Higher Order Beliefs, Confidence, and Business Cycles*

Zhen Huo
University of Minnesota

Naoki Takayama
Cabinet Office, Government of Japan

Please check http://zhenhuo.weebly.com/research.html for the newest version.

Abstract

This paper presents a model of business cycles driven by shocks to agents' beliefs about economic fundamentals. Agents are hit both by common and idiosyncratic shocks. Common shocks act as confidence shocks, which cause economy-wide optimism or pessimism and consequently, aggregate fluctuations in real variables. Idiosyncratic shocks generate dispersed information, which prevents agents from perfectly inferring the state of the economy. Crucially, asymmetric information induces the infinite regress problem, that is, agents need to forecast the forecasts of others. We develop a method that can solve the infinite regress problem without approximation. Even though agents face a complicated learning problem, the equilibrium policy can be represented by a small number of state variables. Theoretically, we prove that the persistence of aggregate output is increasing in the degree of information frictions and strategic complementarity, and there is a hump-shaped relationship between the variance of output and the variance of the confidence shock. Quantitatively, our model with confidence shocks can explain a number of the key business cycle moments.

Keywords: Higher order beliefs, Infinite regress problem, dispersed information, business cycles, confidence.
JEL classifications: E20, E32, F44

*We are thankful for our advisers José-Víctor Ríos-Rull, Ellen McGrattan, Jonathan Heathcote, and Elena Pastorino for inspiration, encouragement, and support. We are especially grateful for comments from Anmol Bhandari, V.V. Chari, Patrick Kehoe, Fabrizio Perri, Thomas Sargent, and Kei-Mu Yi. We also thank Manuel Amador, Yan Bai, Luigi Bocola, Katarina Borovickova, Jaroslav Borovicka, Fend Dong, Si Guo, Boyan Jovanovic, Kyle Herkenhoff, Guido Lorenzoni, Shihui Ma, Virgiliu Midrigan, Alessandro Peri, Edouard Schaal, Kjetil Storesletten, Gianluca Violante, Chunying Xie, and the participants of the Labor Workshop at the University of Minnesota, Federal Reserve Bank of Minneapolis, and New York University for helpful comments. Zhen Huo thanks the University of Minnesota Doctoral Dissertation Fellowship for financial support.

1 Introduction

Motivated by the Great Recession, there has been an increased interest in business cycles driven by confidence shocks (Lorenzoni, 2009; Angeletos and La'O, 2013; Benhabib, Wang, and Wen, 2014). A confidence shock can be understood as a shock to agents' beliefs about the economic activities that others are capable of. When this shock is correlated across agents, it induces economy-wide optimism or pessimism, and therefore, aggregate fluctuations in the main macro variables. Intuitively, confidence is promising as a source of business cycle fluctuations since it is well known that people's perceptions of business conditions vary dramatically. However, there have been substantial difficulties to incorporate confidence shocks into a rational expectations framework because of the *infinite regress problem* (Townsend, 1983). Namely, with asymmetric information and interconnection between agents' economic activities, agents' payoffs depend on their beliefs about others' actions, and rationality requires agents to forecast the forecast of others. While it is necessary to allow for some persistence in shocks for empirical relevance, rational agents have to keep all the information learned from the past to forecast all higher order beliefs, which leads to an infinite-dimensional state space. The goal of this paper is to overcome this technical difficulty, and to explore whether the confidence shock could be an important factor in accounting for business cycles.

Our first contribution is to solve the infinite regress problem by applying our method developed in Huo and Takayama (2014). It is widely believed that if a rational expectations model involves higher order beliefs and persistent hidden states, the Kalman filter has to be applied to solve the signal extraction problem and to keep track of an infinite number of state variables in order to forecast all higher order beliefs. To short-circuit this problem, the existing literature typically assumes that the information become public after a certain number of periods, or imposes a heterogeneous prior formulation. Instead of modifying the original problem, we confront and solve the infinite regress problem directly. We prove that for any linear rational expectations model with an ARMA signal process, the equilibrium policy rule always allows a finite-state-variable representation.[1] We also provide a procedure to find these state variables and their laws of motion. By using a small set of state variables, agents can perform their best inference in equilibrium, and economists can calibrate or estimate the model as standard DSGE models with perfect information.

The idea is to find the true solution in the space spanned by the entire history of signals in the first place. In this infinite-dimensional state space, we use the Wiener filter to handle the signal extraction problem, as opposed to the standard Kalman filter. It turns out that if the signal process follows an ARMA process, the equilibrium policy will inherit this property and also be of the ARMA type. This implies that information can be summarized in a relatively compact way, and it allows

[1]The linearity may be obtained by log-linearization, and the ARMA process assumption is compatible with the shock structure specified in most macroeconomic models.

us to find a finite-state-variable representation of the equilibrium policy rule. In addition, after we find this representation, the equilibrium is characterized by a simple linear system, and we no longer need to solve any inference problems when simulating the economy.

Our second contribution is to formalize the idea of confidence shocks in a rational expectations model and to apply our method to evaluate its quantitative importance. We first construct an illustrating model with decentralized trading and information frictions, which is based on the structure specified in Angeletos and La'O (2013). The economy consists of a continuum of islands, and the islands differ in their productivity. At every period, each island is randomly matched with another island and trades with it. Households value both domestic and foreign goods, resulting in the local output increasing in their trading partner's output. Information frictions prevent households from observing their trading partner's productivity, and households only receive a noisy signal of this productivity. With a positive (negative) noise, islands tend to overestimate (underestimate) their trading partners' productivity and output, and also to increase their own output due to strategic complementarity. If the noise shock is correlated across islands, then it will cause economy-wide output fluctuations. We label this shock a confidence shock.

When choosing the production level, agents need to infer their trading partner's productivity level, which is equivalent to inferring the confidence shock. However, this is not the end of the inference problem. Note that different islands receive different signals over time, and they will form different inferences about this confidence shock. As a result, agents also need to infer their trading partners' inference of the confidence shock, and all other higher order beliefs. If the confidence shock is persistent, the entire history of signals should be recorded since these signals contain information about the current state of the economy. Even though this is a fairly complicated learning problem, we manage to obtain a sharp analytic solution.

This model economy has two important properties. First, under the assumption that the confidence shock follows an AR(1) process, the aggregate output also follows an AR(1) process. Interestingly, the persistence of the aggregate output is increasing in the degree of strategic complementarity, the value of which is a function of the deep parameters related to preferences and technology. With a stronger interdependence, households respond more aggressively to signals, which magnifies the effects of the confidence shock. The persistence of the aggregate output is also increasing in the degree of information frictions, as it is more difficult to separate the confidence shock from a true productivity shock. Secondly, the unconditional variance of the aggregate output is not monotonically increasing in the variance of the confidence shock. On the one hand, if the variance of the confidence shock is small, the variance of aggregate output is also small since confidence shocks are the only exogenous disturbances. On the other hand, if the variance of the confidence shock is large, agents understand that signals become less useful for information extraction, and they optimally respond less to them. These two competing forces result in a hump-shaped relationship between the variance of output and the variance of the confidence shock. This nonlinearity is absent in standard DSGE models without information friction.

Another important property is that the forecast error is persistent. Supposing the forecast error is absent or there is no information friction, the equilibrium allocation is uniquely pinned down by economic fundamentals, leaving no room for the confidence shock. If we aim to generate persistent aggregate fluctuations, it is important to make sure that the forecast error is long-lasting. In our model, the forecast error is indeed persistent, and agents can never perfectly infer the underlying shocks. This is the result of our information structure, in which there are more shocks than signals, and agents do not have enough information to recover the true state of the economy. By contrast, in Kasa (2000) and Acharya (2013), the number of shocks equals the number of signals, and the forecast error disappears quickly. To ensure the persistent effects of the confidence shock, the information process has to be complicated enough to confuse agents for a relatively long time.

With these insights, we develop a quantitative business cycle model to examine the empirical relevance of the confidence shock. Our quantitative model has three key features: a rich information process, goods market frictions, and endogenous capital accumulation. (1) The rich information process provides the flexibility to pin down the degree of information frictions, which is the key factor in determining the performance of the model. The rational expectations framework allows us to link the signal extraction problem faced by agents in the model with the micro-level data. We set the variance and persistence of noise shocks to match the GDP forecast error in the Survey of Professional Forecasters. (2) Introducing goods market frictions a la Bai, Ríos-Rull, and Storesletten (2011) helps generate endogenous movements of the Solow residual. Goods market frictions create a wedge between potential and realized output. As consumers increase their demand, the utilization rate of potential output also increases, translating into a higher Solow residual. Without the endogenous Solow residual, employment becomes the only driving force of output in the short run, and it leads to the counter-factual prediction that the volatility of employment is much greater than that of output. (3) Capital accumulation brings additional endogenous persistence into the model economy. It also increases the complexity of the signal extraction problem substantially, which prevents us from obtaining an analytic solution. However, we can still represent the equilibrium policy rule by a small number of state variables.

In terms of quantitative performance, we find that the confidence shocks alone can account for much business cycle movement and co-movement. The persistence of main aggregate variables is endogenously determined, which represents from 50% to 80% of their data counterpart under our calibrations of information frictions. The persistence of aggregate variables hinges on the persistence of forecast errors, which are only modestly persistent in the data. This moment, the persistence of forecast errors, imposes an upper bound on the degree of information frictions, and it prevents generating large persistence of aggregate variables in our model with confidence shocks. Compared with standard RBC models driven by TFP shocks, two differences stand out. First, our model driven by confidence shocks generates strong counter-cyclical labor wedges, a moment emphasized by Chari, Kehoe, and McGrattan (2007). Secondly, with confidence shocks, the standard deviation of employment is 60% greater than in the RBC models, and it is much closer to the data.

Related literature From a methodological point of view, our paper is related to the literature that attempts to solve models with higher order beliefs. The most widely used method is truncating the relevant state by assuming all shocks become public information after a finite time or only a finite number of higher order beliefs matter for the equilibrium. With a finite number of state variables, the standard Kalman filter can be applied. This line of literature includes Townsend (1983), Hellwig and Venkateswaran (2009), Lorenzoni (2009), Bacchetta and Wincoop (2006), and Nimark (2008) among others. Using these methods to solve our quantitative model with endogenous capital, the number of state variables needed is fairly large to achieve reasonable accuracy, and it is even more difficult to conduct calibration or estimation. The method we developed in Huo and Takayama (2014) provides the true solution to the model, and it only requires a small number of state variables, which makes calibration or estimation possible. Kasa (2000) and Acharya (2013) also solve models with higher order beliefs without truncation, but in their environment, the number of signals is the same as the number of shocks, and the forecast error is not persistent. Our method allows us to use a general signal process when there are more shocks than signals, and the confidence shock has persistent effects.[2]

Angeletos, Collard, and Dellas (2014) assume agents have heterogeneous prior. This assumption avoids the difficult infinite regress problem, but as acknowledged by the authors, it also abstracts from agents' information extraction process. Under the common prior assumption, our method does not increase the computational difficulty, but allows us to link the model with micro-data and to pin down the degree of information frictions. The cross-sectional evidence on belief dispersion and forecast errors imposes an upper bound on the persistence and volatility of output that can be generated by confidence shocks.

Our quantitative application also complements the literature on aggregate fluctuations driven by shocks to agents' beliefs. In Lorenzoni (2009), Angeletos and La'O (2010), and Blanchard, L'Huillier, and Lorenzoni (2013), there is a shock to aggregate TFP, but agents only observe aggregate TFP contaminated by common noise. Even though this common noise can generate aggregate fluctuations, its effects are bounded above by the variance of the TFP shock. As the variance of the TFP shock approaches zero, agents will not respond to the noise shock. Angeletos and La'O (2013) introduce

[2] In Rondina and Walker (2013), the number of shocks is the same as the number of signals, but they assume that the underlying shock process is not invertible, which leads to persistent forecast error. We think it is more natural to introduce persistent forecast error by allowing more shocks than signals, a feature that is prevalent in signal extrication problems.

additional trading and communication frictions, and as a result, common noise can generate aggregate fluctuations with aggregate fundamentals being fixed. Our model environment is similar to Angeletos and La'O (2013), but we allow persistent common noise. Also, we highlight the role of higher order beliefs in shaping aggregate output. Benhabib, Wang, and Wen (2014) propose another type of environment in which sentiments can generate aggregate fluctuations without resorting to trading or information frictions, and the variance of sentiment shocks is endogenously determined. Unlike our model, agents do not need to solve the infinite regress problem. Our paper is also related to the literature on news shocks, such as Beaudry and Portier (2006), Jaimovich and Rebelo (2009), Barsky and Sims (2012), Schmitt-Grohé and Uribe (2012), and Beaudry, Nam, and Wang (2011) among others.

The rest of the paper is organized as follows. Section 2 sets up a simple economy and describes how the infinite regress problem arises in this environment. We obtain an analytic solution, and discuss various properties of this economy. Section 3 considers the case when agents observe the signal which contains endogenous information. We compare the equilibrium outcome with and without endogenous information. Section 4 explores the quantitative performance of a full-blown model with confidence shocks. Section 5 concludes.

2 An Analytic Model with Higher Order Beliefs

In this section, we present a simple island model to introduce confidence shocks which trigger 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 this model empirically relevant, but it induces the infinite regress problem which is difficult to solve. We apply the method developed in Huo and Takayama (2014) to solve the model and obtain a sharp analytic solution.

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 $N(0, \sigma_a^2)$ but fixed over time. Each island is populated by a continuum of identical households. In each household, there is a head and a shopper. The head decides how much to produce. The shopper then receives the output from the head 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, sometimes we will use j to denote $m(i, t)$ as the index of island i's contemporary trading partner to simplify notation. It should be clear that island

i is matched and trades with a different island j at each period.

We assume that the production plan has to be made at the beginning of a period without perfect knowledge of their trading partner's productivity level. The heads receive noisy signals about $a_{m(i,t)}$ (which will be specified below), and choose their output level conditional on these signals. After production, the two islands matched trade with each other.

The average productivity in the economy is fixed over time, but island i's specific trading partner changes every period. Even though households in each island know that there is no aggregate change of fundamentals, they still face uncertainty due to the decentralized trading arrangement and the communication frictions. The need to infer their trading partner's output and the lack of perfect information leaves room for confidence shocks and also for higher order beliefs.

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

$$x_{it}^1 = a_{m(i,t)} + \xi_t, \tag{2.1}$$

where $a_{m(i,t)} \sim N(0, \sigma_a^2)$. Crucially, we assume that common noise ξ_t follows a persistent process

$$\xi_t = \rho \xi_t + \eta_t, \tag{2.2}$$

where $\rho \in (0, 1)$ and $\eta_t \sim N(0, \sigma_\eta^2)$. A positive (negative) realization of ξ_t makes all agents in the economy overestimate (underestimate) 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 $u_{it} \sim N(0, \sigma_u^2)$ is idiosyncratic noise. The variance of u_{it} determines the degree of information friction in the economy. If $\sigma_u = 0$, then the heads observe ξ_t perfectly, and can figure out their trading partner's productivity using the first signal without error. The learning problem is trivial in this scenario. If $\sigma_u > 0$ but $\rho = 0$, the head faces a static learning problem, because the information is independent of previous periods. If $\sigma_u > 0$ and $\rho > 0$, the head faces a persistent learning problem, which is the focus of this paper.

The household head's information set on island i at time t is all the signals received up to time t

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

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, heads on different islands share heterogeneous information sets. It follows that $\mathbb{E}_{it}[\cdot] \neq \mathbb{E}_{jt}[\cdot]$. After observing the signals, the heads decide the output level Y_{it}, which completes the first stage of a period.

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

Remark The assumption that shoppers die after they trade and consume is only a means to implement the idea that the communication between shoppers and heads is not perfect. Supposing we allow imperfect communication between shoppers and heads, heads will receive another noisy signal on $a_{m(i,t)}$ or ξ_t, but this is equivalent to setting the variance of u_{it} to a smaller value. Therefore, what is really important is how much heads can learn, but not exactly how they learn.

Shoppers' Problem In the transaction stage, goods markets are competitive and the prices for local goods and foreign goods are P_i and P_j respectively.[3] Shoppers receive the output Y_i produced in the first period by heads. The shoppers' problem on island i solves the following static problem

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

subject to

$$P_i C_{ii} + P_j C_{ij} = P_i Y_i,$$

where C_{ii} is local consumption goods and C_{ij} is foreign consumption goods. We adopt a Cobb-Douglas preference structure and use ω to denote the degree of home bias. The first order condition for the shoppers' problem is

$$\frac{C_{ii}}{C_{ij}} = \frac{\omega}{1-\omega}\frac{P_j}{P_i},$$

[3] Because shoppers solve a static problem in the second stage, we use j to denote $m(i,t)$ to simplify the notation.

The goods market clearing condition in equilibrium is

$$\begin{aligned} C_{ii} + C_{ji} &= Y_i, \\ C_{ij} + C_{jj} &= Y_j. \end{aligned}$$

Combining the equilibrium condition and the first order condition for both islands, we have

$$\begin{aligned} C_{ii}^* &= \omega Y_i, \\ C_{ij}^* &= (1-\omega) Y_j. \end{aligned}$$

In equilibrium, local and foreign consumption are equal to a fixed fraction of local and foreign output, thanks to the Cobb-Douglas preference. In addition, we can derive the shadow value of 1 additional unit of local goods for shoppers:

$$\mathcal{U}_i = \left(\frac{C_{ij}}{C_{ii}} \frac{\omega}{1-\omega} \right)^{1-\omega} = \left(\frac{P_i}{P_j} \right)^{1-\omega} = \left(\frac{Y_j}{Y_i} \right)^{1-\omega}. \tag{2.5}$$

The shadow value is the marignal utility of 1 addtional local consumption, and it is clearly increasing in the terms of trade and consequently, increasing in foreign output.

Heads' Problem Household heads chooses how much to produce. They understand that in the second period, shoppers' utility is given by equation (2.5), which depends on not only their own output, but also on their trading partners' output. If there is no information friction ($\sigma_u = 0$), the productivities on both islands becomes common knowledge, and the output level on both islands will only be a function of the fundamentals. When there are information frictions, the output level on island i is determined by the expected output level on island $m(i,t)$.

The heads' problem on island i is choosing a state contingent plan for Y_{it} and labor N_{it} to maximize their expected present value

$$\max_{Y_{it}, N_{it}} \mathbb{E}_{i0} \sum_{t=0}^{\infty} \beta^t \frac{1}{1-\sigma} \left[\left(\frac{P_{it}}{P_{m(i,t)t}} \right)^{1-\omega} Y_{it} - N_{it}^{1+\gamma} \right]^{1-\sigma}$$

subject to

$$Y_{it} = \exp(a_i)\, N_{it}^{\theta}.$$

In the utility function, γ is the inverse of Frisch elasticity, θ determines the labor share, and σ is the risk aversion.[4] Heads assess the value of local output through the shadow value derived in equation

[4]We assume a GHH type utility function to eliminate wealth effects. With the general utility function, when

(2.5). The heads' optimal choice is equating the marginal utility of local output for the shoppers with the marginal disutility of producing the output. When expected $Y_{m(i,t)t}$ increases, the terms of trade improves and the marginal utility of local output also increases, which encourages heads on island i to produce more output. In this sense, there is strategic complementarity between local and foreign output. The first order condition is [5]

$$Y_{it} = \left(\frac{\theta}{1+\gamma}\right)^{\frac{1}{\frac{1+\gamma}{\theta}-\omega}} \exp\left(\frac{1}{1-\frac{\theta}{1+\gamma}\omega}\, a_i\right) \mathbb{E}_{it}[Y_{m(i,t)t}^{1-\omega}]^{\frac{1}{\frac{1+\gamma}{\theta}-\omega}}. \tag{2.6}$$

Standard parametrization ensures that $\gamma > 0$, $\theta \in (0,1)$, and $\omega \in (0,1)$. This implies that $\frac{1}{\frac{1+\gamma}{\theta}-\omega}$, and that the local output is increasing in the expected output $Y_{m(i,t)t}$.

Log-Linearized Economy In this paper, we will work with log-linearzied model. Throughout, we use small letters to denote the log deviation from a variable's steady state value. The log-linearized version of the head's decision rule (2.6) is

$$y_{it} = \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}[y_{m(i,t)t}], \tag{2.7}$$

where

$$\begin{aligned} \alpha_0 &= \frac{1}{1-\frac{\theta}{1+\gamma}\omega}, \\ \alpha_1 &= \frac{1-\omega}{\frac{1+\gamma}{\theta}-\omega}. \end{aligned}$$

As discussed before, α_1 is positive, and y_{it} is increasing in $\mathbb{E}_{it}[y_{m(i,t)t}]$. To guarantee a stable solution, we also restrict our parameter values such that $\alpha_1 < 1$. From now on, we will focus on equation (2.7). Note that the deep parameters related to preferences and technologies are all summarized by α_0 and α_1.

Perfect Information Benchmark Supposing the variance of the idiosyncratic noise u_{it} vanishes, then agents on island i can use the two signals to figure out $a_{m(i,t)}$ and ξ_t perfectly. In this case, there is no information friction. The optimal policy rule (2.7) becomes

$$y_{it} = \alpha_0 a_i + \alpha_1 y_{m(i,t)t}. \tag{2.8}$$

household heads on island i expect $Y_{m(i,t)}$ to be large, they will reduce their labor, which is the opposite of what we want.

[5] In the first order condition, we have already used the equilibrium condition that the individual output choice coincides with the aggregate output level due to the representative agent assumption.

As expected, the output on island i is completely determined by the economic fundamentals

$$y_{it} = \frac{\alpha_0}{1-\alpha_1^2} a_i + \frac{\alpha_0 \alpha_1}{1-\alpha_1^2} a_{m(i,t)}. \tag{2.9}$$

By the law of large number, the aggregate output y_t stays at its steady state

$$y_t = \int y_{it} = 0. \tag{2.10}$$

The confidence shock ξ_t has no effect at all.

2.2 Infinite Regress Problem

When there are information frictions, agents have to infer their trading partners' productivity and output. Higher order beliefs become crucial in determining the production level. By equation (2.7), to infer the output on island $m(i,t)$, island i has to infer the productivity on island $m(i,t)$, which relies on i's prediction of the confidence shock ξ_t. But the same logic also applies to island $m(i,t)$. Therefore, island i needs to infer island $m(i,t)$'s prediction of ξ_t. But so does island $m(i,t)$. It turns out that island i has to predict $m(i,t)$'s prediction of i's prediction of ξ_t, and all other higher order beliefs eventually.

Proposition 2.1. *When $\alpha_1 \in (0,1)$, the optimal output rule is given by*

$$y_{it} = \frac{\alpha_0}{1-\alpha_1^2} a_i + \frac{\alpha_0 \alpha_1}{1-\alpha_1^2} a_{m(i,t)} + \frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k (\xi_t - \mathbb{E}_{it}^k[\xi_t]) \tag{2.11}$$

where

$$\begin{aligned}
\mathbb{E}_{it}^1[\xi_t] &= \mathbb{E}_{it}[\xi_t] \\
\mathbb{E}_{it}^2[\xi_t] &= \mathbb{E}_{it}\mathbb{E}_{m(i,t)t}[\xi_t] \\
\mathbb{E}_{it}^k[\xi_t] &= \mathbb{E}_{it}\mathbb{E}_{m(i,t)t}\mathbb{E}_{it}^{k-2}[\xi_t], \ \textit{for } k = 3, 4, 5 \ldots
\end{aligned}$$

Proof. See Appendix A.1 for the proof. □

Because islands differ in their information sets, the law of iterated expectation does not apply. Confidence shocks have real effects on the economy. More specifically, the effects of confidence shock on island i are captured by the last term of equation (2.11)

$$\frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k (\xi_t - \mathbb{E}_{it}^k[\xi_t]), \tag{2.12}$$

and the aggregate output is

$$y_t = \frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k \left(\xi_t - \int \mathbb{E}_{it}^k[\xi_t] \right). \tag{2.13}$$

Note that the higher order beliefs $\mathbb{E}_{it}^k[\xi_t]$ for $k = \{1, 2, \ldots\}$ are different from ξ_t itself in general, which is the reason why the confidence shock can trigger aggregate fluctuation. If ξ_t is underestimated, then islands tend to overestimate their trading partners' productivities. By strategic complementarity, all the islands increase their own output because they expect a higher output from their trading partners, and a boom occurs.

The difficulty lies in computing the equilibrium policy rule of y_{it}. By Proposition 2.1, y_{it} depends on all the higher order beliefs $\mathbb{E}_{it}^k[\xi_t]$, but computing all the higher order beliefs is a fairly complicated task. The number of state variables needed to infer higher order beliefs is increasing in the order of the belief.

Proposition 2.2. *Given the signal process (2.1) to (2.3), the forecast of* $\left\{ \mathbb{E}_{jt}^1[\xi_t], \mathbb{E}_{jt}^2[\xi_t], \ldots, \mathbb{E}_{jt}^k[\xi_t] \right\}$ *requires* $k+1$ *state variables.*

The state variables in this proposition are the priors of these higher order beliefs. To spell out all the higher order beliefs, island i needs to keep track of an infinite number of state variables, which is the infinite regress problem. In the next section, we define the equilibrium and use the method developed in Huo and Takayama (2014) to solve the infinite regress problem. It turns out that the geometric sum of all higher order beliefs follows a simple ARMA process, and a finite number of state variables is sufficient for agents to choose the optimal output y_{it}.

2.3 Equilibrium

The information set of heads on island i is $\Omega_{it} = (a_i, \{x_{it-\tau}^1\}_{\tau=0}^{\infty}, \{x_{it-\tau}^2\}_{\tau=0}^{\infty})$. Therefore, island i's policy rule belongs to the space spanned by square-summable linear combinations of current and past realizations of x_{it}^1, x_{it}^2, and also by the time independent local productivity a_i

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

where $h_a \in \mathbb{R}$, $h_1(L)$ and $h_2(L)$ are lag polynomials

$$h_1(L) = \sum_{\tau=0}^{\infty} h_{1\tau} L^{\tau},$$
$$h_2(L) = \sum_{\tau=0}^{\infty} h_{2\tau} L^{\tau}.$$

The infinite sequences $\{h_{1\tau}\}_{\tau=0}^{\infty}$ and $\{h_{2\tau}\}_{\tau=0}^{\infty}$ belong to the square-summable space ℓ^2, which guarantees that y_{it} is a covariance-stationary process. The equilibrium is defined as follows

Definition 2.1. *Given the signal process (2.1) to (2.3), the equilibrium of model (2.7) is a policy rule $h = \{h_a, h_1, h_2\} \in \mathbb{R} \times \ell^2 \times \ell^2$, such that*

$$y_{it} = \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}[y_{m(i,t)t}],$$

where

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

The equilibrium policy rule is given by the following theorem.

Theorem 1. *Assume that $\alpha_1 \in (0, 1)$. Given the signal process (2.1) to (2.3), the equilibrium policy rule is given by*

$$h_a = \frac{\alpha_0}{1 - \alpha_1^2 \varphi_1}, \tag{2.14}$$

$$h_1(L) = \frac{h_a \alpha_1 (\varphi_1 - \vartheta L)}{1 - \vartheta L}, \tag{2.15}$$

$$h_2(L) = -\frac{h_a \alpha_1 \varphi_2}{1 - \vartheta L}, \tag{2.16}$$

where $\tau_1 = \frac{\sigma_a^2}{\sigma_\eta^2}$, $\tau_2 = \frac{\sigma_u^2}{\sigma_\eta^2}$, and

$$\varphi_1 = \frac{\rho \tau_1 + \vartheta \tau_2}{\rho(\tau_1 + \tau_2)}, \quad \varphi_2 = \frac{\tau_1(\rho - \vartheta)}{\rho(\tau_1 + \tau_2)}, \tag{2.17}$$

$$\vartheta = \frac{1}{2}\left[\left(\frac{1}{\rho} + \rho + \frac{(1-\alpha_1)(\tau_1 + \tau_2)}{\rho \tau_1 \tau_2}\right) - \sqrt{\left(\frac{1}{\rho} + \rho + \frac{(1-\alpha_1)(\tau_1 + \tau_2)}{\rho \tau_1 \tau_2}\right)^2 - 4}\,\right]. \tag{2.18}$$

The aggregate output follows

$$y_t = \vartheta y_{t-1} + \frac{h_a \alpha_1 \vartheta}{\rho} \eta_t. \tag{2.19}$$

Proof. See Appendix A.2 for proof. □

Even though agents face a fairly complicated learning problem, the equilibrium policy rule is simple. $h_1(L)$ is an ARMA(1,1) process and $h_2(L)$ is an AR(1) process. The aggregate output follows an AR(1) process. To understand the equilibrium policy rule, we discuss the following: the persistence of y_t, the unconditional variance of y_t, and the forecast error of y_t.

Endogenous Persistence of y_t Crucially, the persistence of y_t is given by ϑ in equation (2.18), which also determines the persistence of the effects of the confidence shock. We have derived the following properties for ϑ.

Proposition 2.3. *Assume that $\alpha_1 \in (0,1), \rho \in (0,1), \tau_1 > 0$ and $\tau_2 > 0$. Then ϑ satisfies*

1. $0 < \lambda < \vartheta < \rho$, *where*

$$\lambda = \frac{1}{2}\left[\frac{\tau_1+\tau_2}{\rho\tau_1\tau_2} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_1+\tau_2}{\rho\tau_1\tau_2} + \frac{1}{\rho} + \rho\right)^2 - 4}\right]. \tag{2.20}$$

2. ϑ *is increasing in* α_1 *and*

$$\lim_{\alpha_1 \to 1} \vartheta = \rho$$
$$\lim_{\alpha_1 \to 0} \vartheta = \lambda$$

3. ϑ *is increasing in* τ_1, τ_2 *and* ρ.

Proposition 2.3 states that ϑ is bounded from above by the persistence of the confidence shock ρ. Intuitively, agents gradually learn ξ_t from the signals and once they can infer ξ_t relatively accurately, we return to the perfect information benchmark and the confidence shock will have little effect on output. Consequently, the persistence of output is always smaller than the confidence shock. At the same time, ϑ is also bounded from below by λ. Here, $1-\lambda$ is the Kalman gain in predicting ξ_t using the Kalman filter, and λ controls the persistence of $\mathbb{E}_{it}[\xi_t]$, the forecast of ξ_t. To put it differently, λ determines the speed at which information is revealed, and it serves as the lower bound for the persistence of y_t.

Given the information related parameters ρ, σ_ϵ, σ_u, and σ_η, ϑ is increasing in α_1. As α_1 increases, there is stronger strategic complementarity. Agents respond more aggressively to possible good (bad) trading opportunities. As a result, the effects of confidence shocks last longer. In the extreme case, as α_1 approaches 1, the persistence of y_t approaches the persistence of ξ_t itself. Even though the information obtained by agents does not vary with α_1, the persistence of output chosen by individual agent varies with α_1 because of strategic complementarity.

It is not surprising that the persistence is increasing in τ_1 and τ_2, because the values of these two determine the degree of information frictions. Given the variance of innovation to the confidence shock σ_η^2, as σ_a or σ_u increases, it becomes more difficult to infer the confidence shock ξ_t, and the effects of the confidence shock last longer. Similarly, given the magnitude of idiosyncratic noise, the persistence of output decreases in σ_η.

Unconditional Variance of y_t The following proposition characterizes several properties of the variance of aggregate output:

Proposition 2.4. *Assume that $\alpha_1 \in (0,1), \rho \in (0,1), \tau_1 > 0$ and $\tau_2 > 0$. Then the unconditional variance of output y_t is given by*

$$\mathrm{Var}(y_t) = \frac{1}{1-\vartheta^2}\left(\frac{h_a \alpha_1 \vartheta}{\rho}\right)^2 \sigma_\eta^2, \tag{2.21}$$

and it has the following properties:

1. *There is a hump-shaped relationship between Var(y_t) and the variance of innovation to the confidence shock σ_η^2. Furthermore,*

$$\lim_{\sigma_\eta \to 0} \mathrm{Var}(y_t) = 0$$

$$\lim_{\sigma_\eta \to \infty} \mathrm{Var}(y_t) = 0$$

2. $\mathrm{Var}(y_t)$ *is increasing in α_1, σ_a, σ_u and ρ.*

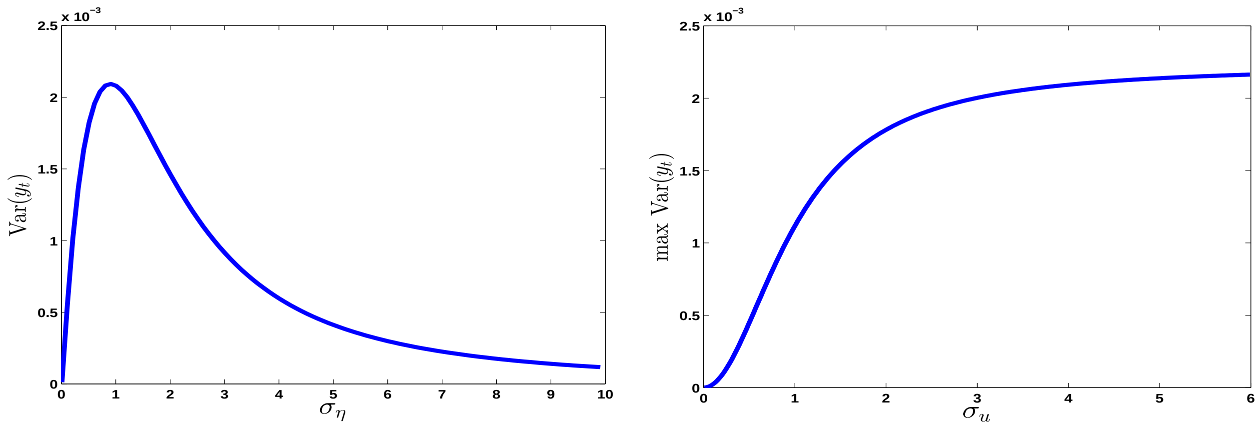
As discussed in the introduction, there are two competing forces that determine the variance of output. The volatility of output tends to increase with σ_η^2 because there are stronger exogenous disturbances. At the same time, with a larger σ_η^2, agents attenuate their response to signals because they understand that signals are less useful for information extraction. We can also define the maximum amount of volatility that can be generated by confidence shocks given certain information frictions

$$\max \mathrm{Var}(y_t) = \max_{\sigma_\eta} \frac{1}{1-\vartheta^2}\left(\frac{h_a \alpha_1 \vartheta}{\rho}\right)^2 \sigma_\eta^2. \tag{2.22}$$

The left graph in Figure 1 shows an example of how the variance of output changes with the variance of η_t, which displays a hump-shaped relationship. The right graph in Figure 1 shows that the maximum of the variance of output is increasing in the variance of idiosyncratic noise, but it is also bounded from above. The upper bound is determined by the underlying distribution of productivity across islands. This graph clearly illustrates that there exists a limit for the effects of confidence shocks on the aggregate economy.

Proposition 2.4 has two implications for our quantitative exercise in the next section. First, given the degree of information frictions, there is an upper bound for the variance of aggregate output by varying the variance of the confidence shock. If the degree of information frictions is relatively low, we may not be able to generate enough volatility of output. Second, there are two different values of

FIGURE 1: Relationship between the and σ_η, and



variance of the confidence shock which can generate the same volatility of output. These two choices of σ_η will imply different degrees of information frictions and consequently, different magnitudes of forecast errors. Both of these two implications indicate that it is crucial to discipline the degree of information frictions in order to evaluate the quantitative importance of the confidence shock.

Persistent Forecasting Error An important feature of the learning problem in this model is that the forecast error is persistent. In Kasa (2000) and Acharya (2013) where the number of signals equals the number of shocks, the forecast error only exists in one period and agents can learn the true state fairly quickly. The reason is that there are enough signals for agents to figure out the true state of the economy. In our economy, there are more shocks than signals. Agents can never infer the state of the economy perfectly and the forecast error is long lasting. This is crucial in generating the persistent effects of the confidence shock, because once the forecast error disappears, the economy returns to the perfect information case and the confidence shock no longer plays a role.

We look in particular at differences between the aggregate output and the average predicted aggregate output

$$y_t - \int \mathbb{E}_{it}[y_t] = h_a \alpha_1 \frac{1 - \frac{\lambda(\tau_1+\tau_2)}{\rho\tau_1\tau_2(1-\vartheta\lambda)} - \lambda L}{\rho(1-\lambda L)(1-\vartheta L)}\, \eta_t, \tag{2.23}$$

which follows an ARMA(2,1) process. This statistic is important in the calibration of the quantitative model. We will use the variance and persistence of this forecast error to pin down the degree of information frictions.

2.4 Example

In this section, we provide an example to show how the simple economy responds to a confidence shock. We choose parameters exogenously and they are summarized in Table 1.[6]

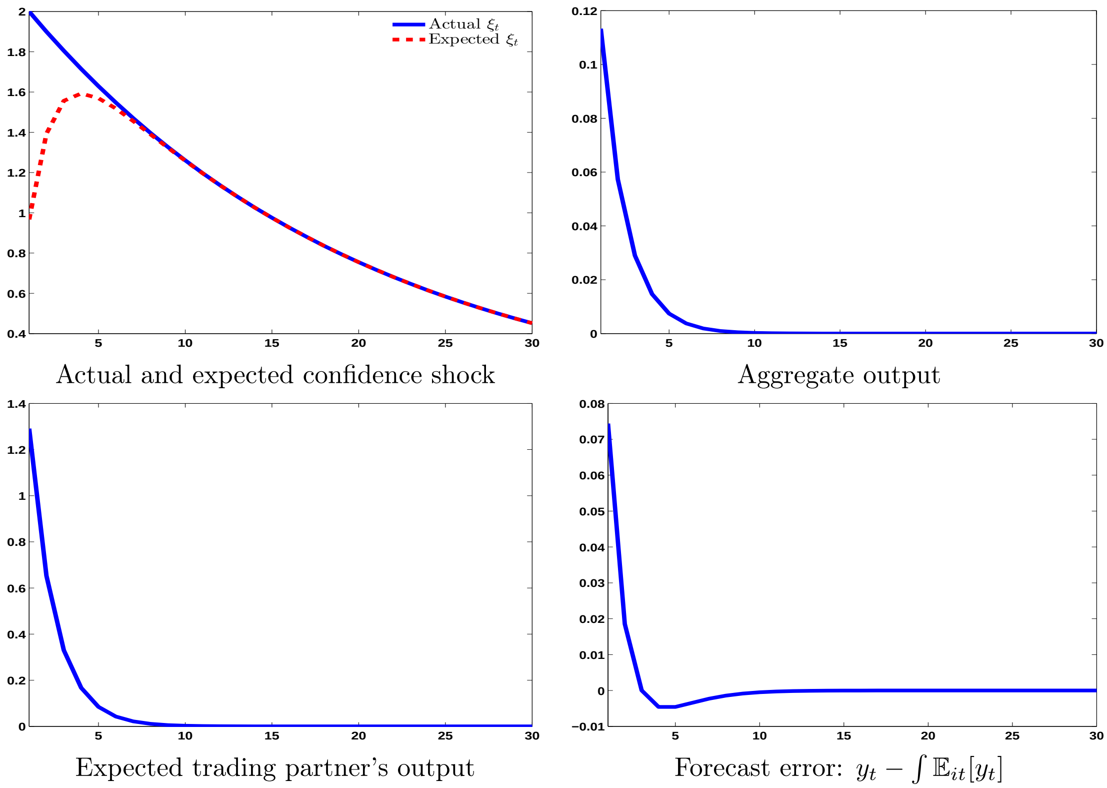
Table 1: Parameters for the Simple Economy

Predetermined	Description	Value
ω	Home bias	0.70
$\frac{1}{\gamma}$	Frisch elasticity	0.55
θ	Labor share	0.68
ρ	Persistence of confidence shock	0.95
σ_η	St.d of confidence shock	1.00
σ_a	St.d of productivity distribution	4.00
σ_u	St.d of noise shock	4.00
Implied	**Description**	**Value**
α_0	Response to own productivity	1.20
α_1	Strategic complementarity	0.09
ϑ	Endogenous persistence of output	0.70

The impulse response to confidence shocks is shown in Figure 2. At 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. Meanwhile, the aggregate output forecast error is persistent, and it resembles the pattern of the actual output. The parameters chosen imply that $\alpha_0 = 1.20$ and $\alpha_1 = 0.09$. The implied endogenous persistence of the output is $\vartheta = 0.70$.

[6] Because there is no intertemporal decision for households, we do not need to specify the discount factor and the risk aversion.

Figure 2: Impulse Response to a Confidence Shock in the Simple Economy

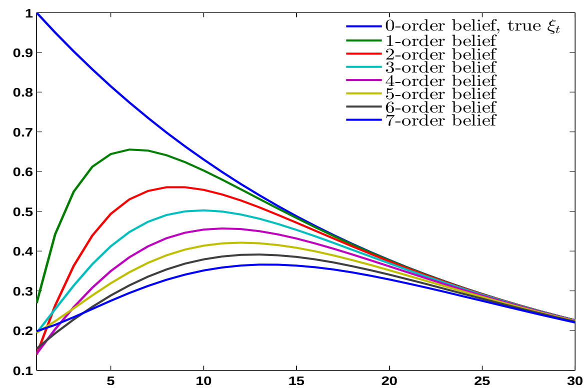


Higher Order Beliefs By Proposition 2.1, the aggregate output can also be written in the form of higher order beliefs

$$y_t = \frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k \left(\xi_t - \int \mathbb{E}_{it}^k[\xi_t] \right). \tag{2.24}$$

The effects of the confidence shock depend on the difference between the confidence shock and the higher order beliefs about the confidence shock. Figure 3 plots the the impulse response of the higher order beliefs. Initially, all the higher order beliefs are smaller than the true ξ_t, which implies that $\xi_t - \int \mathbb{E}_{it}^k[\xi_t] > 0$ and the output y_t will be high in the short run. Gradually, all the higher order beliefs converge to ξ_t, and the output y_t returns to its steady state value. As the order of the beliefs increases, the difference between ξ_t and $\mathbb{E}_{it}^k[\xi_t]$ also becomes greater. However, the effects of these higher order beliefs decay at rate α_1, meaning that as k approaches infinity, the effects of $\mathbb{E}_{it}^k[\xi_t]$ become zero. This intuition is discussed extensively in Nimark (2011).

Figure 3: Impulse Response of Higher Order Beliefs to the Confidence Shock



Heterogeneous Prior In Angeletos and La'O (2013) and Angeletos, Collard, and Dellas (2014), a heterogeneous-prior formulation is applied to avoid the infinite regress problem. The heterogeneous prior assumption works as follows. Assume that agents on island i observe both ξ_t and $a_{m(i,t)t}$ perfectly. However, they believe agents on island $m(i,t)$ observe a_i with bias ξ_t. If agent i's policy rule is

$$y_{it} = f_1 a_i + f_2 a_{m(i,t)} + f_3 \xi_t,$$

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

$$y_{m(i,t)t} = f_1 a_{m(i,t)} + f_2 (a_i + \xi_t) + f_3 \xi_t.$$

In equilibrium,

$$y_{it} = \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}[y_{m(i,t)t}],$$

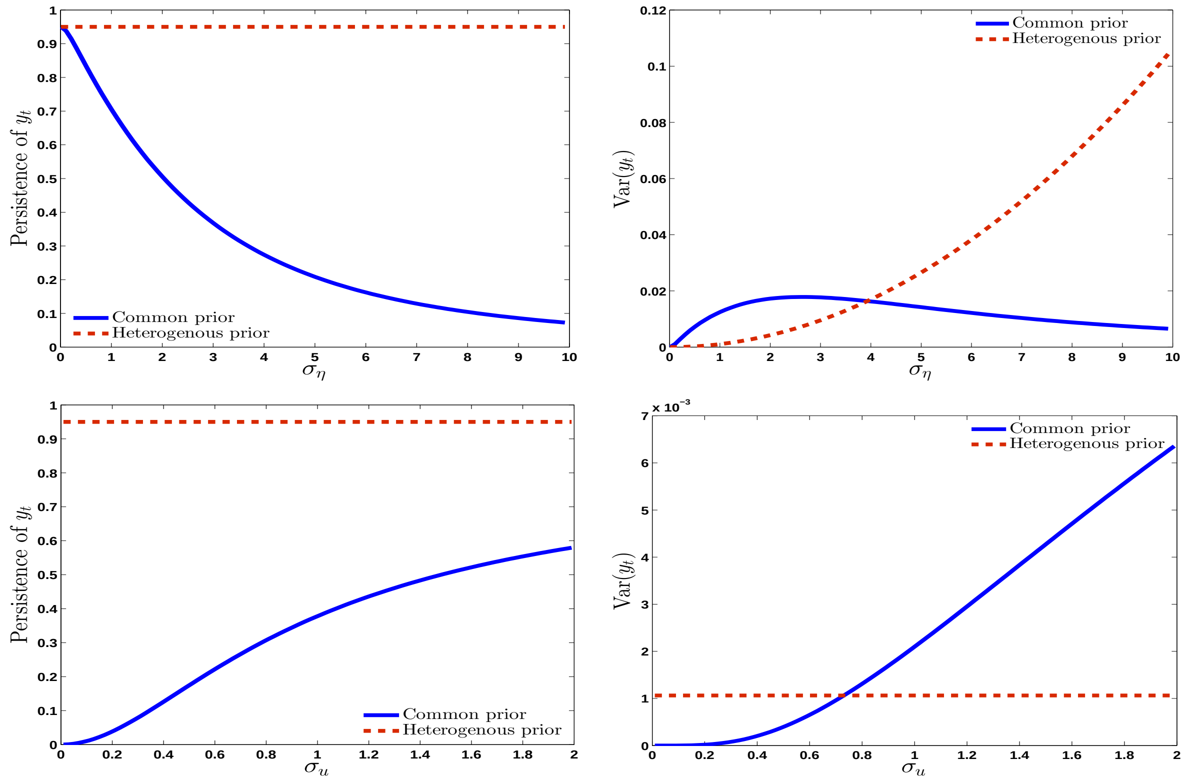
which leads to

$$y_{it} = \frac{1}{1-\alpha^2} a_i + \frac{\alpha}{1-\alpha^2} a_{m(i,t)} + \frac{\alpha_1^2}{(1-\alpha_1^2)(1-\alpha)} \xi_t \tag{2.25}$$

$$y_t = \frac{\alpha^2}{(1-\alpha^2)(1-\alpha)} \xi_t \tag{2.26}$$

By assuming heterogeneous prior beliefs, y_t is perfectly correlated with ξ_t, since the belief process is exogenously given. In Figure 4, we show how the persistence and variance of output vary with the variance of the confidence shock and the variance of idiosyncratic noise. With common prior, as we increase the variance of the confidence shock, τ_1 and τ_2 both decrease, and by Proposition 2.3, the persistence of output also decreases. By Proposition 2.4, there is a hump-shaped relationship between the variance of output and the variance of the confidence shock. In terms of information

Figure 4: Common Prior v.s. Heterogeneous Prior



frictions, both of the persistence and variance of output are monotonically increasing in the variance of idiosyncratic noise σ_u^2. With heterogeneous prior, the persistence of output is independent of the variance of the confidence shock, and the variance of output is monotonically increasing with the variance of the confidence shock. In addition, both of these two statistics are independent of the degree of information frictions.

3 Endogenous Information

In the previous section, the signal process was exogenously determined and independent of agents' actions. An important theme in the literature on dispersed information and higher order beliefs is the role of an endogenous signal in coordinating beliefs and revealing information.[7] In this section, we allow agents to observe signals that contain a variable which is endogenously determined in equilibrium.

More specifically, we allow agents to observe two signals. The first signal is the same as before, which

[7] See Kasa (2000), Hellwig and Venkateswaran (2011), and Rondina and Walker (2013) for example.

is their trading partner's productivity plus the confidence shock. The second signal is the aggregate output with an idiosyncratic noise. The aggregate output is endogenously determined by agents' output choice, but at the same time it serves as a signal for agents to infer the state of the economy. Agents understand that η_t is the underlying shock that drives the confidence shock and the aggregate output. Hence, observing the noisy signal of aggregate output will help them predict η_t and in turn the confidence shock. Formally, the equilibrium with endogenous information is defined as follows.

Definition 3.1. *The equilibrium is an endogenous stochastic process Ω_{it}, a policy rule for individual agents $\phi = \{\phi_a, \phi_1, \phi_2, \phi_3\} \in \mathbb{R} \times \ell^2 \times \ell^2 \times \ell^2$ and the law of motion for aggregate output $\Phi \in \ell^2$, such that*

1. *Information process generating Ω_{it} is given*

$$x_{it}^1 = a_{m(i,t)} + \xi_t, \tag{3.1}$$
$$x_{it}^2 = y_t + u_{it}, \tag{3.2}$$

where

$$\xi_t = \frac{1}{1-\rho L}\eta_t, \tag{3.3}$$
$$y_t = \Phi(L)\eta_t. \tag{3.4}$$

2. *Individual rationality*

$$y_{it} = \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}[y_{m(i,t)}], \tag{3.5}$$

where

$$y_{it} = \phi_a a_i + \phi_1(L) a_{m(i,t)} + \phi_2(L) u_{it} + \phi_3(L)\eta_t. \tag{3.6}$$

3. *Aggregate consistency*

$$\Phi(L) = \phi_3(L). \tag{3.7}$$

The policy rule in this definition is in terms of the underlying shocks. As proved in Huo and Takayama (2014), there is a one-to-one mapping between the policy defined in terms of signals and shocks. With endogenous information, it is more convenient to express the policy rule in terms of shocks, because it clearly separates the idiosyncratic components from the aggregate components. The equilibrium with endogenous information involves two fixed points. The first fixed point is individual rationality. Given the signal process, all islands choose the same policy rule ϕ that solves their optimization problem. Agents need to infer higher order beliefs, and the infinite regress problem is still there. The second fixed point is absent in the equilibrium with exogenous information. It requires that the perceived law of motion for aggregate output be the same as the law of motion for actual aggregate output. This can be viewed as the cross-equation restriction in the sense that agents perceptions are in line with the reality generated by their own actions.

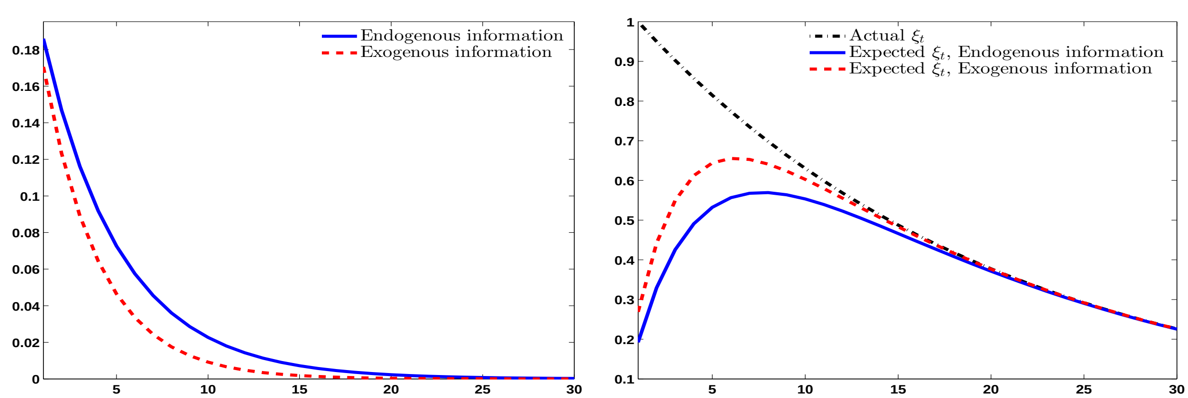
Since there are more shocks than signals, agents cannot infer the shocks perfectly. The information role of output depends on the volatility of output. If the aggregate output is very volatile, then the second signal will be very informative about the confidence shock. However, once agents can learn quickly the state of the economy from aggregate output, the effects of the confidence shock will be very limited, which implies that the aggregate output can not respond to the confidence shock aggressively. Conversely, if there is little movement of aggregate output, then agents will pay little attention to the second signal and attribute a big portion of the confidence shock to their trading partner's productivity. Under this scenario, the confidence shock will generate large movements of aggregate output, which is a contradiction. The argument above provides the intuition for the existence of the equilibrium: there exists a point such that the volatility of aggregate output is neither too large nor too small.

Theorem 2. *If $\alpha_1 \in (0, 1)$, then there exists a unique equilibrium of the model in Definition 3.1.*

Proof. See Appendix A.3 for the proof. □

As shown in Huo and Takayama (2014), even though there exists a unique equilibrium, aggregate output follows an infinite-order process. As a result, no analytic solution is possible any more. We use the method discussed in Huo and Takayama (2014), and approximate the aggregate output by an ARMA (3,2) process. This approximation is close enough to the true solution.

FIGURE 5: Endogenous Information versus Exogenous Information



Output y_t Expected ξ_t

Figure 5 compares the impulse response of the aggregate output to the confidence shock under endogenous information with the one under exogenous information. It can be seen that the output under endogenous information is more responsive to the confidence shock. To understand the results, we need to highlight the information role of the aggregate output. Since α_1 is small in our example, the force of strategic complementarity is weak, and hence the aggregate output is not very volatile. As a result, the endogenous signal $x_{it}^2 = y_t + u_{it}$ conveys less information about the confidence shock compared with the exogenous signal $x_{it}^2 = \xi_t + u_{it}$. Note that in Figure 5 the prediction of the confidence shock is indeed less accurate with endogenous information. Therefore, when α_1 is small, the effects of the confidence shock are greater under endogenous information than under exogenous information. Conversely, if we set α_1 to a large number, the aggregate output will be more volatile than ξ_t itself. It follows that the endogenous signal $x_{it}^2 = y_t + u_{it}$ will contain more information than the exogenous signal $x_{it}^2 = \xi_t + u_{it}$. Consequently, the effects of the confidence shock will be greater with exogenous information.

This example illustrates that whether agents observe exogenous signals or endogenous signals does not really matter. What matters is how much information agents can learn about the underlying state of the economy. At the end of the day, individual agents treat all signals as exogenously given, and we can change the size of the noise shocks to control the amount of information agents can extract. Based on this observation, in our quantitative model, we assume all information follows an exogenous process, but it should be noted that this assumption is not crucial to the purpose of evaluating the role of confidence shocks.

4 Quantitative Model

In this section, we present the full-blown business cycle model driven by confidence shocks. To evaluate its quantitative performance and confront the model with data, several issues need to be addressed. First, the confidence shock itself and the idiosyncratic noises cannot be observed, but we need to pin down the degree of information frictions. Second, because the confidence shock does not affect aggregate technology, the Solow residual remains constant. As a result, all the short-run fluctuations are driven by changes in employment, which is at odds with data. Third, aggregate investment is important in shaping business cycles, and agents constantly make intertemporal decisions. Based on these considerations, we extend the simple model presented in Section 2 along three dimensions: (1)we adopt a more flexible matching process and information structure, which allows us to link the model with the survey data in order to discipline information frictions; (2)we introduce competitive search in the goods market à la Bai, Ríos-Rull, and Storesletten (2011)), which generates endogenous movements of the Solow residual; (3) we allow households to accumulate capital.

4.1 Model

Matching and Information In the simple model, we assume that the matching follows an i.i.d. process, that is, the quality of island i's trading partner in period t is completely independent of its trading partner in period $t-1$. This assumption is convenient for deriving analytic results, but it is far from being realistic. If we interpret an island as an establishment, a firm, or a region, the output or revenue of these entities is typically correlated over time. Meanwhile, the exact form of the matching process is also related to the degree of information frictions. Therefore, we allow the matching process to be persistent. Namely, if island i is matched with a good trading partner today, it is more likely that island i is also matched with a good trading partner tomorrow. Recall that we denote the index of island i's trading partner in period t as $m(i,t)$, and we now assume that the productivity of $m(i,t)$ follows an AR(1) process

$$a_{m(i,t)} = \rho_a a_{m(i,t-1)} + \epsilon_{it}, \tag{4.1}$$

where $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ and $\sigma_\epsilon^2 = (1-\rho_a^2)\sigma_a^2$. Note that the choice of σ_ϵ guarantees that the unconditional variance of $a_{m(i,t)}$ is consistent with σ_a^2. If we set $\rho_a = 0$, it collapses to the original i.i.d. matching process. The following proposition proves the existence of the persistent matching process.

Proposition 4.1. *Let $m(i,t)$ be island i's trading partner at time t and $a_{m(i,t)}$ be its productivity. There exists a stochastic process such that, for all $i \in [0,1)$,*

$$\begin{aligned} a_{m(i,t)} &= \rho_a a_{m(i,t-1)} + \epsilon_{it}, \\ \epsilon_{it} &\sim N(0,\sigma^2) \end{aligned}$$

where $\rho_a \in (0,1)$.

Proof. See Appendix A.4 for the proof. □

Turing to the information part, at the beginning of each period, we still assume that a household head receives two signals. The first signal concerns their trading partner's productivity, but it is contaminated by the common confidence shock

$$x_{it}^1 = a_{m(i,t)} + \xi_t. \tag{4.2}$$

The process of $a_{m(i,t)}$ is specified in equation (4.1). The confidence shock ξ_t follows the same AR(1) process as the simple model

$$\xi_t = \rho\xi_t + \eta_t, \tag{4.3}$$

where $\rho \in (0, 1)$ and $\eta_t \sim N(0, \sigma_\eta^2)$.

The second signal is the confidence shock plus an idiosyncratic shock.

$$x_{it}^2 = \xi_t + u_{it}, \tag{4.4}$$

Different from the simple model, we assume the idiosyncratic shock u_{it} follows a persistent process instead of an i.i.d. process

$$u_{it} = \rho_u u_{it-1} + \upsilon_{it}, \tag{4.5}$$

where $\rho_u \in (0, 1)$ and $\upsilon_{it} \sim N(0, \sigma_\upsilon^2)$. The persistence of the idiosyncratic noise affects the persistence of the forecast error. We will discuss how to set this parameter in the calibration section.

The information set, up to time t, is

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

Competitive Search and Shoppers' Problem In the simple model, the goods market between the two trading partners is frictionless. Shoppers from the two islands meet in a centralized market, and the prices P_i and P_j clear the goods market. To introduce goods market frictions, we now assume that trading takes place at decentralized markets and is subject to additional frictions, which we will describe in detail.

Shoppers serve both as sellers and buyers. 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 household heads. 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 shoppers to exert search efforts 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 shopper expects to find a location with probability $\Psi^d(Q)$ at price P. Not all markets are active. In fact, shoppers understand that there is an equilibrium-determined expected revenue per unit of good, $\zeta = P\ \Psi^f(Q)$, that active markets have to satisfy. As a seller, each shopper is endowed with 1 unit measure of locations. A shopper in market (P, Q) expects to sell her goods with probability $\Psi^f(Q)$ at price P. 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 shoppers on island j choose (P_{jj}, Q_{jj}) and (P_{ji}, Q_{ji}). In equilibrium, $P_{ii}\Psi^f(Q_{ii}) = P_{ji}\Psi^f(Q_{ji}) = \zeta_i$, and $P_{jj}\Psi^f(Q_{jj}) = P_{ij}\Psi^f(Q_{ij}) = \zeta_j$.

Crucially, not all goods can be sold and the produced goods Y_i and Y_j are only potential output. The realized 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

$$\max_{\substack{C_{ii},C_{ij},I_{ii},I_{ij},\\ Q_{ii},Q_{ij},D_{ii},D_{ij}}} \left(\frac{C_{ii}}{\omega}\right)^{\omega} \left(\frac{C_{ij}}{1-\omega}\right)^{1-\omega} - \chi_d D_i \tag{4.6}$$

subject to

$$P_{ii}(C_{ii}+I_{ii}) + P_{ij}(C_{ij}+I_{ij}) = \zeta_i Y_i, \tag{4.7}$$

$$C_{ii}+I_{ii} = D_{ii}\Psi^d(Q_{ii})Y_i, \tag{4.8}$$

$$C_{ij}+I_{ij} = D_{ij}\Psi^d(Q_{ij})Y_j, \tag{4.9}$$

$$P_{ii}\Psi^f(Q_{ii}) = \zeta_i, \tag{4.10}$$

$$P_{ij}\Psi^f(Q_{ij}) = \zeta_j, \tag{4.11}$$

$$I_i = \left(\frac{I_{ii}}{\omega}\right)^{\omega} \left(\frac{I_{ij}}{1-\omega}\right)^{1-\omega}, \tag{4.12}$$

$$D_i = D_{ii} + D_{ij}. \tag{4.13}$$

This calls for several comments. The head of a household now determines both the level of production Y_i and the level of capital investment I_i in the first stage. As a result, the head not only transfers the output Y_i to the shopper, but also requires the shopper to purchase the investment good such that the composite of I_{ii} and I_{ij} satisfies the head's investment demand I_i. The search effort D_i is the new elements in the shoppers' problem, and the variation in D_i leads to changes in the utilization rate. Related to the search effort, as discussed in Huo and Rios-Rull (2014), shoppers with different incomes choose markets with different prices and search intensities.

The equilibrium conditions include

$$Q_{ii} = \frac{T_{ii}}{D_{ii}},\ \ Q_{ij} = \frac{T_{ji}}{D_{ij}},\ \ Q_{ji} = \frac{T_{ij}}{D_{ji}},\ \ Q_{jj} = \frac{T_{jj}}{D_{jj}}, \tag{4.14}$$

$$T_{ii}+T_{ij} = 1, \quad T_{ji}+T_{jj} = 1, \tag{4.15}$$

$$\zeta_i = P_{ii}\Psi^f(Q_{ii}) = P_{ji}\Psi^f(Q_{ji}), \tag{4.16}$$

$$\zeta_j = P_{jj}\Psi^f(Q_{jj}) = P_{ij}\Psi^f(Q_{ij}). \tag{4.17}$$

Implicitly, shoppers also choose to which market they send their locations T_{ii} and T_{ij}, but they are indifferent to the markets since they will obtain the same expected revenue ζ_i.

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

$$\Psi^d(Q) = \nu Q^{1-\mu}, \tag{4.18}$$
$$\Psi^f(Q) = \nu Q^{-\mu}, \tag{4.19}$$

where μ is the matching elasticity. The equilibrium allocations satisfy

$$C_{ii}^* = \omega\nu \left(\frac{\mu\nu}{\chi_d}\right)^{\frac{\mu}{1-\mu}} Y_i^{\frac{1-\mu+\mu\omega}{1-\mu}} Y_j^{\frac{\mu(1-\omega)}{1-\mu}} - \omega \left(\frac{Y_i}{Y_j}\right)^{1-\omega} I_i, \tag{4.20}$$
$$C_{ij}^* = (1-\omega)\nu \left(\frac{\mu\nu}{\chi_d}\right)^{\frac{\mu}{1-\mu}} Y_i^{\frac{\mu\omega}{1-\mu}} Y_j^{\frac{1-\mu\omega}{1-\mu}} - (1-\omega) \left(\frac{Y_i}{Y_j}\right)^{-\omega} I_i, \tag{4.21}$$
$$D_i^* = \left(\frac{\mu\nu}{\chi_d} Y_i^{\omega} Y_j^{1-\omega}\right)^{\frac{1}{1-\mu}}. \tag{4.22}$$

As the production level increases, shoppers purchase more consumption goods. At the same time, they also exert more search efforts. Because the total measure of locations is fixed, more search effort translates into a higher utilization rate and the matching elasticity μ determines the percentage increase of the utilization rate.

The indirect utility function derived from the second stage is

$$\mathcal{U}(Y_i, Y_j, I_i) = \left(\frac{C_{ii}^*}{\omega}\right)^{\omega} \left(\frac{C_{ij}^*}{1-\omega}\right)^{1-\omega} - \chi_d D_i^* = \frac{\mu\chi_d}{1-\mu} \left(\frac{\mu\nu}{\chi_d}\right)^{\frac{1}{1-\mu}} Y_i^{\frac{\omega}{1-\mu}} Y_j^{\frac{1-\omega}{1-\mu}} - I_i, \tag{4.23}$$

which will be used in the heads' problem.

Heads' Problem Compared to the simple model, the only complication of the heads' problem is the addition of investment choice. The heads' problem is choosing a state contingent plan for Y_{it}, K_{it+1} and N_{it} to maximize their expected present value.

$$\max_{Y_{it}, N_{it}, K_{it+1}, I_{it}} \mathbb{E}_{i0} \sum_{t=0}^{\infty} \beta^t \frac{[\mathcal{U}(Y_{it}, Y_{m(i,t)t}, I_{it}) - \chi_n N_{it}^{1+\gamma}]^{1-\sigma}}{1-\sigma}$$

subject to

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

We assume that the investment is subject to a standard capital adjustment cost $\Xi(I_{it}, K_{it})$ with the following functional form

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

To derive the first order conditions, we first substitute the production function into the objective function and define

$$\mathcal{V}(Y_{it}, a_i, K_{it}) = \chi_n \left(\frac{Y_{it}}{\exp(a_i)\; K_{it}^{1-\theta}}\right)^{\frac{1+\gamma}{\theta}}. \tag{4.24}$$

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

$$\mathbb{E}_{it}\left[[\mathcal{U}(Y_{it}, Y_{m(i,t)t}, I_{it}) - \chi_n N_{it}^{1+\gamma}]^{-\sigma}\left(\mathcal{U}_{y_{it}} - \mathcal{V}_{y_{it}}\right)\right] = 0 \tag{4.25}$$

The first order condition with respect to K_{it+1} is

$$\frac{\mathbb{E}_{it}\left[\left[\mathcal{U}(Y_{it}, Y_{m(i,t)t}, I_{it}) - \chi_n N_{it}^{1+\gamma}\right]^{-\sigma}\right]}{1 - \Xi_i(K_{it}, I_{it})} = \beta\mathbb{E}_{it}\Bigg[\left[\mathcal{U}(Y_{it+1}, Y_{m(i,t+1)t+1}, I_{it+1}) - \chi_n N_{it+1}^{1+\gamma}\right]^{-\sigma} \left(\mathcal{U}_{y_{it+1}}(1-\theta)\exp(a_i)K_{it+1}^{-\theta}N_{it+1}^{\theta} + \frac{1-\delta-\Xi_k(K_{it+1}, I_{it+1})}{1-\Xi_i(K_{it+1}, I_{it+1})}\right)\Bigg] \tag{4.26}$$

These two first order conditions are quite similar to those in standard stochastic growth models, except marginal returns to production depend on the heads' expectation of their trading partners' output level. As in a two-country business cycle model, the output and investment decisions both increase with their trading partners' output level.

Log-Linearized Economy Equation (4.25) and (4.26) summarize the heads' decisions. The log-linearized version of these two equations is:

$$\Gamma_1 a_i + \Gamma_2 y_{it} + \Gamma_3 k_{it} + \Gamma_4 \mathbb{E}_{it}[y_{m(i,t)t}] = 0, \tag{4.27}$$

$$\Upsilon_1 k_{it} + \Upsilon_2 k_{it+1} + \Upsilon_3 \mathbb{E}_{it}[y_{m(i,t)t}] + \Upsilon_4 \mathbb{E}_{it}[y_{it+1}] + \Upsilon_5 \mathbb{E}_{it}[k_{it+2}] + \Upsilon_6 \mathbb{E}_{it}[y_{m(i,t+1)t+1}] = 0, \tag{4.28}$$

where $\{\Gamma_1, \ldots, \Gamma_4\}$ and $\{\Upsilon_1, \ldots, \Upsilon_6\}$ are functions of the deep parameters. Similarly to the simple model, the equilibrium is defined as:

Definition 4.1. *Given the signal process (4.1) to (4.5), the equilibrium is policy rules* $h^y = \{h_a^y, h_1^y, h_2^y\} \in \mathbb{R} \times \ell^2 \times \ell^2$ *and* $h^k = \{h_a^k, h_1^k, h_2^k\} \in \mathbb{R} \times \ell^2 \times \ell^2$

$$y_{it} = h_a^y a_i + h_1^y(L)x_{it}^1 + h_2^y(L)x_{it}^2, \tag{4.29}$$

$$k_{it+1} = h_a^k a_i + h_1^k(L)x_{it}^1 + h_2^k(L)x_{it}^2, \tag{4.30}$$

such that equations (4.27) and (4.28) are satisfied.

To solve for the equilibrium, we apply the method developed in Huo and Takayama (2014). The details of the computation can be found in our online appendix.

4.2 Calibration

The model period is a quarter. We separate the parameters into two groups: those in the first group (shown in Table 2) are determined exogenously, and those in the second group (shown in Table 3) are jointly determined by solving a large system: the equations require that the steady-state model statistics equal the targets, and the parameters are the unknowns.

Many parameters of preferences and technology are standard, and we choose them to reflect commonly used values. We set the discount rate β to 0.99, which implies that the rate of return is 4%. We set the risk aversion σ to 1. We choose the Frisch elasticity to be $\frac{1}{\gamma} = 0.55$, which lies between the micro and macro estimates. We choose the labor share $\theta = 0.68$, in line with Ríos-Rull and Santaeulalia-Llopis (2010). The home bias parameter matters for the degree of strategic complementarity. We set $\omega = 0.7$ as our benchmark value.

Turning to the matching process. If we interpret each island as a firm, the persistence of the matching process directly translates into the persistence of the measured firms' profit or productivity even though their technology is unchanged. The empirical estimate of the persistence of the firms' productivity varies in the literature, ranging from 0.5 (Abraham and White, 2006) to 0.8 (Foster, Haltiwanger, and Syverson, 2008) for the United States, and it varies even more when examining other countries (Collard-Wexler, Asker, and Loecker, 2011). We set $\rho_a = 0.7$ and $\sigma_\epsilon = 0.1$, which lie in the middle of various estimates. Note that σ_a is determined residually by $\sigma_a^2 = \frac{\sigma_\epsilon^2}{1-\rho_a^2}$. We will conduct robustness checks for different values of ρ_a and σ_ϵ.

The matching elasticity is particularly important in shaping the endogenous Solow residual. The realized aggregate output is:

$$\overline{y} = \int \Psi^f(q_{ii}) + \int y_i = z + y$$

Here, we use $\overline{y}$ to denote the aggregate output, or realized sales, y to denote the potential output, or produced goods, and z to denote the measured Solow residual. Using equation (4.22), the measured Solow residual is proportional to the potential output

$$z = \int \Psi^f(q_{ii}) \;\; \propto \;\; \frac{\mu}{1-\mu} \int d_i \;\; \propto \;\; \frac{\mu}{1-\mu} y \;\; = \mu \overline{y}$$

Therefore, the matching elasticity μ determines the portion of the output fluctuations which can be attributed to the Solow residual. As a benchmark, we set $\mu = 0.4$.

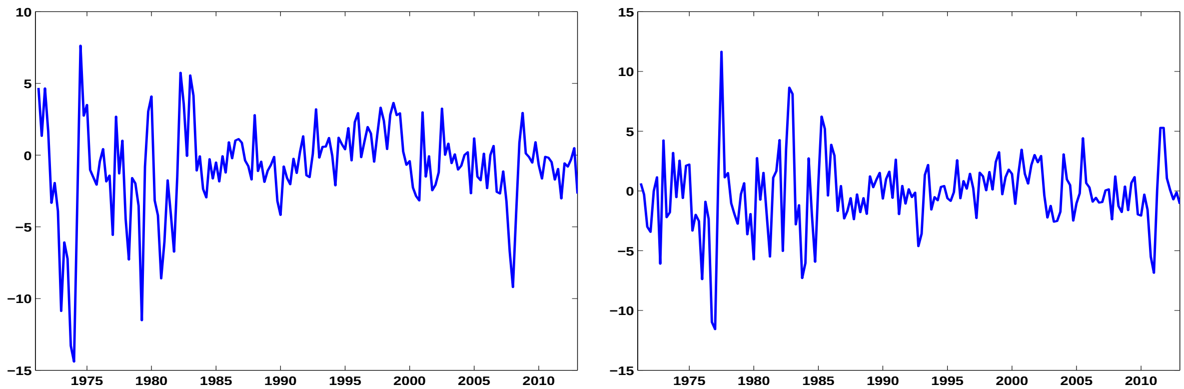
Table 2: Exogenously Determined Parameters

Parameter	Description	Value
β	Discount rate	0.99
σ	Risk aversion	1.00
ω	Home bias	0.70
$\frac{1}{\gamma}$	Frisch elasticity	0.55
θ	Labor share	0.64
μ	Matching elasticity	0.40
ρ	Persistence of confidence shock	0.95
ρ_a	Persistence of matching quality	0.70
σ_a	St.d of island specific productivity	0.14

In terms of the endogenously determined parameters, we associate the parameters with the targets for which they are most directly responsible, even though these parameters are eventually determined simultaneously. We choose χ_n to target the average working time to be 0.4 which only serves as a normalization. We target the capital-output ratio to be 2, which pins downs the capital depreciation rate δ. Two parameters are related to goods market frictions: the units of search costs ξ_d and the matching efficiency ν. We choose the values for them so that the average occupation rate is 81% and the average market tightness is 1. We set the capital adjustment cost ψ to match the relative volatility of investment to output.

An important part of the calibration is to discipline the degree of information frictions. The two related parameters are the persistence of the idiosyncratic noise, ρ_u, and the standard deviation of innovation to the idiosyncratic noise, σ_v. The most direct way to determine information frictions is to look at agents' forecast errors. The Survey of Professional Forecasters (SPF) provides a quarterly survey of macroeconomic forecasts for the economy of the United States. The survey participants have formal, advanced training in economic theory. These survey forecasts are generally better than forecasts generated by econometric models. Figure 6 displays the 1-year-ahead and 1-quarter-ahead mean forecast error for aggregate output. These two forecast error series are similar in terms of standard deviation, but have clear differences in persistence. The persistence for the 1-year-ahead forecast error is 0.58, but it is only 0.28 for the 1-quarter-ahead forecast error. Agents in the

FIGURE 6: Mean Forecast Error in SPF



1-year-ahead GDP forecast error

1-quarter-ahead GDP forecast error

model are interpreted as normal households and firms, who have less information compared with the professional forecasters in the SPF in general. As a benchmark, we calculate the 1-quarter-ahead forecast error in our model, and choose ρ_u and σ_v to match the persistence and the standard error of the 1-quarter-ahead forecast error in the SPF. We label this calibration small-information-friction calibration. We also choose ρ_u and σ_v to match the 1-quarter-ahead forecast error in the model with the 1-year-ahead forecast error in the SPF, which is labelled large-information-friction calibration. We think the persistence of the forecast error for the general public should lie between these two benchmarks.

The last two parameters which determine the process of the confidence shock are ρ and σ_η. The Index of Consumer Sentiment and the Consumer Confidence Index reflect households' views of the aggregate and local economic condition, but they do not directly correspond to the confidence shock in our model. The confidence shock in our model is not observable. It should be clear that ρ itself does not determine the persistence of the equilibrium allocation, and it is the degree of information frictions that eventually determines the persistence. The parameters ρ_u, σ_ν, and ρ cannot be separately identified. We set ρ to 0.95, and vary the other two parameters to determine the information frictions. We choose σ_η such that the volatility of output in our model is the same as that in the data. Note that when we choose σ_η, it is not only a normalization. As shown in Section 2, the volatility of output is not monotonically increasing in σ_η. Supposing information frictions are small, there is a chance that the volatility of output in the model is always smaller than that in the data. If this is the case, we choose σ_η to maximize the volatility of output in the model. As shown in Figure 7, under small-information-calibration, we can account for over 90% of the output volatility, but we cannot match the volatility of output exactly. By allowing large information frictions, the output

Table 3: Endogenously Determined Parameters

Parameter	Value	Target	Value	Model
χ_n	1.02	Average labor	0.40	0.40
χ_d	0.68	Average market tightness	1.00	1.00
ν	0.81	Average utilization rate	0.81	0.81
δ	0.02	Capital-to-output ratio	2.00	2.00
ψ	16.12	Ratio of st.d of investment to output	4.00	3.92
Small information friction				
σ_η	3.68%	St.d of output	1.54%	1.47%
ρ_u	0.15	Persistence of forecast error	0.28	0.29
σ_u	4.27%	St.d of forecast error	0.73%	0.73%
Large information friction				
σ_η	3.17%	St.d of output	1.54%	1.54%
ρ_u	0.56	Persistence of forecast error	0.56	0.59
σ_u	4.12%	St.d of forecast error	0.80%	0.80%

volatility can be matched exactly.

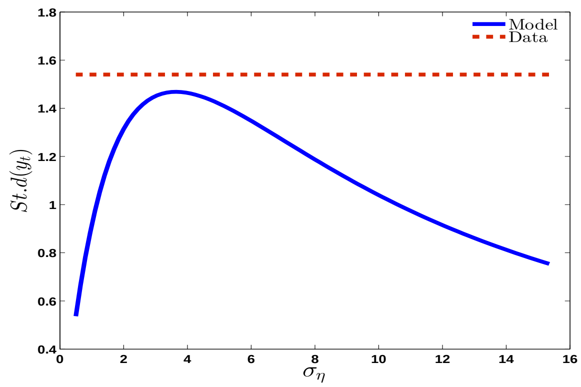
In a standard log-linearized DSGE model, the standard deviation of a shock is independent of policy rules. By contrast, in our model with information frictions, the relative volatility of various shocks, i.e., $\frac{\sigma_\epsilon}{\sigma_\eta}$ and $\frac{\sigma_\nu}{\sigma_\eta}$, does have direct effects on the policy rules, and σ_η have non-linear effects on aggregate variables.

4.3 Results

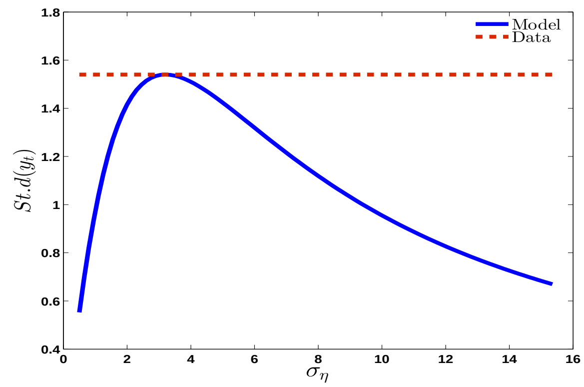
Figure 8 shows the impulse response of the main aggregate variables to the confidence shock under the large-information-friction calibration.[8] At the beginning, agents underestimate the confidence shock and attribute a part of the confidence shock to a good realization of the matching process. As a result, household heads believe that their trading partners' output is higher than average, and it will be so for a while due to the fact that the matching process is persistent. Because of strategic complementarity, believing that there is higher output on other islands leads to a higher output and

[8]In Figure 8, we choose the size of the confidence shock such that the initial response of output is 1.

FIGURE 7: Volatility of Aggregate Output



Small-information-friction calibration



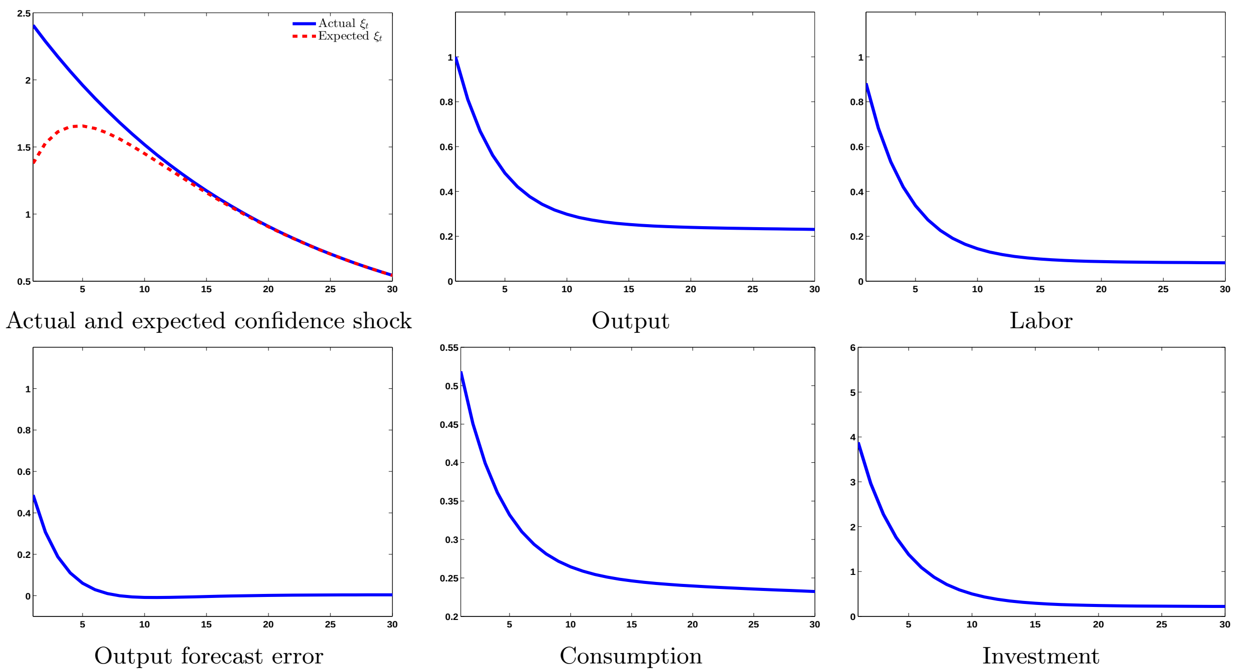
Large-information-friction calibration

investment level on their own islands, and thereby to a high aggregate output and investment. This belief is partially true, since the output on other islands is indeed higher than average. However, it is not because the productivity is higher, but because all the islands are optimistic. After a confidence shock, agents on average overestimate their trading partners' output and underestimate the aggregate output in the short run.

Table 4 compares the business cycle statistics from the data and our model driven by confidence shocks. Models with both large and small information frictions can produce reasonable volatility, even though the volatility is smaller than the data counterpart with small information frictions. From the demand side, the standard deviation of investment is approximately 4 times larger than that of output, similarly to the data. The volatility of consumption is smaller than the volatility of output, but it is less volatile compared with the data. From the supply side, the change in the output can be decomposed into the change in labor and the change in the measured Solow residual. The standard deviation of labor is 70% of its data counterpart, which we think is acceptable given that we choose a relatively low Frisch elasticity. Recall that there is no change in aggregate TFP, the changes in the measured Solow residual are entirely endogenous, driven by shoppers' searching activities. We have chosen the matching elasticity $\mu = 0.4$, which implies that when total output increases by 1%, the measured Solow residual increases by 0.4%. Goods market frictions and shopping efforts can account roughly for half of the observed fluctuations of the Solow residual.

The model cannot generate the same persistence of aggregate variables as in the data. The basic mechanism of the model is that the behavior output mirrors the behavior of forecast errors. In the data, the forecast errors are only modest persistent, which implies that the persistent of output in the

Figure 8: Impulse Response to the Confidence Shock: Large-Information-Friction Calibration



model cannot be too high. With large information frictions, the autocorrelation of various variables is around 0.6, but it declines to around 0.45 with small information frictions. Even though capital introduces additional persistence, this effect is not strong enough to allow the model to achieve the same persistence as in the data. To match the autocorrelation in the data, it seems necessary to include other more persistent real shocks.

Comparison with RBC Models Now we compare our model driven by confidence shocks with the RBC model driven only by TFP shocks. The RBC model we use is the same as our quantitative model presented in Section 4.1 except for three differences: (1)there is no competitive search in the goods market and hence no endogenous Solow residual; (2)there are exogenous shocks to aggregate TFP; and (3) there is no information friction.

We assume that the aggregate TFP shock follows an AR(1) process

$$z_t = \rho_z z_{t-1} + \varsigma_t, \tag{4.31}$$

where $\varsigma_{it} \sim N(0, \sigma_\varsigma^2)$. After subtracting a linear trend, we estimate process (4.31) and obtain $\rho_z = 0.96$ and $\sigma_\varsigma = 0.0078$. With the aggregate TFP shock, the productivity of an individual island's

Table 4: Business Cycle Statistics

	Data	RBC model	Common-prior model, ξ shock		Hetero-prior model
		TFP shock	Small infor friction	Large infor friction	ξ shock
Std. deviation					
Y	1.54	1.16	1.47	1.54	1.54
C	1.26	0.62	0.76	0.81	0.96
I	6.87	4.49	5.79	6.03	4.24
N	1.86	0.41	1.31	1.37	2.20
Z	1.24	0.88	0.59	0.62	—
$\frac{Y}{N}$	0.94	0.74	0.19	0.23	0.67
LW_1	4.87	—	2.94	3.08	2.71
LW_2	3.96	—	2.21	2.30	1.76
Corr with Y					
Y	1.00	1.00	1.00	1.00	1.00
C	0.88	0.99	0.99	0.99	0.99
I	0.91	0.99	1.00	0.99	0.99
N	0.86	1.00	1.00	1.00	1.00
Z	0.77	0.99	1.00	1.00	—
$\frac{Y}{N}$	-0.07	1.00	0.84	0.80	-0.99
LW_1	-0.84	—	-1.00	-1.00	-1.00
LW_2	-0.75	—	-0.99	-0.99	-1.00
Autocorrelation					
Y	0.87	0.74	0.43	0.61	0.70
C	0.88	0.75	0.46	0.64	0.71
I	0.83	0.73	0.42	0.60	0.69
N	0.92	0.74	0.42	0.60	0.69
Z	0.81	0.73	0.43	0.61	—
$\frac{Y}{N}$	0.77	0.74	0.63	0.79	0.69
LW_1	0.92	—	0.42	0.61	0.70
LW_2	0.91	—	0.42	0.60	0.69
Std. deviation of confidence shock					
σ_η	—	—	3.68	3.17	1.42

Note: All variables are HP-filtered logarithms of the original series. The standard deviations are multiplied by 100. LW_1 is the labor wage defined by the standard separable utility function $U(C,N) = \log C - \frac{N^{1+\gamma}}{1+\gamma}$, and $LW_1 = \log(\frac{Y}{N}) - \log(CN^\gamma)$. LW_2 is the labor wage defined by the GHH utility function in this paper, and $LW_2 = \log(\frac{Y}{N}) - \log(N^\gamma)$.

follows

$$z_{it} = a_i + z_t. \tag{4.32}$$

That is, the productivity in each island equals the sum of the island specific productivity and the aggregate TFP. Note that household heads now can observe their trading partners' productivity perfectly.

We set the same exogenously determined parameters as before and calibrate the endogenously determined parameters to the same targets. As can be seen in Table 4, the two models have similar performances in explaining the volatility of consumption and investment. The model with confidence shocks is more successful in accounting for the volatility of labor, a variable that the RBC model has difficulty matching. The RBC model with TFP shocks outperforms the model with confidence shocks in explaining the Solow reisdual and labor productivity, but this is due to the exogenously assumed TFP shock process.

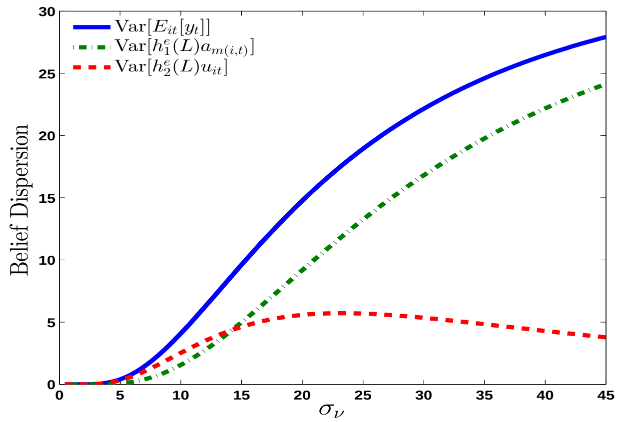
As emphasized by Chari, Kehoe, and McGrattan (2007), standard RBC models fail to capture the pattern of labor wedges. In our model with confidence shocks, the labor wedge is highly countercyclical. The reason is that agents increase or decrease their labor supply not because there is a real change in labor productivity, but because they believe the demand from other islands is high thanks to information frictions. The confidence shock serves as a wedge between labor productivity and the marginal rate of substitution.

Comparison with the Heterogeneous-Prior Formulation To compare with the heterogeneous-prior formulation, we use the baseline model in Angeletos, Collard, and Dellas (2014).[9] In this formulation, the persistence and variance of output are independent of information frictions. With the same confidence shock process, the persistence of various aggregate variables is sympathetically higher than the one in our common-prior formulation. Unlike our common-prior model in which there is an upper bound for the variance of output, one can obtain any variance of output with heterogeneous-prior formulation. The required standard deviation of innovation to the confidence shock is 1.42, which is less than 50% of the one in the common-prior formulation. To capture the effects of information frictions, Angeletos, Collard, and Dellas (2014) choose a relatively low persistence of the confidence shock. Our paper implements this notion by solving the common-prior model and examining whether the forecast errors in the model match the micro data.

4.4 Belief Dispersion

The mean forecast error in the data may be due to information frictions or unexpected stochastic shocks, and the two are difficult to separate. Another way of examining the degree of information frictions is to look at the dispersion of output forecasts among agents, and we label it as belief dispersion. The belief dispersion is increasing in the degree of information frictions, and it is independent of unexpected aggregate stochastic shocks. These two features make it a better indicator of

Figure 9: Belief Dispersion



information frictions.

In our model, the output forecast on island i at time t is given by

$$\mathbb{E}_{it}[y_t] = h_1^e(L)x_{it}^1 + h_2^e(L)x_{it}^2, \tag{4.33}$$

where $h_1^e(L)$ and $h_2^e(L)$ can be derived using the Wiener filter. The mean forecast in the economy is

$$\overline{\mathbb{E}}_{it}[y_t] = \int \mathbb{E}_{it}[y_t] = (h_1^e(L) + h_2^e(L))\xi_t, \tag{4.34}$$

and the variance of forecasts cross all agents is given by

$$\mathrm{Var}[\mathbb{E}_{it}[y_t]] = \mathrm{Var}[h_1^e(L)a_{m(i,t)}] + \mathrm{Var}[h_2^e(L)u_{it}]. \tag{4.35}$$

As expected, the belief dispersion comes from the dispersion of productivity $a_{m(i,t)}$ and the dispersion of idiosyncratic noise u_{it}. Figure 9 presents how $\mathrm{Var}[\mathbb{E}_{it}[y_t]]$, $\mathrm{Var}[h_1^e(L)a_{m(i,t)}]$, and $\mathrm{Var}[h_1^e(L)a_{m(i,t)}]$ vary with the standard deviation to the innovation of idiosyncratic noise σ_ν. The variance of cross-sectional output forecasts is monotonically increasing in σ_ν. By contrast, the part due to the variance of idiosyncratic noise $\mathrm{Var}[h_1^e(L)a_{m(i,t)}]$ displays a hump-shaped relationship with σ_ν. The reason is that as σ_ν becomes larger, agents also optimally respond less to the second signal.

In the Survery of Professional Forecasters, the variance of 1-quarter ahead output forecasts among all forecasters is 0.20%. Under our small-information-friction calibration, the corresponding variance is 0.22%, which is close to the data counterpart. While under our large-information-friction calibration,

[9]The details of the model specification can be found on our online appendix.

the corresponding variance is 0.47%, and it is much larger than the one in the data. Based on this calculation, we think the small-information-calibration should be taken more seriously as a benchmark in evaluating the quantitative performance of the confidence shock.

5 Conclusion

In this paper, we study a business cycle model in which aggregate fluctuations are driven by confidence shocks. Because of asymmetric information, higher order beliefs are crucial in shaping equilibrium outcomes, and the infinite regress problem arises. We use our method developed in Huo and Takayama (2014) to solve the infinite regress problem without approximation. It turns out that the persistence aggregate output is increasing in the degree of information frictions and strategic complementarity. Also, there is an upper bound for the volatility of output that can be obtained by confidence shocks. In our quantitative model, we calibrate the parameters that determine information frictions to match micro-level data. We find that confidence shocks can account for many salient features of business cycles. These results imply that confidence shocks or other non-fundamental shocks can play an important role in explaining the macro economy. We believe the method and the insights discussed in this paper can also be applied to a broad class of models with higher order beliefs.

References

Abraham, A., and K. White (2006): "The Dynamics of Plant-Level Productivity in U.S. Manufacturing," Working Papers 06-20, Center for Economic Studies, U.S. Census Bureau.

Acharya, S. (2013): "Dispersed Beliefs and Aggregate Demand Management," Working paper.

Angeletos, G.-M., F. Collard, and H. Dellas (2014): "Quantifying Confidence," Working paper.

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.

Bacchetta, P., and E. V. Wincoop (2006): "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," *American Economic Review*, 96(3), 552–576.

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.

Beaudry, P., D. Nam, and J. Wang (2011): "Do Mood Swings Drive Business Cycles and is it Rational?," NBER Working Papers, No. 17651.

Beaudry, P., and F. Portier (2006): "Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 96(4), 1293–1307.

Benhabib, J., P. Wang, and Y. Wen (2014): "Sentiments and Aggregate Demand Fluctuations," Forthcoming, *Econometrica*.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013): "News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, 103(7), 3045–70.

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

Collard-Wexler, A., J. Asker, and J. D. Loecker (2011): "Productivity Volatility and the Misallocation of Resources in Developing Economies," Discussion paper, Working Paper 17175, National Bureau of Economic Research.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Hellwig, C., and V. Venkateswaran (2009): "Setting the right prices for the wrong reasons," *Journal of Monetary Economics*, 56, S57–S77.

——— (2011): "Hayek Vs Keynes: Dispersed Information and Market Prices in a Price-Setting Model," Discussion paper, Working Paper.

Huo, Z., and J.-V. Rios-Rull (2014): "Tightening Financial Frictions on Households, Recessions, and Price Reallocations," Forthcoming, *Review of Economic Dynamics*.

Huo, Z., and N. Takayama (2014): "Rational Expectations Models with Higher Order Beliefs," Manuscript, University of Minnesota.

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

Kasa, K. (2000): "Forecasting the Forecasts of Others in the Frequency Domain," *Review of Economic Dynamics*, 3(4), 726–756.

Keynes, J. M. (1936): *The General Theory of Employment Interest and Money*. Fondo de Cultura Economica.

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

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

——— (2011): "Dynamic higher order expectations," Economics Working Papers 1118, Department of Economics and Business, Universitat Pompeu Fabra.

Ríos-Rull, J.-V., and R. Santaeulalia-Llopis (2010): "Redistributive Shocks and Productivity Shocks," *Journal of Monetary Economics*, 57(8), 931–948.

Rondina, G., and T. B. Walker (2013): "Information Equilibria in Dynamic Economies," Working paper.

Schmitt-Grohé, S., and M. Uribe (2012): "What's News in Business Cycles," *Econometrica*, 80(6), 2733–2764.

Townsend, R. M. (1983): "Forecasting the Forecasts of Others," *Journal of Political Economy*, 91(4), 546–88.

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

Appendix

A Proof of Theorems and Propositions

A.1 Proof of Proposition 2.1

Proof. Let j denote $m(i,t)$. With the optimal output rule (2.7), successive iteration leads to

$$
\begin{aligned}
y_{it} &= \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}\left[y_{jt}\right] \\
&= \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}\left[\alpha_0 a_j + \alpha_1 \mathbb{E}_{jt}\left[y_{it}\right]\right] \\
&= \alpha_0 a_i + \alpha_0\alpha_1 \mathbb{E}_{it}\left[a_j\right] + \alpha_1^2 \mathbb{E}_{it}\mathbb{E}_{jt}[y_{it}] \\
&= \alpha_0 a_i + \alpha_0\alpha_1 \mathbb{E}_{it}\left[a_j\right] + \alpha_1^2 \mathbb{E}_{it}\mathbb{E}_{jt}[\alpha_0 a_i + \alpha_1 \mathbb{E}_{it}\left[y_{jt}\right]] \\
&= \alpha_0 a_i + \alpha_0\alpha_1^2 \mathbb{E}_{it}\mathbb{E}_{jt}[a_i] + \alpha_0\alpha_1 \mathbb{E}_{it}\left[a_j\right] + \alpha_1^3 \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}[y_{jt}] \\
&= \alpha_0 a_i + \alpha_0\alpha_1^2 \mathbb{E}_{it}\mathbb{E}_{jt}[a_i] + \alpha_0\alpha_1 \mathbb{E}_{it}\left[a_j\right] + \alpha_0\alpha_1^3 \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}[a_j] + \alpha_1^4 \mathbb{E}_{it}\mathbb{E}_{jt}\mathbb{E}_{it}\mathbb{E}_{jt}[y_{it}] \\
&\;\vdots \\
&= \alpha_0 \sum_{k=0}^{\infty} \alpha_1^{2k} \mathbb{E}_{it}^{2k}[a_i] + \alpha_0 \sum_{k=0}^{\infty} \alpha_1^{2k+1} \mathbb{E}_{it}^{2k+1}[a_j].
\end{aligned}
$$

Given that $\alpha_1 \in (0,1)$ and the modulus of the expectation is bounded from above, the summation in the last line is well defined. The expectation operator $\mathbb{E}_{it}^k$ stands for higher order beliefs and is given by

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

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

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

More compactly,

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

The the output in island i is

$$\begin{aligned}
y_{it} &= \alpha_0 \sum_{k=0}^{\infty} \alpha_1^{2k} \mathbb{E}_{it}^{2k}[a_i] + \alpha_0 \sum_{k=0}^{\infty} \alpha_1^{2k+1} \mathbb{E}_{it}^{2k+1}[a_j] \\
&= \frac{\alpha_0}{1-\alpha_1^2} a_i + \frac{\alpha_0 \alpha_1}{1-\alpha_1^2} x_{it}^1 - \frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k \mathbb{E}_{it}^k [\xi_t] \\
&= \frac{\alpha_0}{1-\alpha_1^2} a_i + \frac{\alpha_0 \alpha_1}{1-\alpha_1^2} a_j + \frac{\alpha_0 \alpha_1}{1-\alpha_1^2} \xi_t - \frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k \mathbb{E}_{it}^k [\xi_t] \\
&= \frac{\alpha_0}{1-\alpha_1^2} a_i + \frac{\alpha_0 \alpha_1}{1-\alpha_1^2} a_j + \frac{\alpha_0}{1+\alpha_1} \sum_{k=1}^{\infty} \alpha_1^k (\xi_t - \mathbb{E}_{it}^k [\xi_t])
\end{aligned}$$

□

A.2 Proof of Theorem 1

Proof. The signal process in our simple economy can be written as

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

where we have normalized the shock process to be with unit variance. By the Canonical Factorization Theorem discussed in Huo and Takayama (2014), the matrices for the fundamental representation are

$$B(z) \quad = \quad \frac{1}{1-\rho z} \begin{bmatrix} 1 - \frac{\tau_1 \rho + \lambda \tau_2}{\tau_1 + \tau_2} z & \frac{\tau_1 \rho - \lambda \tau_1}{\tau_1 + \tau_2} z \\ \frac{\tau_2 \rho - \lambda \tau_2}{\tau_1 + \tau_2} z & 1 - \frac{\tau_2 \rho + \lambda \tau_1}{\tau_1 + \tau_2} z \end{bmatrix},$$

$$V^{-1} \quad = \quad \frac{1}{\rho(\tau_1 + \tau_2)} \begin{bmatrix} \frac{\tau_1 \rho + \lambda \tau_2}{\tau_1} & \lambda - \rho \\ \lambda - \rho & \frac{\tau_2 \rho + \lambda \tau_1}{\tau_2} \end{bmatrix},$$

where $\tau_1 = \frac{\sigma_a^2}{\sigma_\eta^2}$ and $\tau_2 = \frac{\sigma_u^2}{\sigma_\eta^2}$. τ_1 and τ_2 are the relative variance of idiosyncratic shocks to the confidence shock.[10] λ is given by

$$\lambda \quad = \quad \frac{1}{2} \left[\frac{\tau_1 + \tau_2}{\rho \tau_1 \tau_2} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_1 + \tau_2}{\rho \tau_1 \tau_2} + \frac{1}{\rho} + \rho \right)^2 - 4} \right].$$

In equilibrium

$$y_{it} = \alpha_0 a_i + \alpha_1 \mathbb{E}_{it}[y_{m(i,t)t}].$$

[10] Since we assume $\sigma_\eta = 1$, it follows that $\tau_1 = \sigma_a^2$ and $\tau_2 = \sigma_u^2$.

We are looking for policy rule

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

such that the equilibrium condition is satisfied. To predict $y_{m(i,t)t}$, it is equivalent to forecast

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

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. The Wiener-Hopf prediction formula gives

$$\mathbb{E}_{it}[a_{m(i,t)}] = \frac{1}{1-\lambda L}\begin{bmatrix} \frac{\tau_1\rho+\tau_2\lambda}{\rho(\tau_1+\tau_2)} - \lambda L \\ \frac{\tau_1(\lambda-\rho)}{\rho(\tau_1+\tau_2)} \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix},$$

$$\mathbb{E}_{it}\left[\frac{h_1(L)+h_2(L)}{1-\rho L}\eta_t\right] = \frac{1}{1-\lambda L}\begin{bmatrix} \frac{\lambda}{\rho\tau_1(L-\lambda)}\left(L[h_1(L)+h_2(L)] - \lambda[h_1(\lambda)+h_2(\lambda)]\frac{1-\rho L}{1-\rho\lambda}\right) \\ \frac{\lambda}{\rho\tau_2(L-\lambda)}\left(L[h_1(L)+h_2(L)] - \lambda[h_1(\lambda)+h_2(\lambda)]\frac{1-\rho L}{1-\rho\lambda}\right) \end{bmatrix} \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix}.$$

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

$$\begin{aligned}
& h_a a_i + h_1(z)x_{it}^1 + h_2(L)x_{it}^2 \\
= \;& \alpha_0 a_i \\
+ \;& \alpha_1 h_a \begin{bmatrix} \frac{\frac{\tau_1\rho+\tau_2\lambda}{\rho(\tau_1+\tau_2)} - \lambda L}{1-\lambda L} \\ \frac{\tau_1(\lambda-\rho)}{\rho(\tau_1+\tau_2)}\frac{1}{1-\lambda L} \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix} + \alpha_1 h_1(0)a_i \\
+ \;& \alpha_1 \begin{bmatrix} \frac{\lambda}{\rho\tau_1}\frac{L}{(1-\lambda L)(L-\lambda)}h_1(L) - \frac{\lambda^2}{\rho\tau_1}\frac{1}{1-\rho\lambda}\frac{1-\rho L}{(1-\lambda L)(L-\lambda)}h_1(\lambda) \\ \frac{\lambda}{\rho\tau_2}\frac{L}{(1-\lambda L)(L-\lambda)}h_1(L) - \frac{\lambda^2}{\rho\tau_2}\frac{1}{1-\rho\lambda}\frac{1-\rho L}{(1-\lambda L)(L-\lambda)}h_1(\lambda) \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix} \\
+ \;& \alpha_1 \begin{bmatrix} \frac{\lambda}{\rho\tau_1}\frac{z}{(1-\lambda L)(L-\lambda)}h_2(L) - \frac{\lambda^2}{\rho\tau_1}\frac{1}{1-\rho\lambda}\frac{1-\rho z}{(1-\lambda L)(L-\lambda)}h_2(\lambda) \\ \frac{\lambda}{\rho\tau_2}\frac{L}{(1-\lambda L)(L-\lambda)}h_2(L) - \frac{\lambda^2}{\rho\tau_2}\frac{1}{1-\rho\lambda}\frac{1-\rho L}{(1-\lambda L)(L-\lambda)}h_2(\lambda) \end{bmatrix}' \begin{bmatrix} x_{it}^1 \\ x_{it}^2 \end{bmatrix}
\end{aligned}$$

By the Reise-Fisher Theorem, the following system in the analytic function space has to be true

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

where

$$C(z) = \begin{bmatrix} 1 - \alpha_1\frac{\lambda}{\rho\tau_1}\frac{z}{(1-\lambda z)(z-\lambda)} & -\alpha_1\frac{\lambda}{\rho\tau_1}\frac{z}{(1-\lambda z)(z-\lambda)} \\ -\alpha_1\frac{\lambda}{\rho\tau_2}\frac{z}{(1-\lambda z)(z-\lambda)} & 1 - \alpha_1\frac{\lambda}{\rho\tau_2}\frac{z}{(1-\lambda z)(z-\lambda)} \end{bmatrix}$$

$$d(z) \quad = \quad \begin{bmatrix} h_a\alpha_1 \frac{\frac{\tau_1\rho+\tau_2\lambda}{\rho(\tau_1+\tau_2)}-\lambda z}{1-\lambda z} - \alpha_1 \frac{\lambda^2}{\rho\tau_1}\frac{1}{1-\rho\lambda}\frac{1-\rho z}{(1-\lambda z)(z-\lambda)}[h_1(\lambda)+h_2(\lambda)] \\ h_a\alpha_1 \frac{\tau_1(\lambda-\rho)}{\rho(\tau_1+\tau_2)}\frac{1}{1-\lambda z} - \alpha_1 \frac{\lambda^2}{\rho\tau_2}\frac{1}{1-\rho\lambda}\frac{1-\rho z}{(1-\lambda z)(z-\lambda)}[h_1(\lambda)+h_2(\lambda)] \end{bmatrix}$$

To solve for $h_1(z)$ and $h_2(z)$, we use Cramer's rule, which requires the determinant of $C(z)$.

$$\begin{aligned} \det C(z) &= 1 - \alpha_1 \left[\frac{\lambda(\tau_1+\tau_2)}{\rho\tau_1\tau_2} \frac{z}{(1-\lambda z)(z-\lambda)} \right] \\ &= \frac{\rho\tau_1\tau_2(1-\lambda z)(z-\lambda) - \alpha_1\lambda(\tau_1+\tau_2)z}{\rho\tau_1\tau_2(1-\lambda z)(z-\lambda)} \\ &= \frac{-\lambda\left[z^2 - \left(\frac{1}{\lambda}+\lambda-\frac{\alpha_1(\tau_1+\tau_2)}{\rho\tau_1\tau_2}\right)z+1\right]}{(1-\lambda z)(z-\lambda)}. \end{aligned}$$

The determinant of $C(z)$ has two roots which are reciprocal for each other. The inside root is

$$\vartheta = \frac{\left(\frac{1}{\rho}+\rho+\frac{(1-\alpha_1)(\tau_1+\tau_2)}{\rho\tau_1\tau_2}\right) - \sqrt{\left(\frac{1}{\rho}+\rho+\frac{(1-\alpha_1)(\tau_1+\tau_2)}{\rho\tau_1\tau_2}\right)^2 - 4}}{2}$$

Therefore

$$\det C(z) = \frac{\frac{\lambda}{\vartheta}(z-\vartheta)(1-\vartheta z)}{(1-\lambda z)(z-\lambda)}$$

Using Cramer's rule,

$$h_1(z) \quad = \quad \frac{\det \begin{bmatrix} d_1(z) & -\alpha_1 \frac{\lambda}{\rho\tau_1}\frac{z}{(1-\lambda z)(z-\lambda)} \\ d_2(z) & 1-\alpha_1\frac{\lambda}{\rho\tau_2}\frac{z}{(1-\lambda z)(z-\lambda)} \end{bmatrix}}{\det C(z)}$$

To make sure $h_1(z)$ does not have poles in the unit circle, we need to choose $h_1(\lambda)+h_2(\lambda)$ to remove the pole at ϑ, which requires

$$\det \begin{bmatrix} d_1(\vartheta) & -\alpha_1 \frac{\lambda}{\rho\tau_1}\frac{\vartheta}{(1-\lambda\vartheta)(\vartheta-\lambda)} \\ d_2(\vartheta) & 1-\alpha_1\frac{\lambda}{\rho\tau_2}\frac{\vartheta}{(1-\lambda\vartheta)(\vartheta-\lambda)} \end{bmatrix} = 0$$

Note that evaluating z at ϑ, we have

$$d_1(\vartheta) + d_2(\vartheta) = 0.$$

We can then solve for $h_1(\lambda)+h_2(\lambda)$ as a function of h_a.

$$h_1(\lambda)+h_2(\lambda) = \frac{h_a(\vartheta-\lambda)\left(\frac{\lambda}{\rho}-\lambda\vartheta\right)}{\frac{\lambda^2}{\rho}\frac{1}{1-\rho\lambda}(1-\rho\vartheta)(\frac{1}{\tau_1}+\frac{1}{\tau_2})} = \frac{h_a(\vartheta-\lambda)(1-\rho\lambda)\tau_1\tau_2}{\lambda(\tau_1+\tau_2)}$$

Using this result, it follows that

$$\det \begin{bmatrix} d_1(z) & -\alpha_1 \frac{\lambda}{\rho\tau_1} \frac{z}{(1-\lambda z)(z-\lambda)} \\ \\ d_2(z) & 1-\alpha_1 \frac{\lambda}{\rho\tau_2} \frac{z}{(1-\lambda z)(z-\lambda)} \end{bmatrix}$$
$$= \frac{1}{(1-\lambda z)(z-\lambda)} \alpha_1 h_a^y(-\lambda)(z-\vartheta)\left(z - \frac{\rho\tau_1 + \vartheta\tau_2}{(\tau_1+\tau_2)\vartheta\rho}\right).$$

Therefore,

$$h_1(z) = \frac{\alpha_1 h_a \vartheta \left(\frac{\rho\tau_1+\vartheta\tau_2}{(\tau_1+\tau_2)\vartheta\rho} - z\right)}{1-\vartheta z}.$$

Similarly, we can solve for $h_2(z)$ as

$$h_2(z) = -\frac{\alpha_1 h_a \frac{\tau_1(\rho-\theta)}{\rho(\tau_1+\tau_2)}}{1-\vartheta z}.$$

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

$$h_a = \alpha_0 + \alpha_1 h_1(0) = \alpha_0 + \alpha_1^2 h_a \frac{\rho\tau_1 + \vartheta\tau_2}{(\tau_1+\tau_2)\rho} = \frac{\alpha_0}{1-\alpha_1^2 \frac{\rho\tau_1+\vartheta\tau_2}{(\tau_1+\tau_2)\rho}}.$$

□

A.3 Proof of Theorem 2

Proof. Let $\phi = \{\phi_a, \phi_1, \phi_2, \phi_3\} \in \mathbb{R} \times \ell^2 \times \ell^2 \times \ell^2$. The norm of ϕ can de defined as

$$\|\phi\| = \sqrt{\sigma_a^2\phi_a^2 + \sigma_a^2 \sum_{i=0}^{\infty} \phi_{1i}^2 + \sigma_u^2 \sum_{i=0}^{\infty} \phi_{2i}^2 + \sigma_\eta^2 \sum_{i=0}^{\infty} \phi_{3i}^2}.$$

Given an arbitrary ϕ, let

$$\Phi(L) = \phi_3(L)$$

Then the signal process is well defined.Let

$$y_{it}^{\phi} = \phi_a a_i + \phi_1(L) a_{m(i,t)} + \phi_2(L) u_{it} + \phi_3(L)\eta_t,$$

and the optimal linear forecast is given by

$$\mathbb{E}_{it}[y_{m(i,t)t}^{\phi}] = \widehat{\phi}_a a_i + \widehat{\phi}_1(L) a_{m(i,t)} + \widehat{\phi}_2(L) u_{it} + \widehat{\phi}_3(L)\eta_t.$$

If $y_{it}^{\phi} = \alpha_1 a_i + \alpha_1 \mathbb{E}_{it}[y_{m(i,t)t}^{\phi}]$, then ϕ and Φ consist an equilibrium.

Define the operator $\mathcal{T} : \mathbb{R} \times \ell^2 \times \ell^2 \times \ell^2 \to \mathbb{R} \times \ell^2 \times \ell^2 \times \ell^2$ as

$$\mathcal{T}(\phi) = \mathcal{T}(\{\phi_a, \phi_1, \phi_2, \phi_3\}) = (\{\alpha_0 + \alpha_1\widehat{\phi}_a, \alpha_1\widehat{\phi}_1, \alpha_1\widehat{\phi}_2, \alpha_1\widehat{\phi}_3\}).$$

The equilibrium is a fixed point of the operator $\mathcal{T}$. If we can show that $\mathcal{T}$ is a contraction mapping, it is sufficient to prove the theorem.

Let $\phi \in \mathbb{R} \times \ell^2 \times \ell^2 \times \ell^2$ and $\psi \in \mathbb{R} \times \ell^2 \times \ell^2 \times \ell^2$. The distance between ϕ and ψ is

$$\|\phi - \psi\| = \sqrt{\sigma_a^2(\phi_a - \psi_a)^2 + \sigma_a^2 \sum_{i=0}^{\infty}(\phi_{1i} - \psi_{1i})^2 + \sigma_u^2 \sum_{i=0}^{\infty}(\phi_{2i} - \psi_{2i})^2 + \sigma_\eta^2 \sum_{i=0}^{\infty}(\phi_{3i} - \psi_{3i})^2}.$$

The distance between $\mathcal{T}(\phi)$ and $\mathcal{T}(\psi)$ is

$$\|\mathcal{T}(\phi) - \mathcal{T}(\psi)\| = |\alpha_1| \sqrt{\sigma_a^2(\widehat{\phi}_a - \widehat{\psi}_a)^2 + \sigma_a^2 \sum_{i=0}^{\infty}(\widehat{\phi}_{1i} - \widehat{\psi}_{1i})^2 + \sigma_u^2 \sum_{i=0}^{\infty}(\widehat{\phi}_{2i} - \widehat{\psi}_{2i})^2 + \sigma_\eta^2 \sum_{i=0}^{\infty}(\widehat{\phi}_{3i} - \widehat{\psi}_{3i})^2}.$$

Note that the variance of a variable is always larger than the variance of its predictor

$$\begin{aligned}
&\mathrm{Var}[y^{\phi-\psi}_{m(i,t)t}] \\
=&\mathrm{Var}[(\phi_a - \psi_a)a_{m(i,t)} + (\phi_1(L) - \psi_1(L))a_{m(m(i,t),t)} + (\phi_2(L) - \psi_2(L))u_{m(i,t)t} + (\phi_3(L) - \psi_3(L))\eta_t] \\
=&\sigma_a^2(\phi_a - \psi_a)^2 + \sigma_a^2 \sum_{i=0}^{\infty}(\phi_{1i} - \psi_{1i})^2 + \sigma_u^2 \sum_{i=0}^{\infty}(\phi_{2i} - \psi_{2i})^2 + \sigma_\eta^2 \sum_{i=0}^{\infty}(\phi_{3i} - \psi_{3i})^2 \\
=&\|\phi - \psi\|^2 \\
\\
\geq&\mathrm{Var}[\mathbb{E}_{it}[y^{\phi-\psi}_{jt}]] \\
=&\mathrm{Var}[(\widehat{\phi}_a - \widehat{\psi}_a)a_i + (\widehat{\phi}_1(L) - \widehat{\psi}_1(L))a_{m(i,t)} + (\widehat{\phi}_2(L) - \widehat{\psi}_2(L))u_{it} + (\widehat{\phi}_3(L) - \widehat{\psi}_3(L))\eta_t] \\
=&\sigma_a^2(\widehat{\phi}_a - \widehat{\psi}_a)^2 + \sigma_a^2 \sum_{i=0}^{\infty}(\widehat{\phi}_{1i} - \widehat{\psi}_{1i})^2 + \sigma_u^2 \sum_{i=0}^{\infty}(\widehat{\phi}_{2i} - \widehat{\psi}_{2i})^2 + \sigma_\eta^2 \sum_{i=0}^{\infty}(\widehat{\phi}_{3i} - \widehat{\psi}_{3i})^2 \\
=&\|\mathcal{T}(\phi) - \mathcal{T}(\psi)\|^2 \frac{1}{|\alpha_1|^2}.
\end{aligned}$$

Therefore, $\|\mathcal{T}(\phi) - \mathcal{T}(\psi)\| \leq \alpha_1 \|\phi - \psi\|$ when $\alpha_1 \in (0, 1)$. The operator $\mathcal{T}$ is a contraction mapping. There exists a unique fixed point. □

A.4 Proof of Proposition 4.1

Proof. Let $m(i, t)$ be island i's partner at time t and $a_{m(i,t)}$ be its productivity. We want to guarantee that there exists stochastic process such that, for all $i \in [0, 1)$,

$$\begin{aligned}
a_{m(i,t)} &= \rho a_{m(i,t-1)} + \epsilon_t, \\
\epsilon_t &\sim N(0, \sigma^2)
\end{aligned}$$

where $\rho \in (0, 1)$.

Without loss of generality, we can assume that at some t every island $x \in [0, \frac{1}{2})$ meets an island $m(x, t) = x + \frac{1}{2}$ and vice versa. Define a shift operator as

$$a \oplus \mathrm{b} \quad \equiv \quad a - \frac{1}{2} + b - \frac{1}{2}\left\lfloor 2(a - \frac{1}{2} + b)\right\rfloor,$$

where $\lfloor c \rfloor$ is the largest integer not exceeding c. Then, for all $n \in Z_+$, for all $x \in [0, \frac{1}{2})$, let

$$m(x, t + n + 1) \quad = \quad m(x, t + n) \oplus \Delta,$$

where $\Delta \in R$, and $\Delta \notin Q$. As for $x \in [\frac{1}{2}, 1)$, vice versa. In a discrete analog with countably infinite islands, the next partner island is obvious e.g. its neighbor to the left or right. Here, however, there is no naturally next number to x, and hence we need to guarantee that there exists a step size Δ such that, for all $x \in [\frac{1}{2}, 1)$,

$$a_{x\oplus\Delta} - \rho a_x \quad \sim \quad N(0, \sigma^2),$$

and similarly for $x \in [0, \frac{1}{2})$. This is not an obvious task.

Now, there exists an Ornstein-Uhlenbeck process $\{Z_x\}$ obeying

$$\begin{aligned} dZ_x &= -\hat{\rho} Z_x + \hat{\sigma} dW_x, \\ Cov[Z_y, Z_x] &= \frac{\hat{\sigma}^2}{2\hat{\rho}} \exp(-\hat{\rho}|y - x|), \end{aligned}$$

where $\{W_x\}$ is the Wiener process and its discrete analog (an AR(1) process) is written as

$$\begin{aligned} z_n &= \kappa_N z_{n-1} + \sqrt{1 - \kappa_N^2}\hat{\epsilon}_n, \\ \kappa_N &= \exp(-\frac{\hat{\rho} X}{N}), \\ \hat{\epsilon}_n &\sim N(0, \frac{\hat{\sigma}^2}{2\hat{\rho}}), \end{aligned}$$

where $n = 1, ..., N$ and N is a large number[11]. Then, let

$$\begin{aligned} X &= \frac{1}{2} \\ \Delta &= \frac{X}{N}, \\ \rho &= \kappa_N, \\ \sigma &= \hat{\sigma}\sqrt{\frac{\Delta(\rho^2 - 1)}{2 \log \rho}}. \end{aligned}$$

It follows that

$$\begin{aligned} z_n &= \rho z_{n-1} + \sqrt{1 - \rho^2}\hat{\epsilon}_n, \\ &= \rho x_{n-1} + \epsilon_n, \end{aligned}$$

[11] Finch, Steven (2004) "Ornstein-Uhlenbeck Process, " mimeo.

and this can be interpreted as a discrete analog of a_x. The corresponding Ornstein-Uhlenbeck process is rewritten as

$$dZ_x = \frac{\log \rho}{\Delta} Z_x + \sigma \sqrt{\frac{2 \log \rho}{\Delta(\rho^2 - 1)}} dW_x,$$

and hence

$$Cov[Z_{x+\Delta}, Z_x] = \frac{\rho \sigma^2}{1 - \rho^2}.$$

Note this is identical to the first auto-correlation of the discrete analog and W_x is normally distributed so is the sum of innovation of Z_x between $x + \Delta$ and x. Therefore, if we assume $a_x = Z_{x-\frac{1}{2}}$ for $x \in [\frac{1}{2}, 1)$ and similarly for $x \in [0, \frac{1}{2})$ (with another identical stochastic process), the step size we want is Δ, given no wrap-around happens at $x = 1$, and the wrap-around can be ignored when $\Delta \to 0$. $\square$