bmatrix} \frac{\lambda\tau_\varepsilon}{\rho} \frac{1}{(1-\lambda L)(L-\lambda)} (h_1(L) + h_2(L)) - \frac{\lambda \tau_\varepsilon}{\rho} \frac{1}{1-\rho\lambda} \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} (h_1(\lambda) + h_2(\lambda)) \\ \frac{\lambda\tau_u}{\rho} \frac{1}{(1-\lambda L)(L-\lambda)} (h_1(L) + h_2(L)) - \frac{\lambda \tau_u}{\rho} \frac{1}{1-\rho\lambda} \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} (h_1(\lambda) + h_2(\lambda)) \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix} \\ \mathbb{E}_{it}[y_{it+1}] &= \begin{bmatrix} \frac{h_1(L)}{L} - \frac{h_1(0)}{L} + (h_1(0) + h_2(0)) \frac{\lambda\tau_\varepsilon}{(1-\rho\lambda)(1-\lambda L)} \\ \frac{h_2(L)}{L} - \frac{h_2(0)}{L} + (h_1(0) + h_2(0)) \frac{\lambda\tau_u}{(1-\rho\lambda)(1-\lambda L)} \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix} \\ \mathbb{E}_{it}[k_{it+2}] &= \begin{bmatrix} \frac{g_1(L)}{L} - \frac{g_1(0)}{L} + (g_1(0) + g_2(0)) \frac{\lambda\tau_\varepsilon}{(1-\rho\lambda)(1-\lambda L)} \\ \frac{g_2(L)}{L} - \frac{g_2(0)}{L} + (g_1(0) + g_2(0)) \frac{\lambda\tau_u}{(1-\rho\lambda)(1-\lambda L)} \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix}. \end{aligned}$$

As a result, the policy functions need to satisfy

$$\mathbf{C}(L) \begin{bmatrix} h_1(L) & h_2(L) & g_1(L) & g_2(L) \end{bmatrix}' = \mathbf{D}(L),$$

where $\mathbf{C}(L)$ and $\mathbf{D}(L)$ are given by

$$\mathbf{C}(L) = \begin{bmatrix} 1 - \mathbf{\Gamma}_1(L) \frac{\lambda\tau_a}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & -\mathbf{\Gamma}_1(L) \frac{\lambda\tau_a}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & -\mathbf{\Psi}_{12}(L) & 0 \\ -\mathbf{\Gamma}_1(L) \frac{\lambda\tau_u}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & 1 - \mathbf{\Gamma}_1(L) \frac{\lambda\tau_u}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & 0 & -\mathbf{\Psi}_{12}(L) \\ -\mathbf{\Psi}_{21}(L) - \mathbf{\Gamma}_2 \frac{\lambda\tau_a}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & -\mathbf{\Gamma}_{21}(L) \frac{\lambda\tau_a}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & 1 - \mathbf{\Psi}_{22}(L) & 0 \\ -\mathbf{\Gamma}_2(L) \frac{\lambda\tau_u}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & -\mathbf{\Psi}_{21}(L) - \mathbf{\Gamma}_2(L) \frac{\lambda\tau_u}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & 0 & 1 - \mathbf{\Psi}_{22}(L) \end{bmatrix},$$

and

$$
\mathbf{D}(L) = \begin{bmatrix} \boldsymbol{\Gamma}_1(1) h_a \left(1 - \frac{\lambda\tau_a}{\rho(1-\lambda L)(1-\rho\lambda)}\right) - \boldsymbol{\Gamma}_1(1) q_1 \lambda \tau_a \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} \\ \boldsymbol{\Gamma}_1(1) h_a \left(-\frac{\lambda\tau_u}{\rho(1-\lambda L)(1-\rho\lambda)}\right) - \boldsymbol{\Gamma}_1(1) q_1 \lambda \tau_u \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} \\ [\boldsymbol{\Gamma}_2(L)]_+ h_a \left(1 - \frac{\lambda\tau_a}{\rho(1-\lambda L)(1-\rho\lambda)}\right) - \boldsymbol{\Gamma}_2(\lambda^{-1}) q_1 \tau_a \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} - f_1 \frac{1}{L} + q_2 \frac{\lambda\tau_a}{(1-\rho\lambda)(1-\lambda L)} \\ [\boldsymbol{\Gamma}_2(L)]_+ h_a \left(-\frac{\lambda\tau_u}{\rho(1-\lambda L)(1-\rho\lambda)}\right) - \boldsymbol{\Gamma}_2(\lambda^{-1}) q_1 \tau_u \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} - f_2 \frac{1}{L} + q_2 \frac{\lambda\tau_u}{(1-\rho\lambda)(1-\lambda L)} \end{bmatrix},
$$

with the related scalars defined as

$$
\begin{aligned}
f_1 &= (\boldsymbol{\Psi}_{21}(1) - [\boldsymbol{\Psi}_{21}(L)]_+(1)) h_1(0) + (\boldsymbol{\Psi}_{22}(1) - [\boldsymbol{\Psi}_{22}(L)]_+(1)) g_1(0), \\
f_2 &= (\boldsymbol{\Psi}_{21}(1) - [\boldsymbol{\Psi}_{21}(L)]_+(1)) h_2(0) + (\boldsymbol{\Psi}_{22}(1) - [\boldsymbol{\Psi}_{22}(L)]_+(1)) g_2(0), \\
q_1 &= h_1(\lambda) + h_2(\lambda), \\
q_2 &= f_1 + f_2.
\end{aligned}
$$

We have used the notation + to indicate the annihilation operator. For example, $[\boldsymbol{\Psi}_{21}(L)]_+ = -\frac{1-\omega}{\varkappa_3}$.

The scalar $h_1(0)$ needs to satisfy the condition that

$$
h_1(0) = \left. \frac{\det \begin{bmatrix} \mathbf{D}(L) & \mathbf{C}_2(L) & \mathbf{C}_3(L) & \mathbf{C}_4(L) \end{bmatrix}}{\det \mathbf{C}(L)} \right|_{L=0} = \Gamma_1(0) h_a + \frac{(h_1(\lambda) + h_2(\lambda) - h_a)\lambda\tau_a}{\rho(1-\lambda\rho)}. \tag{A.3}
$$

Given the policy functions, the aggregate outcome y_t and k_{t+1} can be expressed as

$$
\begin{aligned}
y_t &= (h_1(L) + h_2(L))\xi_t \\
k_{t+1} &= (g_1(L) + g_2(L))\xi_t.
\end{aligned}
$$

Utilizing the definition of $\mathbf{C}(L)$ and $\mathbf{D}(L)$ and defining $\tau \equiv \tau_a + \tau_u$, we have

$$
\mathbf{T}(L) \begin{bmatrix} h_1(L) + h_2(L) \end{bmatrix} = \boldsymbol{d}(L),
$$

where

$$
\mathbf{T}(L) = \begin{bmatrix} 1 - \boldsymbol{\Gamma}_1(L) \frac{\lambda\tau}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & -\boldsymbol{\Psi}_{12}(L) \\ -\boldsymbol{\Psi}_{21}(L) - \boldsymbol{\Gamma}_2 \frac{\lambda\tau}{\rho} \frac{L}{(1-\lambda L)(L-\lambda)} & 1 - \boldsymbol{\Psi}_{22}(L) \end{bmatrix} = \mathbf{I} - \boldsymbol{\Psi}(L) - \frac{L(1-\lambda\rho)(\rho-\lambda)}{\rho(1-\lambda L)(L-\lambda)} \begin{bmatrix} \boldsymbol{\Gamma}(L) & \mathbf{0} \end{bmatrix},
$$

and

$$
\boldsymbol{d}(L) = \begin{bmatrix} h_a \boldsymbol{\Gamma}_1(1) \left(1 - \frac{\lambda}{\rho(1-\lambda L)(1-\rho\lambda)}\right) - \boldsymbol{\Psi}_{12}(1) q_1 \lambda \tau \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} \\ h_a \boldsymbol{\Gamma}_1(1) \left(1 - \frac{\lambda\tau}{\rho(1-\lambda L)(1-\rho\lambda)}\right) - \boldsymbol{\Gamma}_2(\lambda^{-1}) h_1 \tau \frac{1-\rho L}{(1-\lambda L)(L-\lambda)} - q_2 \left(\frac{1}{L} - \frac{\lambda\tau}{(1-\rho\lambda)(1-\lambda L)}\right) \end{bmatrix}.
$$

The degree of the numerator of $\det[\mathbf{T}(L)]$ is 4, and there are two constants q_1 and q_2 to be determined. Denote the inverse of the outside roots of the determinant of $\mathbf{T}(L)$ as ϑ_1 and ϑ_2, and the inside roots as ζ_1 and ζ_2. The constants q_1 and q_2 have to be set to eliminate the poles at ζ_1 and ζ_2, which leads to

$$\det\begin{bmatrix}\mathbf{d}(\zeta_1) & \begin{bmatrix}-\mathbf{\Psi}_{12}(\zeta_1)\\ 1-\mathbf{\Psi}_{22}(\zeta_1)\end{bmatrix}\end{bmatrix}=0, \qquad \text{and} \qquad \det\begin{bmatrix}\mathbf{d}(\zeta_2) & \begin{bmatrix}-\mathbf{\Psi}_{12}(\zeta_2)\\ 1-\mathbf{\Psi}_{22}(\zeta_2)\end{bmatrix}\end{bmatrix}=0. \tag{A.4}$$

The constants $\{h_a, g_a, q_1, q_2, h_1(0)\}$ then can be obtained by solving the linear system (A.2), (A.3), and (A.4). By the Vieta's formula, $\zeta_1+\zeta_2$ and $\zeta_1\zeta_2$ can be substituted by ϑ_1 and ϑ_2, and $\{h_a, g_a, q_1, q_2\}$ can therefore be expressed as functions of $\{\vartheta_1, \vartheta_2\}$.

After obtaining these constants, the policy function $h_1(L)+h_2(L)$ and $g_1(L)+g_2(L)$ can be obtained accordingly by the Cramer's rule

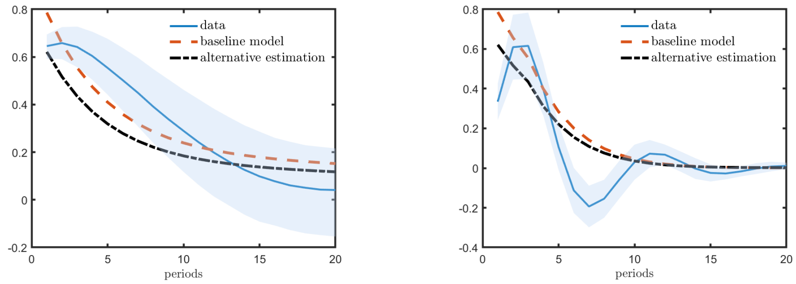
$$\begin{aligned}\frac{h_1(L)+h_2(L)}{1-\rho L} &= \mu_y\frac{1-r_yL}{(1-\vartheta_1L)(1-\vartheta_2L)},\\ \frac{g_1(L)+g_2(L)}{1-\rho L} &= \mu_k\frac{1-r_kL}{(1-\vartheta_1L)(1-\vartheta_2L)(1-\lambda L)},\end{aligned}$$

where

$$\begin{aligned}\mu_y &= \frac{\mathbf{\Gamma}_1(1)(h_a\lambda+q_1(\rho-\lambda))}{\rho},\\ \mu_k &= \frac{q_2+\mathbf{\Gamma}_1(1)(\mathbf{\Psi}_{21}(1)-[\mathbf{\Psi}_{21}(L)]_+(1))\left((q_1-h_a)\frac{\lambda}{\rho}-q_1\right)}{\mathbf{\Psi}_{22}(1)-[\mathbf{\Psi}_{22}(L)]_+(1)},\\ r_y &= -\frac{h_a\vartheta_1\vartheta_2\left(\mathbf{\Psi}_{12}(1)(\mathbf{\Gamma}_2(1)-[\mathbf{\Gamma}_2(L)]_+)-\mathbf{\Gamma}_1(1)(\mathbf{\Psi}_{22}(1)-[\mathbf{\Gamma}_{22}(L)]_+)\right)}{\mu_y\rho\left((\mathbf{\Psi}_{22}(1)-[\mathbf{\Gamma}_{22}(L)]_+)+\Psi_{12}(1)(\mathbf{\Psi}_{21}(1)-[\mathbf{\Gamma}_{21}(L)]_+)\right)},\\ r_k &= -\frac{h_a\lambda\vartheta_1\vartheta_2\left((\mathbf{\Gamma}_2(1)-[\mathbf{\Gamma}_2(L)]_+)+\mathbf{\Psi}_{12}(1)(\mathbf{\Psi}_{21}(1)-[\mathbf{\Gamma}_{21}(L)]_+)\right)}{\mu_k\rho\left((\mathbf{\Psi}_{22}(1)-[\mathbf{\Gamma}_{22}(L)]_+)+\Psi_{12}(1)(\mathbf{\Psi}_{21}(1)-[\mathbf{\Gamma}_{21}(L)]_+)\right)}.\end{aligned}$$

B. Supplementary Materials for Section 4

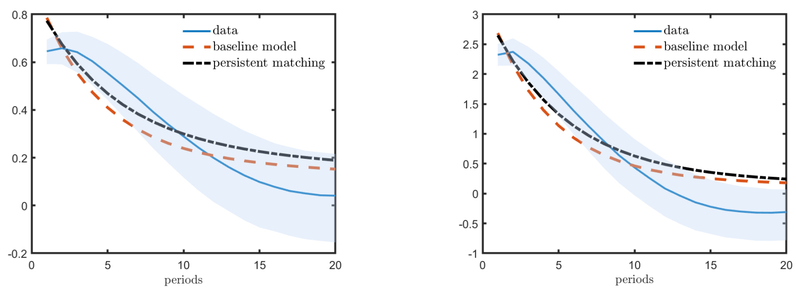
B.1 Additional Tables and Figures



(a) output (b) output forecast error

Figure A1: IRFs: Matching Different Forecasting Horizons

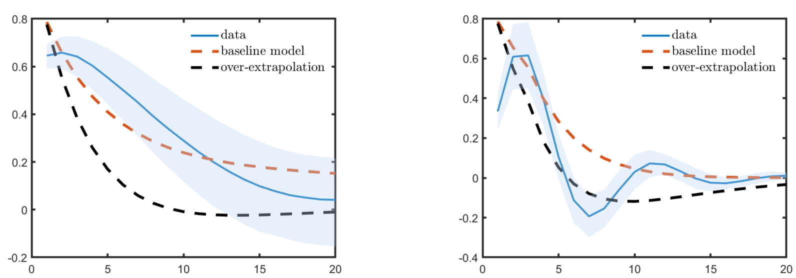
Notes: The red-dashed lines correspond to the baseline estimation which matches 3-step ahead forecast error. The black broken lines correspond to the estimation which matches the average of 1-step to 3-step ahead forecast errors.



(a) output (b) investment

Figure A2: IRFs: Persistent Matching

Notes: The red-dashed lines correspond to the baseline model with i.i.d matching process. The black broken lines correspond to results with persistent matching process.



(a) output (b) output forecast error

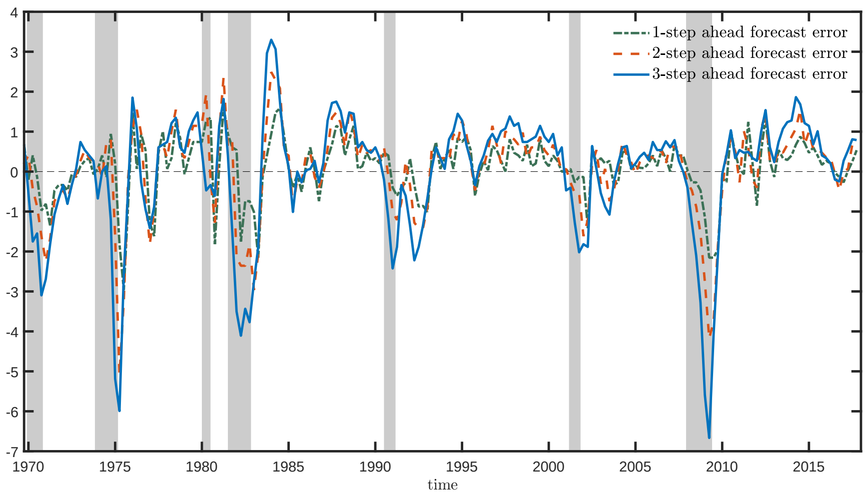
Figure A3: IRFs: Over-Extrapolation

Notes: The red-dashed lines correspond to the baseline parameterization. The black broken lines correspond to results with over-extrapolation of the confidence shock process.

Table A1: Business Cycle Statistics in the Model with Endogenous TFP

	Standard deviation			Corr w/ output			Auto-correlation		
	data	baseline	endo TFP	data	baseline	endo TFP	data	baseline	endo TFP
Y	1.48	0.94	0.98	1.00	1.00	1.00	0.85	0.63	0.64
C	5.63	3.21	4.04	0.93	0.99	0.99	0.86	0.62	0.64
I	0.81	0.19	0.74	0.82	0.79	0.99	0.88	0.81	0.64
N	1.43	2.80	1.01	0.84	0.98	0.99	0.86	0.62	0.64
$\mathcal{M}$	3.00	6.10	1.85	-0.72	-0.97	-0.99	0.84	0.62	0.64

Figure A4: Forecast Errors with Different Horizons



Notes: This figure displays the rescaled forecast errors of unemployment rate. The 3-step ahead forecast error is used in the baseline estimation, and the average of the three series is used in estimation in subsection 4.5.

B.2 Model with Goods Market Frictions

In this part, we describe the model economy where endogenous TFP arises due to goods market frictions.

In the second stage, shoppers serve both as buyers and sellers. As sellers, each shopper is endowed with a unit measure of location and they can choose in which market to sell the goods inherited from their producers. As buyers, shoppers have to consume the goods produced by others but not by themselves, similarly to Trejos and Wright (1995). Goods market frictions require buyers to exert search effort to find the locations of others.

Different markets are indexed by their price and market tightness (P, Q), where market tightness is defined as the ratio of the measure of location to the measure of search effort. Exerting one unit of search effort in market (P, Q), a buyer expects to find a location with probability $\Psi^d(Q)$ at price P. At the same time, a seller in in market (P, Q) expects to sell her goods with probability $\Psi^f(Q)$ at price P. In equilibrium, not all markets are active. In fact, it is understood that there is an equilibrium-determined expected revenue per unit of good, $\zeta = P\ \Psi^f(Q)$, that active markets have to satisfy.

Because there are two different types of goods, local goods Y_i and foreign goods Y_j, there are two equilibrium-determined expected revenues ζ_i and ζ_j. Buyers on island i choose the local market (P_{ii}, Q_{ii}) and foreign markets (P_{ij}, Q_{ij}), while buyers on island j choose (P_{jj}, Q_{jj}) and (P_{ji}, Q_{ji}). In equilibrium, sellers have to be indifferent between allocating their locations to domestic customers and foreign customers, resulting in

$$
\begin{aligned}
P_{ii}\Psi^f(Q_{ii}) &= P_{ji}\Psi^f(Q_{ji}) = \zeta_i, \\
P_{jj}\Psi^f(Q_{jj}) &= P_{ij}\Psi^f(Q_{ij}) = \zeta_j.
\end{aligned}
$$

We assume that the matching function in the goods market is of Cobb-Douglas form

$$
\Psi^d(Q) = \nu Q^{1-\kappa}, \qquad \Psi^f(Q) = \nu Q^{-\kappa},
$$

where κ is the matching elasticity and ν is a constant that determines the average matching probability.

It is important to note that not all goods can be sold and the produced goods Y_i and Y_j are only potential output. The actual output depends on the probability Ψ^f that goods are purchased, which is determined by the amount of search effort. This probability Ψ^f can be understood as the utilization rate, and we will show that it increases with the production level of Y_i and Y_j. When the production level changes, the amount of search effort and the utilization rate also change, generating endogenous movements of the measured Solow residual.

The shoppers' problem on island i can be written as

$$
\max_{\substack{C_{ii}, C_{ij}, I_{ii}, I_{ij}, \\ Q_{ii}, Q_{ij}, D_{ii}, D_{ij}}} \left(\frac{C_{ii}}{\omega}\right)^{\omega} \left(\frac{C_{ij}}{1-\omega}\right)^{1-\omega} - \chi_d D_i
$$

subject to

$$
\begin{aligned}
P_{ii}(C_{ii} + I_{ii}) + P_{ij}(C_{ij} + I_{ij}) &= \zeta_i Y_i, \\
C_{ii} + I_{ii} &= D_{ii}\Psi^d(Q_{ii})Y_i, \\
C_{ij} + I_{ij} &= D_{ij}\Psi^d(Q_{ij})Y_j, \\
P_{ii}\Psi^f(Q_{ii}) &= \zeta_i, \\
P_{ij}\Psi^f(Q_{ij}) &= \zeta_j, \\
I_i &= \left(\frac{I_{ii}}{\omega}\right)^{\omega} \left(\frac{I_{ij}}{1-\omega}\right)^{1-\omega}, \\
D_i &= D_{ii} + D_{ij}.
\end{aligned}
$$

The equilibrium conditions include

$$
\begin{gathered}
Q_{ii} = \frac{T_{ii}}{D_{ii}}, \;\; Q_{ij} = \frac{T_{ji}}{D_{ij}}, \;\; Q_{ji} = \frac{T_{ij}}{D_{ji}}, \;\; Q_{jj} = \frac{T_{jj}}{D_{jj}}, \\
T_{ii} + T_{ij} = 1, \quad T_{ji} + T_{jj} = 1.
\end{gathered}
$$

The equilibrium allocation satisfies

$$D_i^* = \left(\frac{\mu \nu}{\chi_d} Y_i^{\omega} Y_j^{1-\omega} \right)^{\frac{1}{1-\mu}}.$$

The linearized version of aggregate search effort is $d_t = \frac{1}{1-\kappa} y_t$. The measured Solow residual is the average probability a goods is sold, which is increasing in the total search effort

$$z_t = \kappa d_t = \frac{\kappa}{1-\kappa} y_t.$$

Table A1 shows the business cycle moments for the model with goods market frictions.