Macro-Financial Volatility under Dispersed Information*

Jianjun Miao† Jieran Wu‡ Eric R. Young§

First version: September 2016
This version: May 2020

Abstract

We provide a production-based asset pricing model with dispersed information and small deviations from full rational expectations. In the model, aggregate output and equity prices depend on the higher-order beliefs about aggregate demand and individual stochastic discount factors. We prove that equity price volatility becomes arbitrarily large as the volatility of idiosyncratic shocks diverges to infinity due to the interaction of signal-extraction with idiosyncratic trading decisions, while aggregate output volatility falls. We propose a two-step spectral factorization method that permits closed-form solutions in the frequency domain applicable to a wide range of models with more hidden states than signals. Our model can quantitatively match output and equity volatilities observed in US data.

Keywords: Dispersed Information, Frequency Domain Analysis, Higher-order Beliefs, Asset Pricing, Business Cycles

JEL Classifications: E32, E44, G12, G14.

*We are grateful to Florian Scheuer (editor) and three anonymous referees for their constructive suggestions and comments. We also would like to thank Jess Benhabib, Jim Dolmas, Kenneth Kasa, Giacomo Rondina, Laura Veldkamp, Todd Walker, Pengfei Wang for comments. We also thank seminar and conference participants at the LAEF conference, the Texas Monetary Conference, CICM, and the AFR Summer Institute of Economics and Finance for helpful comments. All errors are the responsibility of the authors.

†Department of Economics, Boston University, miaoj@bu.edu

‡School of Economics, Zhejiang University, jw5ya@zju.edu.cn

§Department of Economics, University of Virginia, and School of Economics, Zhejiang University, ey2d@virginia.edu

1 Introduction

One prominent and persistent puzzle in finance is the observation by Shiller (1981) that aggregate stock prices are too volatile relative to the expected present value of dividends. Another and perhaps more important aspect of this equity volatility puzzle is that macroeconomic quantities – aggregate output, consumption, and dividends – are too smooth relative to equity prices, and their fluctuations are largely disconnected. In actual economies all these quantities are endogenous and respond to the same shocks that drive equity price movements. The goal of our paper is to understand whether a production-based asset pricing model is able to deliver both smooth aggregate quantities and volatile equity prices, without introducing complex (non–stationary) exogenous shocks or nonstandard preferences.[1] We provide a positive answer to this question by developing a model of a dispersed-information island economy along the lines of Angeletos and La'O (2010), extended to include a centralized stock market.

Our model consists of a continuum of islands, each populated by a continuum of identical households and firms. Island-specific total factor productivity (TFP) consists of an aggregate and an idiosyncratic component. However, agents on each island cannot separately observe each component, only their sum. Near-rational households (Hassan and Mertens (2017)) consume an aggregate of island-specific goods and trade shares in an aggregate equity market.

Our main result is that as the idiosyncratic TFP volatility approaches infinity, an endogenous unit root arises in the log-linearized solutions for investors' shareholdings and the aggregate equity price, causing the equity price to become infinitely volatile. This arises despite the fact that the aggregate equity price only responds to aggregate TFP shocks, not idiosyncratic ones. The key is that the response coefficient endogenously varies with the idiosyncratic TFP volatility – higher idiosyncratic volatility increases the sensitivity of equity prices to aggregate shocks due to a feedback loop between the idiosyncratic shareholdings and the aggregate equity price volatility. This theoretical result has an appealing quantitative implication: we can choose a relatively low volatility of aggregate shocks to match the low volatility of aggregate consumption and choose a relatively high volatility of idiosyncratic shocks to match the high volatility of equity prices as in the data.

On the other hand, we find that higher-order expectations under dispersed information always reduce the volatility of business cycle fluctuations in the real economy,

[1] See Campbell (1999) and Cochrane (2011) for surveys on these alternative approaches.

provided production decisions display strategic complementarity. We establish this result by showing that the volatility of aggregate output under full information gives an upper bound for that under dispersed information. The key assumption for this result is that agents in the economy are *informationally small*: the idiosyncratic shock component of the private signal washes out in the aggregate. As a result, higher-order beliefs about aggregate TFP are necessarily less volatile than aggregate TFP itself, as agents do not need to predict the expectations of any individual agent about the real economy but only the market average.

Maintaining dynamic and persistent information frictions is crucial for our results regarding the volatility of financial variables. Persistent differences in information generally lead to the technical problem of "forecasting the forecasts of others" (Townsend (1983)): the state space for the model solution contains an infinite number of higher-order expectations. Solving the model in the time domain then becomes infeasible, so we use frequency-domain methods to circumvent this problem; under some special assumptions, we can even provide analytical characterization of the equilibrium. For general cases, we develop a numerical method to solve the model.

In the log-linearized equilibrium, the equity price is equal to the sum of the discounted average forecast of the individual stochastic discount factors (SDFs) and the discounted average forecast of future dividends. Due to dispersed information, the average forecast of the individual SDFs is not equal to the forecast of the average SDFs. As a result, variation in the distribution of individual consumption matters for equity prices. Since individual shareholdings and labor supply affect individual consumption and SDFs, their responses to idiosyncratic shocks affect equity volatility. As a result the effect on the equity price is different from the effect on aggregate output, which depends on the average forecast of aggregate demand instead of individual behavior.

As noted above, agents on each island observe only their island-specific TFP and the equity price, but cannot disentangle idiosyncratic and aggregate TFP movements. To understand why this environment leads to highly volatile equity prices, suppose that the volatility of the idiosyncratic component of TFP is arbitrarily large and agents only observe the island-specific TFP (that is, for now they ignore the price when forming expectations). If aggregate productivity increases unexpectedly, each agent misinterprets this change in TFP as an unexpected increase in idiosyncratic productivity. As a result, any particular agent increases his shareholdings permanently to smooth consumption as in the full information case (Hall (1978)), while simultaneously believing some other agents will decrease asset demand by the same amount; that is, the belief is that the price will not change. Because all agents make this same mistaken inference, aggregate asset demand rises and market clearing requires the equilibrium price to rise permanently as the idiosyncratic TFP volatility tends to infinity. Correlated estimation errors under dispersed information cause permanent shifts in shareholdings to be transmitted into permanent shifts in equity prices, even if the aggregate TFP shock is independently and identically distributed (IID). In other words, the equity price inherits a unit root from the individual shareholdings.

However, if agents observe the price change, they can infer that the TPF shock must have been at least partly aggregate and hence the unit root in the equity price can be eliminated. To see why, consider the effect of a rising price on asset demand. As the price rises, agents move along their demand curve as shares become more expensive (a standard wealth effect). Moreover, when the aggregate TFP is IID or mean-reverting, agents understand this fact and expect future prices to fall relative to the current price, inducing a decline in asset demand since they will be cheaper to purchase tomorrow. The result is that the price does not inherit the unit root from individual shareholdings, the true state is revealed, and equity price volatility will be low.

To prevent prices from fully revealing the aggregate TFP, we follow Hassan and Mertens (2017) and assume that investors are near rational and make correlated forecast errors.[2] Now consider the agent's learning process in response to the rising price; agents will assign some weight to aggregate productivity rising and some to the correlated error rising. Each agent believes the error does not apply to him, only to everyone else; as a result, they individually believe that the price tomorrow will be higher, increasing current demand. The result is that the price does not mean-revert and again inherits the unit root from idiosyncratic shareholdings.

Note that the previous result is not simply a result of adding more unknown states than signals; while both lead to forecasting errors, only the near-rational shock leads to the right kind of error. For example, suppose that we introduce an asset supply/noise trader shock, as is common in the noisy rational expectations literature. In that case, the rising price will be attributed partly to a rise in the aggregate TFP and partly to a fall in asset supply. But both shocks imply that future prices will fall relative to current prices and agents reduce (shift inward) asset demand, eliminating the unit root. Although agents still cannot distinguish between the two shocks, they understand that these two shocks alter their desired trading decisions in the same way.

[2]Unlike the model of Hassan and Mertens (2017) with physical capital where the equity price is mainly driven by the near-rational errors through amplification, in our model the aggregate TFP shock plays a dominant role in driving the equity price. See Section 2 for a detailed discussion.

We establish our main results by assuming that the common forecast error follows a special process so that the equilibrium can be characterized by analytic rational functions in the frequency domain (ARMA (p,q) processes in the time domain). If we relax this assumption, the equilibrium cannot be characterized by rational functions, so we resort to numerical methods using rational functions as approximations.[3] Our numerical solutions show that equity price volatility increases quickly with idiosyncratic TFP volatility even for a very small common forecast error. Using calibrated parameter values for aggregate and idiosyncratic TFP volatilities, we show that our model can match both the output and equity price volatilities in the data.

Our model mechanisms are different from much of the existing literature on asset pricing under dispersed information in three important ways.[4] First, in our environment with a continuum of informationally small agents, higher-order beliefs dampen aggregate output volatility, but generate large fluctuations in the financial market (as opposed to models with informationally-large agents). Second, most of the literature studies endowment economies in which consumption and dividends are exogenously given; obviously these papers cannot address the issue of why macroeconomic quantities are extremely smooth relative to equity prices. Finally, many papers in this literature assume constant exogenous SDFs, whereas our SDFs are endogenous, heterogeneous, and time-varying; as noted already, this feature is key for our result.

We also make an important methodological contribution by extending the existing literature on models with dispersed information to non-square economies (those with more hidden states than signals). We apply a two-step spectral factorization method from Rozanov (1967) to solve economic problems with non-square signal systems, which is of independent interest. The restriction that the numbers of signals and shocks are the same is quite limited: given a restriction to square systems, the equilibrium will be fully revealing unless there is non-invertibility from signals to shocks as in Rondina and Walker (2015).

Our approach provides an alternative to the state-space approach applied by Huo and Takayama (2018). Their approach is numerically convenient since it can be solved using fast Riccati equation methods and the Kalman filter. The main drawback of that approach is that analytical solutions for the Riccati equations rarely exist. As such, obtaining sharp analytical results like the ones we have becomes infeasible. In

[3]See related results in Makarov and Rytchkov (2012) and Huo and Takayama (2018).

[4]Closely related papers include Kasa, Walker, and Whiteman (2014), Rondina and Walker (2015), Angeletos and La'O (2010, 2013), and Benhabib, Wang, and Wen (2015). See Sections 4 and 5 for a more detailed discussion.

contrast, our approach delivers closed-form solutions in a wider range of models, which may be helpful for illuminating economic intuition and mechanisms. The downside is that, for complicated models, our approach is substantially more burdensome, so we suggest researchers apply our methods in simple models and use the approach of Huo and Takayama (2018) for more complicated ones.

2 Basic Intuition

We use a simple two-period model of an endowment economy to illustrate the basic intuition behind our analysis. There is a continuum of agents indexed by $i \in I = [0, 1]$ who trade a single stock with a unit supply in period 1. The stock pays random dividends D in period 2. Each agent i is endowed with one unit of the stock and random labor income L_i in period 1. He derives utility from consumption C_{i1} and C_{i2} in the two periods according to the function

$$\mathcal{E}_i \left[\frac{C_{i1}^{1-\gamma}}{1-\gamma} + \beta \frac{C_{i2}^{1-\gamma}}{1-\gamma} \right],$$

where $\mathcal{E}_i$ denotes the subjective expectation operator given agent i's information, $\beta \in (0, 1)$ is the subjective discount factor, and γ is the coefficient of relative risk aversion. His budget constraints are given by

$$C_{i1} + QS_i = Q + L_i, \quad C_{i2} = DS_i,$$

where Q and S_i denote the stock price and shareholdings, respectively.

Dividends and labor income satisfy

$$\log D = \log \bar{D} + x_d \varepsilon_a, \;\; \log L_i = \log \bar{L} + x_l \varepsilon_i,$$

where $\bar{D}$, x_d, $\bar{L}$, and x_l are exogenous constants, and ε_a and ε_i are independent normal random variables with means zero and variances σ_a^2 and σ_i^2. The labor income shock is purely idiosyncratic such that $\int_I \epsilon_i di = 0$.

At the beginning of period 1, each agent i receives an exogenous signal $x_i = \epsilon_a + \epsilon_i$, but does not observe ϵ_a and ϵ_i separately. Agents do not communicate their signals to each other. Based on his private signal x_i and the equity price Q, each agent i trades on the stock market. At the end of period 1, labor income realizes and agent chooses consumption C_{i1}. At the beginning of period 2, the random labor income and dividends are realized and agent i chooses consumption C_{i2} out of dividend income. In equilibrium $\int_I S_i di = 1$. It is straightforward to show that the deterministic equilibrium with $\epsilon_a = \epsilon_i = 0$ is given by $\bar{S}_i = 1,\ \bar{C}_{i1} = \bar{L},\ \bar{C}_{i2} = \bar{D},\ \bar{Q} = \beta\left(\bar{L}/\bar{D}\right)^{\gamma}\bar{D}$.

In the stochastic case agent i' utility maximization leads to the Euler equation

$$Q = \mathcal{E}_i\left[M_i D\right].$$

where $M_i = \beta\left(C_{i1}/C_{i2}\right)^{\gamma}$ denotes the stochastic discount factor (SDF). Following Hassan and Mertens (2017), each agent makes a small error when forming his expectation. Specifically, let

$$\mathcal{E}_i\left(\cdot\right) = \mathbb{E}_i\left(\cdot\right)U_i,$$

where $\mathbb{E}_i$ denotes the rational expectation operator conditional on agent i's information $\{x_i, Q\}$ and U_i is a small exogenous error that shifts his conditional expectations. Assume that

$$\log U_i = u + v_i,$$

where u and v_i are independent normal random variables with means zero and variances σ_u^2 and σ_v^2. Here u represents aggregate errors and v_i represents idiosyncratic errors satisfying $\int_I v_i di = 0$. When $U_i = 1$ for all i, agents have full rational expectations.

Introducing near-rational forecast errors in the model injects additional noise into the equity price, which prevents prices from being fully revealing. In the literature, there are many candidate shocks available to serve this purpose, e.g., a noise trader shock to the asset supply. Small deviations from the optimal forecasts will play an important role in the dynamic setting where interactions between the stock price and trading behavior becomes the key to understanding stock price fluctuations.

Now we log-linearize the stochastic equilibrium around the deterministic equilibrium and use a lower case variable to denote its log deviation from its deterministic equilibrium value. We then obtain the log-linearized Euler equation

$$q = \mathbb{E}_i[d] + \mathbb{E}_i[m_i] + u + v_i,\ \ m_i = \gamma(c_{i1} - c_{i2}). \tag{1}$$

Next we substitute the log-linearized budget constraints into the SDF and use the Euler equation to derive the log-linearized trading strategy

$$s_i = \frac{\mathbb{E}_i[(1-\gamma)d] - q + u + v_i}{\gamma(1+\bar{Q}/\bar{L})} + \frac{\mathbb{E}_i[l_i]}{1+\bar{Q}/\bar{L}}. \tag{2}$$

This expression is akin to Merton's (1969) result: the trading strategy consists of a mean-variance efficient component and a hedging component against idiosyncratic labor income.

Aggregating equation (2) over $i \in [0,1]$ and using the log-linearized market-clearing condition $\int_I s_i di = 0$, we obtain

$$q = (1-\gamma)\,\bar{\mathbb{E}}[d] + \gamma\bar{\mathbb{E}}[l_i] + u, \tag{3}$$

where $\bar{\mathbb{E}}\left[\cdot\right] \equiv \int \mathbb{E}_i\left[\cdot\right] di$ denotes the average expectation operator.

To solve the model, we conjecture that the equity price takes the form:

$$q = q_a \epsilon_a + q_u u, \tag{4}$$

where q_a and q_u are nonzero constants to be determined. Then the information set can be normalized to $\{\hat{q}, x_i\}$, where $\hat{q} = \epsilon_a + (q_u/q_a)\,u \equiv \epsilon_a + \hat{u}$. The presence of common forecast errors prevents equity prices to fully reveal the aggregate dividend information.

By the Gaussian projection theorem,

$$\mathbb{E}_i[d] = x_d(\tau_q \hat{q} + \tau_x x_i) \implies \bar{\mathbb{E}}[d] = \tau_q x_d(\epsilon_a + \hat{u}) + \tau_x x_d \epsilon_a, \tag{5}$$

where the noise-to-signal ratios are defined as

$$\tau_q = \frac{\sigma_a^2 \sigma_i^2}{g^2(\sigma_a^2\sigma_u^2 + \sigma_u^2\sigma_i^2) + \sigma_a^2\sigma_i^2} \in (0,1), \;\; \tau_x = \frac{g^2\sigma_a^2\sigma_u^2}{g^2(\sigma_a^2\sigma_u^2 + \sigma_u^2\sigma_i^2) + \sigma_a^2\sigma_i^2} \in (0,1),$$

and $g \equiv q_u/q_a$ will be determined in the equilibrium.

A direct comparison of the two expectations in (5) implies that, if agents are informationally small, the variance of the market average forecast of aggregate fundamentals is smaller than that of the individual forecast in that $Var\left(\bar{\mathbb{E}}[d]\right) < Var\left(\mathbb{E}_i[d]\right)$. It is also easy to check that $Var\left(\mathbb{E}_i[d]\right) < Var\left(d\right)$. We will show that this dampening result applies to our general dynamic model when aggregate fundamentals are endogenous (see Lemma 2). An immediate implication is that dispersed information does not generate large equity price volatility if $\gamma = 0$ (agents are risk neutral). In this case it follows from (3) that equity volatility is bounded by the dividend volatility given the small variations in forecast errors u. Thus we need risk aversion $\gamma > 0$ and hence volatile SDFs.

Consider the second term on the right side of equation (3), which comes from the average forecast of individual SDFs. If agents can communicate with each other so that information is homogenous, this term will vanish: $\bar{\mathbb{E}}[l_i] = \mathbb{E}_i[\int_I l_i di] = 0$. Under dispersed information without communication, we have

$$\mathbb{E}_i[l_i] = x_l\left[-\tau_q \hat{q} + (1-\tau_x)x_i\right] \implies \bar{\mathbb{E}}[l_i] = -\tau_q x_l(\epsilon_a + \hat{u}) + (1-\tau_x)x_l \epsilon_a. \tag{6}$$

A high equity price $\hat{q}$ may be due to a high dividend shock ϵ_a. Agent i may believe the labor income shock ε_i to be low given a fixed signal $x_i = \epsilon_a + \epsilon_i$, which explains the negative coefficient of $\hat{q}$ in (6). Thus learning from prices dampens the effect of idiosyncratic shocks.

Plugging (5) and (6) into (3) yields

$$q = \left[(1-\gamma)x_d\tau_a + \gamma x_l \tau_i\right]\epsilon_a + \left[(1-\gamma)g\tau_q x_d - \gamma g \tau_q x_l + 1\right]u, \tag{7}$$

where $\tau_a = \tau_q + \tau_x \in (0,1)$ and $\tau_i = 1 - (\tau_q + \tau_x)$. Matching coefficients in (4) yields a cubic equation for g :

$$\left[(1-\gamma)x_d\sigma_a^2\sigma_u^2 + \gamma x_l(\sigma_a^2\sigma_u^2 + \sigma_u^2\sigma_i^2)\right]g^3 - \left(\sigma_a^2\sigma_u^2 + \sigma_u^2\sigma_i^2\right)g^2 + \gamma x_l \sigma_a^2\sigma_i^2 g - \sigma_a^2\sigma_i^2 = 0.$$

Substituting (7) into (2) yields the equilibrium trading strategy

$$s_i = \frac{(1-\gamma)x_d\tau_x + \gamma x_l(1-\tau_x)}{\gamma(1+\bar{Q}/\bar{L})}\epsilon_i + \frac{1}{\gamma(1+\bar{Q}/\bar{L})}v_i,$$

which only responds to idiosyncratic labor income shocks and idiosyncratic forecast errors. The presence of idiosyncratic forecast errors prevents shareholdings from fully revealing the idiosyncratic labor income realization.

Equation (7) shows that small common errors in forecasting leads to non-fundamental deviations in the equilibrium stock price, as emphasized by Hassan and Mertens (2017). They also show that small common errors in household expectations weaken the stock market's capacity to aggregate dispersed information. We argue that their results rely on the average forecast of the aggregate shock, whose effect corresponds to the response coefficient τ_a in (7). In contrast, our model features uninsured idiosyncratic labor income shocks and hence the equilibrium equity price also depends on the average forecast of these shocks. This effect corresponds to the response coefficient τ_i in (7).

To relate this result to Hassan and Mertens (2017), we consider the impact of σ_i on τ_a and τ_i, illustrated in Figure 1. The figure shows that τ_a decreases with σ_i as in Hassan and Mertens (2017). However, τ_i increases with σ_i. Intuitively, if agents are unable to distinguish between aggregate and idiosyncratic shocks and make errors in forecasting, the equilibrium price is not fully revealing and agents have to solve a signal extraction problem. If σ_i is higher, the agents put a larger weight on the idiosyncratic labor income shock and a smaller weight on the aggregate dividend shock. However, the additional volatility due to the large idiosyncratic shock only has a limited effect since $\tau_i \in (0,1)$. Even if idiosyncratic shocks are arbitrarily volatile, aggregation cancels them out and τ_i approaches the upper bound of one. Unless we assume a very high value of x_l, the quantitative effect on equity prices will be small.

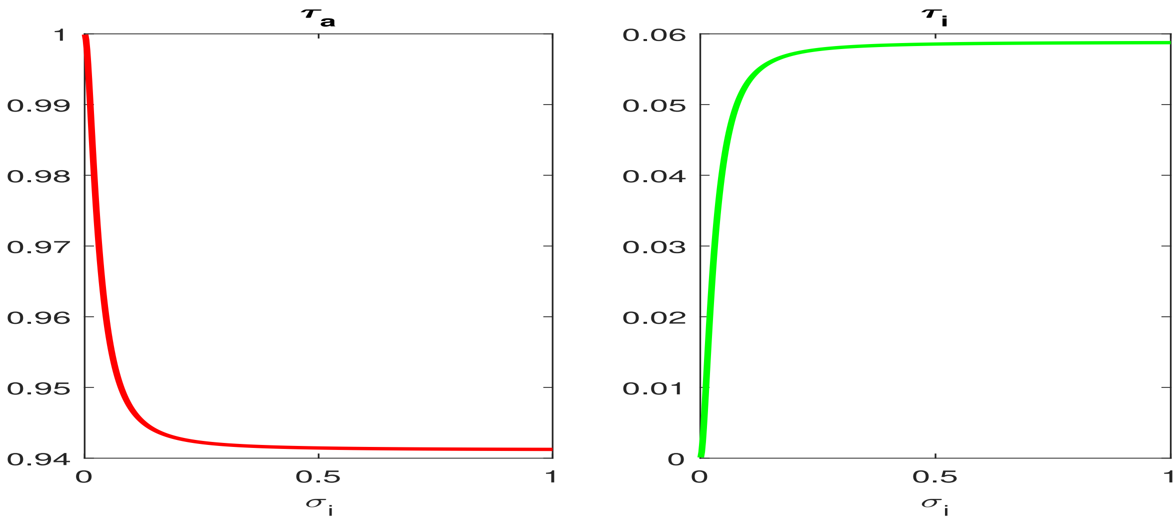

Figure 1: The impact of idiosyncratic volatility σ_i on τ_a and τ_i. Parameter values are $x_l = x_d = 1$, $\gamma = 0.4$, $\sigma_a = 0.1$, and $\sigma_u = 0.01$.

In the next section we extend this simple example to an infinite-horizon setup. We will endogenize labor income and dividends by introducing the production side of the economy so that x_d and x_l are endogenous. In the infinite-horizon model individual SDFs depend on future individual consumption which in turn depends on future trading strategies and labor income. Thus equity prices depend on the higher-order beliefs about the average forecasts of future individual shareholdings and labor income. Interpreted through the lens of the two-period model, this dynamic interaction makes shareholdings and equity prices highly persistent and generates a positive connection between σ_i and x_l that causes equity volatility to increase without bound as $\sigma_i \to \infty$.

3 Model

We consider a variation of the classical dispersed-information business cycle model of Angeletos and La'O (2010). The economy consists of a continuum of islands with a Lebesgue measure over $I = [0, 1]$. Information is dispersed across islands. There is a representative household and a representative firm on each island. Each firm is monopolistically competitive and produces a specialized good using labor input only, while households have Dixit-Stiglitz preferences over varieties. Labor is immobile across islands, but consumption goods of all varieties are freely mobile. The equity market is operated through a mutual fund which owns the firms and issues equity shares to households. The stock price therefore reflects the average valuation of firms in the economy. We normalize the aggregate stock supply to one.

3.1 Households

A representative household on each island $i \in I$ derives utility from the composite consumption good $\{C_{it}\}$ and labor supply $\{N_{it}\}$ according to the utility function of Greenwood, Huffman, and Hercowitz (1988):

$$\mathcal{E}_i \left[\sum_{t=0}^{\infty} \beta^t \log \left(C_{it} - \frac{N_{it}^{1+\phi}}{1+\phi} \right) \right],$$

where $\mathcal{E}_i$ denotes household i's subjective expectation operator, $\beta \in (0,1)$, $\phi > 0$,

$$C_{it} = \left[\int_I C_{it} (j)^{\frac{\varsigma-1}{\varsigma}} dj \right]^{\frac{\varsigma}{\varsigma-1}},$$

and $C_{it}(j)$ denotes the consumption of good j demanded by the household on island i. Here $\varsigma > 1$ denotes the inter-island elasticity of substitution that determines the degree of strategic complementarity.

The household faces the following intertemporal budget constraint

$$\int_I C_{it}(j) P_t(j) dj + Q_t S_{it+1}^h = S_{it}^h \ (Q_t + D_t) + W_{it} N_{it}, \tag{8}$$

where $P_t(j)$, Q_t, S_{it}^h, D_t, and W_{it} represent the price of good j, the stock price, share holdings, aggregate dividends, and the wage rate in island i, respectively.

To simplify the forecasting problem, we assume that each household i consists of two family members, an investor and a shopper. They have different information sets and do not communicate with each other. In each period t the investor's information set consists of the current and past TFP shocks A_{it}, wages W_{it}, and stock prices Q_t. Given this information set, the investor chooses labor supply and shareholdings.[5] The first-order conditions are given by

$$W_{it} = N_{it}^{\phi}, \tag{9}$$

[5] We can introduce bonds (or other assets) into the model and/or allow agents to observe additional signals (such as bond prices and trading volume). In these cases, we would need to insert more unobserved shocks to prevent information revelation, but the main results of our paper would survive, just with an attendant increase in algebraic and computational burden.

$$\mathcal{E}_{it}\left[M_{it+1}\left(Q_{t+1}+D_{t+1}\right)\right]=Q_t, \tag{10}$$

where the SDF M_{it+1} is given by

$$M_{it+1}=\beta\frac{\left(C_{it}-N_{it}^{1+\phi}/\left(1+\phi\right)\right)}{C_{it+1}-N_{it+1}^{1+\phi}/\left(1+\phi\right)}.$$

Our chosen utility function implies that the labor supply in (9) is independent of C_{it} and hence simplifies our analysis, but it is not crucial for our main results (see Appendix A).

As in the two–period example, we assume that investors are near rational and each investor i's subjective expectations satisfy $\mathcal{E}_{it}\left[\cdot\right]=\mathbb{E}_{it}\left[\cdot\right]U_{it}$, where $\mathbb{E}_{it}$ denotes the rational expectation operator conditional on the investor's information at time t and U_{it} is a small exogenous error that shifts the subjective conditional expectations. Let U_{it} satisfy

$$\log U_{it}=u_t+v_{it}, \tag{11}$$

where the aggregate component u_t satisfies

$$u_t=u(L)\epsilon_{ut}, \tag{12}$$

and the idiosyncratic component satisfies

$$\int_I v_{it}di=0. \tag{13}$$

Here $u(L)$ is a square-summable, one-sided lag polynomial and ε_{ut} and v_{it} are independent Gaussian white noises with variances σ_u^2 and σ_v^2.

The shopper collects dividends D_t and purchases consumption good $C_{it}\left(j\right)$ after observing the product prices $P_t\left(j\right)$ for all j and the aggregate price level P_t. The shopper does not face a forecasting problem. The first-order condition is

$$C_{it}\left(j\right)=\left[\frac{P_t\left(j\right)}{P_t}\right]^{-\varsigma}C_{it}, \tag{14}$$

where the aggregate price index $P_t\equiv\left[\int_I P_t\left(j\right)^{1-\varsigma}\ di\right]^{\frac{1}{1-\varsigma}}$ satisfies$\int_I C_{it}\left(j\right)P_t\left(j\right)dj=P_tC_{it}$. We normalize the price index P_t to one so that the budget constraint (8) becomes

$$C_{it}+Q_tS_{it+1}^h=S_{it}^h\ \left(Q_t+D_t\right)+W_{it}N_{it}. \tag{15}$$

Aggregating (14) over $i\in I$ yields the total demand for good $j\in[0,1]$,

$$Y_{jt}=\int_I C_{it}\left(j\right)di=\left[P_t\left(j\right)\right]^{-\varsigma}Y_t, \tag{16}$$

where Y_t denotes aggregate consumption

$$Y_t=\int_I C_{it}di\equiv C_t. \tag{17}$$

3.2 Firms

The representative firm on island $i \in [0,1]$ operates a production technology given by

$$Y_{it} = A_{it} N_{it}^{\alpha},\ \alpha \in (0,1)\,, \tag{18}$$

where A_{it} satisfies

$$A_{it} = A_t \exp\left(\epsilon_{it}\right). \tag{19}$$

Here A_t represents the aggregate component that affects all firms in all islands and ϵ_{it} represents the idiosyncratic component that is independent of A_t and affects the firm in island i only. Investors on island i observe A_{it} at time t, but cannot distinguish between the aggregate and idiosyncratic components. Let

$$\log A_t = a(L)\epsilon_{at}, \tag{20}$$

where ϵ_{at} and ϵ_{it} are independent Gaussian white noises with variances σ_a^2 and σ_i^2, respectively. They are also independent of near-rational shocks u_t and v_{it}. Here $a(L)$ denotes a one-sided, square-summable lag polynomial. Moreover, assume that the law of large number (LLN) holds for ϵ_{it} so that

$$\int_I \epsilon_{it} di = 0. \tag{21}$$

In each period t the firm's information set consists of the current and past TFP shocks A_{it}, wages W_{it}, and stock prices Q_t. Given this information set the firm chooses labor demand to solve the static profit maximization problem

$$\pi_{it} = \max_{N_{it}}\ \mathbb{E}_{it}\left[P_t\left(i\right)\right] Y_{it} - W_{it} N_{it}$$

subject to the demand schedule in (16) for $j = i$. Since the production and labor demand choice is made before observing the output price $P_t\left(i\right)$, the firm needs to form static conditional expectations about the price $P_t\left(i\right)$. Since Y_{it} and N_{it} are observable choice variables, the firm essentially forms conditional expectations about the aggregate demand Y_t. Simple algebra yields the labor demand condition

$$\alpha\left(1 - \frac{1}{\varsigma}\right) \frac{Y_{it}^{\left(1-\frac{1}{\varsigma}\right)} \mathbb{E}_{it}\left[Y_t^{\frac{1}{\varsigma}}\right]}{N_{it}} = W_{it}. \tag{22}$$

For simplicity we assume that firms are fully rational and do not make forecasting errors. Introducing forecasting errors affects profits π_{it} and hence dividends and stock prices. The firms' forecasting errors would therefore play a similar role to the households' forecasting errors in equation (10).

It follows from equations (18) and (22) that observing the local wage W_{it} is equivalent to observing the local productivity shock A_{it}. Thus we can write the information set in the conditional expectation operators $\mathbb{E}_{it}$ and $\mathcal{E}_{it}$ as $\{X_{i,t-k}\}_{k=0}^{\infty}$, where the signal vector is $X_{it} = [A_{it}, Q_t]^{\intercal}$.

3.3 Equilibrium Characterization in the Time Domain

There is one aggregate mutual fund that issues equity shares and collects dividends from individual islands. The aggregate dividend satisfies $D_t = \int_I \pi_{it} di$ and aggregate output satisfies $Y_t = \int_I Y_{it} di$. The mutual fund distributes the dividend to households. The market-clearing condition for the stock is given by

$$\int_I S_{it+1}^h \, di = 1, \forall t \tag{23}$$

-

A competitive equilibrium with dispersed information is characterized by a system of 9 equations (9), (10), (14), (15), (16), (17), (18), (22), and (23) for 9 variables W_{it}, N_{it}, S_{it}^h, C_{it}, $C_{it}(j)$, Y_{it}, $P_t(j)$, Q_t, and Y_t, where D_t satisfies

$$\int_I W_{it} N_{it} di + D_t = Y_t. \tag{24}$$

This equation follows from aggregating (15) using (17) and (23).

Since the equilibrium system is nonlinear and does not admit an explicit solution, we derive a log-linearized approximate system (see Appendix A). We use lower case variables to denote log deviations from the non-stochastic steady state. We impose the following assumption on the parameters so that there exists a unique deterministic steady-state equilibrium.

Assumption 1 *The parameter values satisfy $\alpha, \beta \in (0,1)$, $\phi > 0$, $\varsigma > 1$.*

We first use (9), (18), and (22) to eliminate W_{it} and N_{it} to derive

$$y_{it} = \frac{1}{\xi} a_{it} + \theta \mathbb{E}_{it} [y_t], \tag{25}$$

and

$$y_{it} = a_{it} + \alpha n_{it}, \tag{26}$$

where we define

$$\xi \equiv \frac{1+\phi-\alpha\left(1-1/\varsigma\right)}{1+\phi} > 0,\;\; \theta \equiv \frac{\alpha}{\alpha+\left(1-\alpha+\phi\right)\varsigma} \in \left(0,1\right).$$

The parameter θ describes the degree of strategic complementarity (see Angeletos and La'O (2013) and Huo and Takayama (2018)). Aggregating (25) over $I=[0,1]$, we have

$$y_t = \frac{1}{\xi}\int_I a_{it}di + \theta\overline{\mathbb{E}}_{it}\left[y_t\right], \tag{27}$$

where the average conditional expectation operator is defined as $\overline{\mathbb{E}}_t\left[\cdot\right] \equiv \int_I \mathbb{E}_{it}\left[\cdot\right]di$.

Log-linearizing (10) and (15) yields

$$q_t = \mathbb{E}_{it}\left[m_{it+1}\right] + \mathbb{E}_{it}\left[\beta q_{t+1} + \left(1-\beta\right)d_{t+1}\right] + u_t + v_{it}, \tag{28}$$

where

$$\mathbb{E}_{it}\left[m_{it+1}\right] = \alpha_2 s_{it}^h - \alpha_1 s_{it+1}^h + \mathbb{E}_{it}\left[\alpha_3 s_{it+2}^h + \Delta b_{it+1}\right], \tag{29}$$

and

$$b_{it} = \alpha_4 d_t + \alpha_5 n_{it},\;\; \Delta b_{it+1} \equiv b_{it} - b_{it+1}. \tag{30}$$

Notice that s_{it}^h and s_{it+1}^h are in agent i's information set at time t. Unlike the two-period model, agent i's Euler equation depends on his future consumption so that his expected SDF depends on his forecast of his future shareholdings, labor income, and dividends. Using (9) and (24) we obtain

$$\alpha_6 d_t + \alpha_7 n_t = y_t, \tag{31}$$

where $n_t = \int_I n_{it}di$. Expressions for the coefficients $\alpha_1, \alpha_2, ..., \alpha_7$ can be found in Appendix A. Define the parameter $\lambda_s \equiv \alpha_2/\alpha_1$. In Appendix A we show the following lemma, which is important for our unit root results and also holds for general utility functions.

Lemma 1 *Under Assumption 1,* $\alpha_1, \alpha_2, ..., \alpha_7 > 0$, $\lambda_s \in \left(1/2, 1\right)$, *and* $\alpha_1 = \alpha_2 + \alpha_3$.

Aggregating (28) and using (23) and (29), we show that equity prices satisfy

$$q_t = \overline{\mathbb{E}}_t\left[\alpha_3 s_{it+2}^h + \Delta b_{it+1}\right] + \overline{\mathbb{E}}_t\left[\beta q_{t+1} + \left(1-\beta\right)d_{t+1}\right] + u_t. \tag{32}$$

The first term on the right-hand side of the second equality is the average forecast of the individual SDFs, which depend on future aggregate dividends, individual shareholdings, and individual labor income. Iterating (32) forward, we find that the equity price is determined by an infinite number of forward-looking higher-order expectations about aggregate dividends and individual shareholdings and labor income.

In summary, we characterize the log-linearized equilibrium by a system of 6 equations (25), (26), (27), (28), (31), and (32) for 6 variables $y_{it}, n_{it}, y_t, s_{it}^h, d_t$, and q_t. We are looking for causal covariance stationary equilibrium processes.

3.4 Full Information Benchmark

Before solving for the equilibrium under dispersed information, we present the equilibrium under full information. In this case all agents have the same information about all shocks. They have rational expectations except when forecasting future stock market conditions because they make small forecast errors due to the near rational shock. Hence equations (27) and (32) become

$$y_t = \frac{1}{\xi} a_t + \theta \mathbb{E}_t [y_t], \tag{33}$$

$$q_t = \mathbb{E}_t [\Delta b_{t+1}] + \mathbb{E}_t [\beta q_{t+1} + (1-\beta) d_{t+1}] + u_t, \tag{34}$$

where $b_t = \alpha_4 d_t + \alpha_5 n_t$ and $\mathbb{E}_t$ denotes the rational expectation operator given all available information.

It follows that

$$c_t^{FI} = y_t^{FI} = \frac{1}{(1-\theta)\xi} a_t, \tag{35}$$

where a variable with a superscript "FI" denotes its full information value. We then use (26) and (31) to derive

$$n_t^{FI} = \frac{1-(1-\theta)\xi}{\alpha(1-\theta)\xi} a_t, \qquad d_t^{FI} = \frac{\alpha - \alpha_7 [1-(1-\theta)\xi]}{\alpha\alpha_6 (1-\theta)\xi} a_t.$$

Applying the method of undetermined coefficients and the Hansen and Sargent (1980) prediction formula to (34) yields

$$q_{\boldsymbol{t}}^{FI} = c_t^{FI} + \frac{u(L)L - \beta u(\beta)}{L-\beta} \epsilon_{ut}.$$

Thus, given a small forecast error u_t, the model under full information cannot simultaneously generate smooth consumption (output) and highly volatile equity prices. For example, $q_t^{FI} = c_t^{FI} + u_t$ when $u_t = \epsilon_{ut}$.

Next we investigate individual trading behavior, which lies on the heart of our model mechanism. A subtle but important observation in the full information case is that the processes of individual consumption and shareholdings contain a unit root due to consumption smoothing (Hall (1978)). Applying the method of undetermined coefficients to (28) under full information and using Lemma 1 yield

$$s_{it+1}^{h,FI} = s_{it}^{h,FI} + \chi_s \epsilon_{it} + \frac{1}{\alpha_2} v_{it}, \quad \chi_s \equiv \frac{\alpha_5(1/\xi - 1)}{\alpha\alpha_2}.$$

This in turn implies that individual consumption possesses contain a random walk component using the log-linearized budget constraint:

$$c_{it}^{FI} = c_{it-1}^{FI} + y_t^{FI} - y_{t-1}^{FI} + \chi_c \epsilon_{it} + \left(\frac{D}{C}\chi_s - \chi_c\right)\epsilon_{it-1} - \frac{Q}{\alpha_2 C} v_{it} + \frac{Q+D}{\alpha_2 C} v_{it-1},$$

where $\chi_c \equiv (WN/C)\,(1+\phi)\big[(1/\xi - 1)/\alpha\big] - \chi_s(Q/C)$, and W, N, Q, D, and C are steady state values given in Appendix A.

This result is similar to that in Graham and Wright (2010), where the LLN condition (21) and the full-information assumption ensure that permanent shifts in idiosyncratic consumption and shareholdings cancel out in the aggregate. In particular,

$$\int_I \mathbb{E}_t \left[s_{it+2}^h\right] di = \mathbb{E}_t \int_I \left[s_{it+2}^h\right] di = 0.$$

Under dispersed information, however, this interchange of integration operators is invalid because agents have different information sets, and the interconnection between shareholding choices and the equity price leads to our key results for the financial market.

4 Business Cycle Volatility

In this section we show that output volatility under dispersed information is lower than that under full information without explicitly solving the model.

To analyze the log-linearized equilibrium system under dispersed information, we need to deal with the problem of forecasting the forecast of others as revealed by equations (27) and (32). To see this point, iterating (27) yields

$$y_t = \frac{1}{\xi}\sum_{k=0}^{\infty} \theta^k\, \overline{\mathbb{E}}_t^{(k)} \left[\int_I a_{it} di\right] + \lim_{k\to\infty} \theta^k \overline{\mathbb{E}}_t^{(k)}\left[y_t\right], \quad \overline{\mathbb{E}}_t^{(k)}\left[\cdot\right] = \underbrace{\int_I \mathbb{E}_{it} \int_I \mathbb{E}_{it} \cdots \int_I}_{k} \mathbb{E}_{it}\left[\cdot\right] \underbrace{di \cdots di}_{k}$$

where $\overline{\mathbb{E}}_t^{(k)}$ denotes the k-order average expectation is the repeated integral. Under dispersed information, aggregate output depends on an infinite number of higher-order expectations. Solving these higher-order expectations in the time domain is challenging.

Therefore we adopt the frequency domain approach discussed in Online Supplementary Appendices S2 and S3.

Conjecture that the solution for output in island i takes the following form

$$y_{it} = M_y^a(L)\,\epsilon_{at} + M_y^i(L)\,\epsilon_{it} + M_y^u(L)\,\epsilon_{ut}, \tag{36}$$

where the corresponding z-transforms $M_y^a(z)$, $M_y^u(z)$, and $M_y^i(z)$ are some analytic functions in $\mathbf{H}^2(\mathbb{D})$.[6] Then aggregate output satisfies

$$y_t = \int_I y_{it}di = M_y^a(L)\,\epsilon_{at} + M_y^u(L)\,\epsilon_{ut}. \tag{37}$$

We first present a lemma characterizing the property of the variance of higher-order expectations, which is central for determining business cycle volatility when information is dispersed.

Lemma 2 *Under Assumption 1, we have*

$$Var\left(\overline{\mathbb{E}}_t[y_t]\right) < Var\left(\mathbb{E}_{it}[y_t]\right) \leq Var(y_t).$$

Lemma 2 shows that the variance of the average expectations about aggregate output is smaller than the variance of individual expectations about aggregate output, when individual agents' effect on the aggregate equilibrium is infinitesimal so that the LLN can be applied. This feature is in sharp contrast with models that assume finitely-many uninformed agents, such as Kasa, Walker, and Whiteman (2014) and Albuguerque and Miao (2014).

Using the preceding lemma, we show in Appendix B that the business cycle volatility is dampened under dispersed information relative to a full-information environment.

Theorem 1 *Under Assumption 1, the variance of output under dispersed information is bounded above by the variance under full information*

$$Var\left(y_t^{FI}\right) > Var(y_t).$$

The previous literature has demonstrated a related result, including Morris and Shin (2002) and Angeletos and La'O (2013). Here we simply provide an easy way of proving this result without having to explicitly deal with an infinite number of correlated higher-order expectations. This theorem is applicable to general information

[6] Here $\mathbf{H}^2(\mathbb{D})$ denotes the Hardy space for the open unit disk $\mathbb{D}$ of the complex space and $\|\cdot\|_{\mathbf{H}^2}$ denotes its norm. See Online Supplementary Appendix S2.

structures, exogenous or endogenous, univariate or multivariate. Adding confidence or noise shocks would also not change the results.

Here the presence of higher-order beliefs and the forecasting the forecasts of others problem dampens business cycle fluctuations. In the real economy, the effect of dispersed information and higher-order expectations works through two channels. The first channel is associated with slow learning of the unobserved states. Slow learning creates inertia in endogenous variables, and more importantly in the higher-order average expectations of model variables, which leads to low volatility. The second channel is associated with the forecasting the forecasts of others. Agents have a speculative motive if other agents overreact to news. This channel is strong for informationally-influential participants in models with finitely many agents (Kasa, Walker, Whiteman (2014) and Albuguerque and Miao (2014)). It is also at work in the heterogeneous prior setup (Harrison and Kreps (1978) and Scheinkman and Xiong (2003)). When each agent is informationally negligible as in our model, the second channel completely vanishes since there is no need to forecast any particular agent's forecast due to the law of large numbers. What matters is the forecast of the average. Thus the first channel dominates and leads to the volatility bounds we deliver above.

We also note the limitation of Theorem 1. The underlying dampening result depends on the presence of a beauty-contest type of production decisions with strategic complementarity. In this sense, the real block of our model is closely related to Angeletos and La'O (2010), in which production decisions are made prior to the realization of aggregate demand, which is the key behind the dampening result.[7]

5 Equity Price Volatility

We now turn to the financial side of the model. The main result of this section is that equity volatility will converge to infinity as the variance of the idiosyncratic TFP shock converges to infinity. In contrast to the previous section, we need to derive an explicit model solution to establish this result. We will also prove the existence and uniqueness of equilibrium by extensively using the frequency domain methods described in Appendices S3 and S2.

[7]On the other hand, Angeletos and La'O (2013), Benhabib et al (2015), and Chahrour and Gaballo (2018) show that macroeconomic fluctuations under dispersed information can be amplified via a form of non-fundamental volatility or learning from price. The theoretical volatility bound can also be overturned in cases in which higher-order uncertainty about micro shocks is correlated, as in Angeletos and La'O (2013) and Huo and Takayama (2018)

5.1 Equilibrium Solution

We rewrite (32) as

$$q_t = \int_I \chi_{it} di + u_t, \tag{38}$$

where we define

$$\chi_{it} \equiv \mathbb{E}_{it}\left[\alpha_3 s^h_{it+2} + \Delta b_{it+1}\right] + \mathbb{E}_{it}\left[\beta q_{t+1} + (1-\beta)\, d_{t+1}\right]. \tag{39}$$

The information set consists of the history of signals $X_{it} = [a_{it}, q_t]^\intercal$. Conjecture that

$$\chi_{it} = \pi_1(L) a_{it} + \pi_2(L) q_t, \tag{40}$$

where the analytic functions $\pi_1(z)$ and $\pi_2(z)$ are endogenously determined in $\mathbf{H}^2(\mathbb{D})$.

It follows from (38) that

$$q_t = \frac{\pi_1(L) a(L)}{1-\pi_2(L)} \epsilon_{at} + \frac{u(L)}{1-\pi_2(L)} \epsilon_{ut}. \tag{41}$$

The lag polynomial $\pi_1(L)$ characterizes how the dispersed information about TFP shocks affects equity prices, while $1/\left[1-\pi_2(L)\right]$ characterizes the effect of endogenous learning from equity prices.

To verify the conjecture in (40), we use (41) and the Wiener-Hopf prediction formula to compute the conditional expectations in (40).[8] To apply this formula, we write the signal representation as

$$X_{it} = H(L)\eta_{it} \equiv \begin{bmatrix} a(L) & 1 & 0 \\ \frac{\pi_1(L)a(L)}{1-\pi_2(L)} & 0 & \frac{u(L)}{1-\pi_2(L)} \end{bmatrix} \begin{bmatrix} \epsilon_{at} \\ \epsilon_{it} \\ \epsilon_{ut} \end{bmatrix}, \tag{42}$$

which is a non-square system containing endogenous functions. To derive transparent analytical solutions, we impose the following assumption:

Assumption 2 *Let $u(z) = \pi_1(z)$ and $a(z) = 1$.*

The assumption of IID TFP shocks is for simplicity and can be easily relaxed. The assumption of $u(z) = \pi_1(z)$ follows from Taub (1989) and Rondina and Walker (2015) and substantially simplifies the computation of the spectral factorization and the Wold representation. We can express the equilibrium conditions as a system of linear functional equations for $\pi_1(z)$ and $\pi_2(z)$, allowing us to establish the equilibrium

[8] In Online Supplementary Appendix S3 we provide the details of this formula.

existence and uniqueness and analyze the key model mechanism transparently in the frequency domain. In the next section we relax Assumption 2 and derive numerical results.

In our analytical solution and numerical procedures, one of the key steps is to find the Wold fundamental representation or the spectral factorization for the model's signal structure. We make an important methodological contribution by providing a two–step triangular spectral factorization method based on Rozanov (1967). This method delivers a closed-form spectral factorization up to the solution of complex polynomial equations. Using Rouché's theorem and the fundamental theorem of algebra, we are able to characterize the location and magnitude of the polynomial roots. These roots in turn determine the equilibrium existence and uniqueness as well as its dynamic properties.[9] In the online supplementary appendix S3 and S4, we supply the mathematical details of this approach, including a working example in which we derive the factorization step-by-step as a guide for interested readers.

Conjecture that the equilibrium individual shareholdings satisfy

$$s^h_{it+1} = M^i_s(L)\,\epsilon_{it} + M^v_s(L)\,v_{it}, \tag{43}$$

where $M^i_s(z), M^v_s(z) \in \mathbf{H}^2(\mathbb{D})$. In equilibrium, individual shareholdings can only respond to idiosyncratic TFP shocks and idiosyncratic forecast errors, because the aggregate number of shares is fixed. The following result delivers the link between equity prices and individual shareholdings.

Lemma 3 *Under Assumptions 1 and 2, we have*

$$M^i_s(z) = \frac{\pi_1(z)}{\alpha_1 - \alpha_2 z}. \tag{44}$$

This lemma shows that the exposure of an investor's shareholdings to the idiosyncratic TFP shock is closely related to the equity price exposure to the aggregate TFP shock due to investor's dispersed information about the two components of the shocks. Thus, if investors make large adjustments of their shareholding positions, the response of equity prices to aggregate TFP shocks will also be large. However, this relation vanishes under full-information as analyzed in Section 3.4, because cross-sectional aggregation neutralizes the effect of individual trading decisions on equity prices.

[9] In a technical appendix available upon request, we use these tools to characterize an exogenous information model with a quartic polynomial system.

Theorem 2 *Under Assumptions 1 and 2, there is a unique equilibrium under dispersed information in which* $\pi_1(z)$ *and* $\pi_2(z)$ *are rational analytical functions if the function* $\frac{\pi_1(z)}{1-\pi_2(z)} \in \mathbf{H}^2(\mathbb{D})$ *has no roots in the open unit disk.*

In Appendix C we provide an explicit solution to the equilibrium. The equilibrium is characterized by rational analytic functions $\pi_1(z)$ and $\pi_2(z)$ in the closed unit disk, which corresponds to ARMA(p,q) representations in the time domain. Despite the presence of the infinite number of higher-order expectations formed by agents, the ARMA(p,q) representation allows us to compute the equity price volatility in closed-form via the integral method and Parseval's theorem. More importantly, the explicit expression also highlights some crucial analytical properties of the equity price fluctuations under dispersed information. We are particularly interested in the properties of equity prices as $\sigma_i \to \infty$.

5.2 Equity Volatility

We decompose the equity price in (41) as $q_t = q_t^f + q_t^n$, where

$$q_t^f \equiv \frac{\pi_1(L)}{1-\pi_2(L)}\epsilon_t^a \text{ and } q_t^n = \frac{u(L)}{1-\pi_2(L)}\epsilon_{ut}$$

represent the components driven by the fundamental TFP shock and the common forecast error, respectively. In Appendix C we prove the following result.

Theorem 3 *Under Assumptions 1 and 2, we have*

$$\lim_{\sigma_i\to\infty} \pi_1(1) = \infty; \qquad \lim_{\sigma_i\to\infty} Var\left(q_t^f\right) = \sigma_a^2 \lim_{\sigma_i\to\infty} \left\|\frac{\pi_1(z)}{1-\pi_2(z)}\right\|_{\mathbf{H}^2}^2 = \infty. \tag{45}$$

Although idiosyncratic TFP shocks have no effect on the equity price, the equity price becomes arbitrarily volatile as the volatility of the idiosyncratic shock approaches infinity for any finite $\sigma_a > 0$ and $\sigma_u > 0$. Therefore, our model has the potential to generate a highly volatile equity price, because the idiosyncratic TFP volatility is much larger than the aggregate TFP volatility in the data.

To understand the economic mechanism generating the high equity price volatility, we rewrite (32) as $q_t = \int_I \mathbb{E}_{it}[\beta q_{t+1} + (1-\beta) d_{t+1}] di + \int_I \mathbb{E}_{it} m_{it+1} di + u_t$, where we can show that

$$\int_I \mathbb{E}_{it}[m_{it+1}] di = \int_I \mathbb{E}_{it}\left[\alpha_3 s_{it+2}^h + \Delta b_{it+1}\right] di. \tag{46}$$

Iterating forward gives

$$
\begin{aligned}
q_t &= \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t ... \overline{\mathbb{E}}_{t+k} \left[m_{it+k+1} \right] \\
&\quad + (1-\beta) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t ... \overline{\mathbb{E}}_{t+k} \left[d_{t+k+1} \right] + \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t ... \overline{\mathbb{E}}_{t+k} \left[u_{t+k+1} \right].
\end{aligned}
$$

Thus the equity price consists of a present-value component under a constant SDF (an infinite sum of higher-order expectations about future aggregate dividends), a component equal to the infinite sum of higher-order expectations about individual SDFs, and a non-fundamental component due to common forecast errors.

Using the intuition developed in Sections 2 and 4, we know that the present value component cannot generate a large volatility, as the higher-order expectations about future aggregate dividends are smoother than aggregate dividends. In other words, higher-order expectations about aggregate variables and the failure of the law of iterated expectations do not lead to excess volatility per se. We thus focus on the second component, which depends on the average forecast of future individual shareholdings and labor income by (46). Unlike in the case of full information studied in Section 3.4, the average forecast of individual shareholdings is not equal to the forecast of the average shareholdings,

$$
\int_I \mathbb{E}_{it} \left[s^h_{it+2} \right] di \neq \mathbb{E}_{it} \int_I \left[s^h_{it+2} \right] di = 0.
$$

Correlated movements in the average expectation of individual shareholdings now affect aggregate equity prices.

Individual equity trading decisions only respond to idiosyncratic TFP shocks, and not aggregate TFP shocks, because assets are in fixed supply. Investors interpret a change in the TFP signal as an idiosyncratic shock to their budget sets. As idiosyncratic TFP volatility σ_i tends to infinity, individual shareholding volatility also tends to infinity because the shareholding process contains a unit root as in the full information case. This unit root is transmitted to the equity price in response to the aggregate TFP shock by Lemma 3. Formally, $M^i_s(1) \to \infty$ if and only if $\pi_1(1) \to \infty$.

To gain intuition about where this unit root comes from, consider the following thought experiment. Suppose that the economy receives a positive innovation to a_t; since σ_i is arbitrarily large, an agent observing the increase in a_{it} will mistakenly attribute it to ϵ_{it}. This changes agent i's permanent income unexpectedly, leading him to raise his shareholdings permanently due to his consumption smoothing motive as in the full information case. At the same time, the agent believes there exist other agents whose idiosyncratic demand for shares has fallen by the exact same amount (leaving aggregate demand unchanged) if he does not observe the price change. Because all agents make this same mistaken inference, aggregate demand rises and market clearing requires the equilibrium price to rise permanently as $\sigma_i \to \infty$.[10]

However, agents also observe the price change and can use that signal to infer that there must have been an aggregate shock. By (40) and (41) the learning effect is reflected by the denominator $1 - \pi_2(L)$. If $1 - \pi_2(1) \to \infty$ as $\sigma_i \to \infty$, the unit root for $\pi_1(z)$ would cancel out. The intuition for this adjustment is that the price here plays two roles – it clears markets and it provides information. For now, suppose the near rational error is not present in the model. Then the equity price information fully reveals the aggregate TFP shock. As q_t rises, agents move along their demand curve as usual. But the demand curve also shifts inward as q_t rises, because agents know that the TFP shock is mean-reverting, so future equity prices will be lower than the current price, reducing their current demand. As a result, the unit root from individual shareholdings does not get transmitted into the equity price.

With near rational errors this process gets short-circuited. In this case, agents will optimally assign weights to a rise in a_t and a rise in u_t; since agents believe that u_t applies only to other agents, each individual will overestimate the future value of stocks (through some combination of overestimating q_{t+1}, d_{t+1}, and future consumption), which prevents their demand curve from falling as q_t rises. The learning effect from equity prices is always weaker than the information conveyed by the TFP signal so that the unit root associated with the TFP shock will survive. Therefore when the aggregate TFP shock hits the economy, investors' expectations about future trading decisions adjust in a simultaneous and persistent manner, leading to high equity price volatility.

Formally, using the Wiener-Hopf prediction formula, we have

$$\begin{aligned}\mathbb{E}_{it}\left[s^h_{it+2}\right] &= \frac{\tau_1}{\alpha_3 L}\left[\frac{(1-\lambda_s)\,\pi_1(L)}{1-\lambda_s L} - (1-\lambda_s)\,\pi_1(0)\right] a_{it} \\ &\quad - \frac{\tau_2}{\alpha_3 L}\left[\frac{(1-\lambda_s)\,\pi_1(L)}{1-\lambda_s L} - (1-\lambda_s)\,\pi_1(0)\right] \frac{1-\pi_2(L)}{\pi_1(L)} q_t,\end{aligned}$$

where τ_1 and τ_2 are the signal-to-noise ratios (see equations (C.2) and (C.3) in Ap-

[10] We prove this result under exogenous information formally in a technical appendix available upon request.

pendix C):

$$\tau_1 = \frac{\sigma_i^2}{\sigma_i^2 + \left(\sigma_a^{-2} + \sigma_u^{-2}\right)^{-1}}, \quad \tau_2 = \tau_1 \frac{\sigma_a^2}{\sigma_a^2 + \sigma_u^2}.$$

Due to the common forecast error $\sigma_u > 0$, we have $\tau_2 < \tau_1$. If $\sigma_i \to \infty$, we have $\tau_1 \to 1$, but $\tau_2 \to \sigma_a^2 / \left(\sigma_a^2 + \sigma_u^2\right) \in (0, 1)$. Thus the expression on the second line of the equation above associated with learning from prices does not fully offset the expression on the first line.

Note that the previous result is not simply a result of adding more unknown shocks than signals; while both lead to forecasting confusion, only the near-rational shock leads to the right kind of confusion. A common practice in the noisy rational expectation literature is to introduce an asset supply/noise trader shock. In that case, the rising price will be attributed partly to a rise in aggregate TFP and partly to a fall in asset supply. But both shocks imply that future prices will fall relative to current prices, causing agents to reduce (shift inward) their asset demand, which eliminates the confusion effect and also the unit root. The key intuition is that agents understand how changes in the aggregate asset supply shock affect their individual shareholding decisions. In this sense, the equity price contains enough information for agents to eliminate correlated movements in the average expectation of individual shareholdings. Note that the equilibrium is still not fully revealing, as agents will be unsure about the source of the price change; however, they are certain that the shock is an aggregate one and fully understand how this aggregate change affects their trading decisions.

To summarize, our model captures three distinct forces that act on real and financial volatility. First, since higher-order expectations about the aggregate shock are less volatile than the shock itself as in Morris and Shin (2002) and Bergemann and Morris (2013), real business cycles are dampened. The second mechanism is the confusion between the aggregate and idiosyncratic shock, which dates back to Lucas (1972). The third mechanism is the higher-order expectation of idiosyncratic shocks, which produces extra volatility that would not be possible under homogeneous information. Indeed, our main contribution is to show how the interactions of these last two forces, propagated in the feedback loop between equity price and individual trading, lead to high volatility. A related result can be found in Bergemann, Heumann, and Morris (2015), who show that beauty-contest models can lead to unbounded aggregate volatility. The fundamental difference between our theoretical mechanism and their approach is that they choose the signal structure (i.e., weights in the signal) to maximize volatility, whereas we allow agents to optimally filter the shocks with given signal structure.[11]

6 Discussions

6.1 Numerical Results

One side effect of the assumption of $u(z) = \pi_1(z)$ is that the volatility of the non-fundamental component of the equity price also approaches infinity as $\sigma_i \to \infty$. To isolate this effect, we relax Assumption 2 by assuming that u_t and a_t follow independent AR(1) processes.

Assumption 3 *$u(z) = 1/(1-\rho_u z)$ and $a(z) = 1/(1-\rho_a z)$, where $\rho_a, \rho_u \in [0,1)$.*

Now the equilibrium system cannot be reduced to a system of linear functional equations for $\pi_1(z)$ and $\pi_2(z)$ and hence $\pi_1(z)$ and $\pi_2(z)$ cannot be represented by analytic rational functions. It is well known that any non-rational analytic functions can be approximated by rational functions with arbitrary accuracy (Rudin (1987)). Using this fact, we compute the model numerically by using rational functions to approximate $\pi_1(z)$ and $\pi_2(z)$. In the numerical computation, our spectral factorization method also displays its advantages. It allows us to derive an almost-analytical spectral factor matrix and solve for the non-linear equilibrium fixed point problem in a clear, algebraic form. In this way, we minimize the "black box" in numerical computation so that we are able to see whether the algorithm is correct and whether the model mechanism works.[12]. In Supplementary Appendix S1 we provide the equilibrium system and the numerical algorithm we use to solve the model.

To derive quantitative implications, we calibrate the model parameters assuming one model period corresponds to a quarter. We set the subjective discount factor $\beta = 0.99$, the elasticity of output with respect to labor $\alpha = 0.67$, the persistence of the aggregate TFP shock $\rho_a = 0.8$, and the volatility of the aggregate TFP shock $\sigma_a = 0.7\%$. We also set the inter-island elasticity of substitution $\varsigma = 9$ to generate a steady-state markup of 12.5%, and set $\phi = 2$ to generate a Frisch elasticity of labor

[11] More specifically, they allow the signal weights to be chosen as functions of idiosyncratic volatility and the maximal volatility is attained when the weight of idiosyncratic shock goes to zero, whereas the idiosyncratic volatility in our model only shows up in signal-to-noise ratio as we assume exogenously fixed equal weights in the TFP signal.

[12] Compared with the case in Section 5, we are also able to illustrate why the equilibrium cannot be represented using finite-state representation in terms of rational ARMA(p,q) processes. We leave these details to Online Supplementary Appendix S1.

supply equal to 0.5, consistent with Galí (2015), Angeletos and La'O (2013), and King and Rebelo (2000). As baseline values, we set the idiosyncratic volatility $\sigma_i = 5\%$ and the persistence and volatility of the common forecast error $\rho_u = 0.05$, $\sigma_u = 0.04\%$. It is a well-known empirical fact that idiosyncratic productivity shocks are far more volatile than the aggregate shocks, and our choice of σ_i falls well within the range reported by early literature.[13] The implied ratio of the unconditional volatility of the common forecast error and the unconditional total volatility of the aggregate and idiosyncratic TFP shocks is 0.65%. This small forecast error is consistent with the estimate in Hassan and Mertens (2017). In the Online Supplementary Appendix we show that aggregate output and equity volatilities are independent of the idiosyncratic forecast error volatility. We thus do not need to assign a value for σ_v for our numerical solutions.

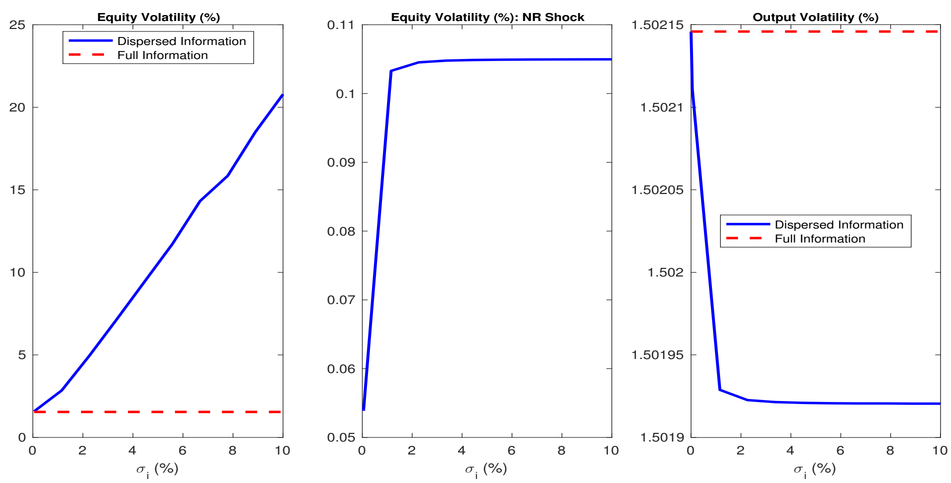

Figure 2: The effect of idiosyncratic TFP volatility σ_i on equity and output volatility.

Our baseline calibration implies a quarterly output volatility of 1.5% and a quarterly equity volatility of 10.5%; the empirical counterparts are 1.61% and 12.04%, respectively.[14] Figure 2 presents the effect of idiosyncratic TFP volatility on equity price volatility. If $\sigma_i = 0$, the model with dispersed information reduces to the one

[13]See Comin and Philippon (2005) and Franco and Philippon (2007).

[14]Aggregate output data (real GDP) is taken from the FRED database, and we convert the aggregate data into per capita terms using the civilian non-institutional population from the Bureau of Labor Statistics (BLS) website. Total hours are defined as total non-farm business hours divided by the population. The monthly equity price series for the S&P 500 Composite Price Index is logged, HP-filtered and averaged over quarterly frequencies (obtained from CRSP). All data cover the period 1968Q4 to 2013Q4.

with full information. As σ_i increases from 0 to 10%, equity volatility rises quickly, but output volatility declines slowly. The component (q_t^n) of equity volatility contributed by the common forecast error increases with σ_i and accounts for a very small fraction of total equity volatility (less than 1%). Thus a very small near-rational error can produce high equity price volatility for reasonable values of σ_i.

Since our model results are driven by information dispersion and confusion, we also evaluate the degree of information frictions in the model. Following Coibion and Gorodnichenko (2012), we measure information frictions using the ratio of the elasticity of the average forecasts of aggregate output with respect to the aggregate TFP shock to the elasticity of actual output with respect to the aggregate TFP shock, $\delta = \frac{\partial \overline{\mathbb{E}}_t[y_t]}{\partial \epsilon_{at}} / \frac{\partial y_t}{\partial \epsilon_{at}}$.[15] This ratio serves as a proxy for the amount of confusion agents face. Figure 3 plots δ against idiosyncratic volatility σ_i (left panel) and its relation to the equity price volatility (right panel). An increase in σ_i raises information frictions so that δ falls and equity price volatility rises as discussed before.

An important observation from this exercise is that the required amount of information confusion needed to generate a high equity price volatility is quite small (δ being around 0.885). Confusion is limited in equilibrium because the equity price partially reveals information and the near-rational shock is small. Compared with the data, we only need a small amount of confusion to obtain our main results. In a comparative statics exercise that is available upon request, we find that the patterns discovered in Figure 3 remain the same as we change the degree of strategic complementarity θ (since θ is not a model primitive, but rather a reduced-form parameter, we can achieve this variation in multiple ways).

6.2 Macro-Financial Disconnection

So far we have demonstrated the sharp differences between the financial market and the real economy in terms of the volatility changes induced by dispersed information. In this section we address another important aspect of the empirical data: the weak correlation between the two. In our baseline model, volatilities in both sides of the economy are driven primarily by the same aggregate TFP shock, a feature that ensures the model's analytical tractability but also leads to the counterfactual prediction that the equity price and output are perfectly correlated. This issue, however, can be resolved by considering a simple modification to the basic model with limited stock market participation.

[15] We greatly appreciate one of the referees for suggesting this exercise.

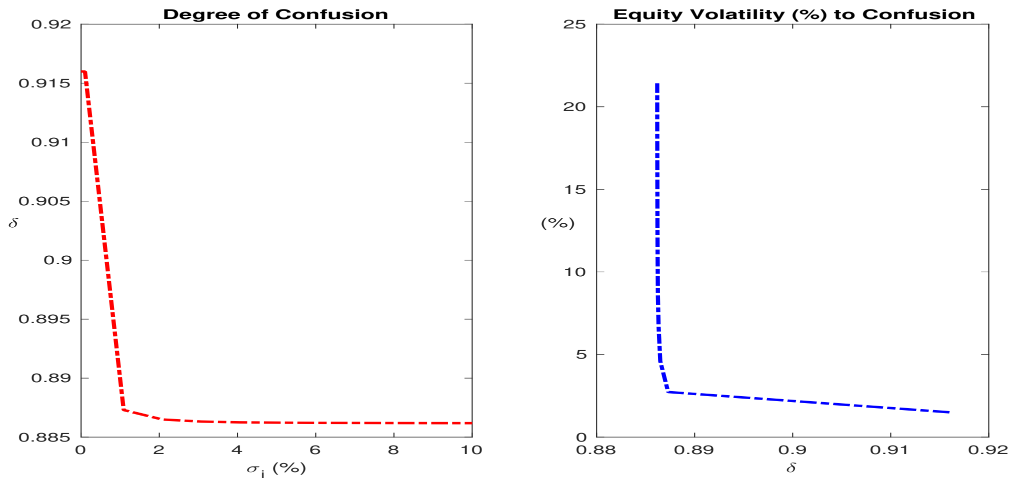

Figure 3: Degree of Information Confusion and Equity Volatility

Suppose that the continuum of islands $I = [0, 1]$ is partitioned into two groups $I = I_p \cup I_n$. Let $I_p = [0, \kappa)$ denote the interval of islands that participate in the stock market trading, while $I_n = [\kappa, 1]$ denotes the interval of non–participants with $\kappa \in (0, 1)$ being the stock-market participation rate. The household problem on participating islands I_p is identical to the baseline model, while on non–participating islands only local labor income is used to finance consumption. Instead of receiving an aggregate TFP shock, participating and non–participating islands are now subject to different group–specific (log-linearized) TFP shocks which are common within groups: $a_t^x = \rho_a a_{t-1}^x + \epsilon_{at}^x$; $x = p, n$, where ϵ_{at}^p and ϵ_{at}^n are IID Gaussian innovations with mean zero and variance σ_a^2. Assume that ϵ_{at}^p and ϵ_{at}^n are independent for all t.

We modify the model's information structure as follows. The information set for investors and firms on island $i \in I_p$ is characterized by $\mathcal{I}_{it}^p = \left\{X_{i,t-k}^p\right\}_{k=0}^{\infty}$ with the signal vector $X_{it}^p = [a_{it}^p, q_t]^\intercal$, which is identical to the baseline model except that the aggregate TFP shock is replaced by the group-specific TFP shock. We maintain the assumption of information separation between investor and shopper within each household. On the other hand, workers and firms on island $j \in I_n$ are able to observe their group-specific TFP shock a_t^n and all idiosyncratic TFP shocks on non-participating islands perfectly. Islands in I_n cannot observe stock prices or dividends, and only shoppers on those islands observe goods prices $\{p_t(i)\}_{i \in I}$. Information is segregated between the two groups.

In the Online Supplementary Appendix S1.3, we show that the equity price in this case is only affected by the participants' TFP shocks ϵ_{at}^p and the near rational errors. In addition, the dynamic interactions between shareholding choices and the aggregate equity price will generate the same unit root property in the equity price as in the basic model. On the other hand, aggregate output is given

$$y_t = \kappa\left[M_y^p(L)\epsilon_{at}^p + M_y^u(L)\epsilon_{ut}\right] + (1-\kappa)\frac{1}{\xi\left[1-\theta(1-\kappa)\right](1-\rho_a L)}\epsilon_{at}^n, \qquad (47)$$

where $M_y^p(L)$ and $M_y^u(L)$ are the decision rules of the participating islands, which are determined by an equilibrium condition that is almost equivalent to the one in the basic model except for the appearance of κ (and appropriate changes in the steady-state coefficients). We leave all the mathematical details to Supplementary Appendix S1.3.

We provide intuition using equation (47). Suppose that κ is small and σ_i is large. Then the modified model is able to generate large equity volatility due to the information confusion mechanism, and the equity price is driven almost entirely by participants' common TFP shocks ϵ_{at}^p. Meanwhile, aggregate output remains smooth and is driven predominantly by non-participants' common TFP shocks ϵ_{at}^n, since κ is small. Given that $cov\left(\epsilon_{at}^p, \epsilon_{at}^n\right) = 0$, the modified model is able to generate a large equity price volatility, while the equity price and output are weakly correlated.

7 Conclusion

We have developed a model of a production economy with dispersed information that features smooth aggregate consumption (output) dynamics and highly volatile equity prices. The key elements of our model are not assumptions on nonstandard preferences, bubbles, or sentiments, but the introduction of dispersed information, near rational expectations, incomplete markets, and the endogeneity of SDFs that are time-varying and heterogeneous across population. The key for our model result is due to the different impact of the higher-order beliefs about the average forecasts of aggregate demand and the individual SDFs, together with the dynamic interaction between shareholdings and equity prices. From a technical point of view, we have proposed a two-step spectral factorization method in the frequency domain, which can be applied to many other contexts that involves solving signal extraction problems with non-square systems.

References

Albuquerque, Rui, and Jianjun Miao, 2014, Advance Information and Asset Prices, *Journal of Economic Theory* 149, 236-275.

Angeletos, George-Marios, and Jennifer La'O, 2010, Noisy Business Cycles, in *NBER Macroeconomics Annual 2009*, Volume 24, pp. 319–378. University of Chicago Press.

Angeletos, George-Marios, and Jennifer La'O, 2013, Sentiments, *Econometrica* 81, 739–779.

Benhabib, Jess, Pengfei Wang, and Yi Wen, 2015, Sentiments and Aggregate Demand Fluctuations, *Econometrica* 83, 549-585.

Bergemann, Dirk, Tibor Heumann, and Stephen Morris, 2015, Information and Volatility, *Journal of Economic Theory* 158, 427-465.

Bergemann, Dirk, and Stephen Morris, 2013, Robust Predictions in Games with Incomplete Information, *Econometrica* 81, 1251-1308.

Campbell, John Y., 1999, Asset Prices, Consumption and the Business Cycle, in John B. Taylor, and Michael Woodford, eds.: *Handbook of Macroeconomics*, vol. 1 (Elsevier Science, North-Holland, Amsterdam).

Chahrour, Ryan, and Gaetano Gaballo, 2018, Learning from Prices: Amplification and Sentiments, Boston College Working Paper 873.

Cochrane, John, H., 2011, Discount Rates, *Journal of Finance* 66, 1047-1108.

Coibion, Olivier, and Yuriy Gorodnichenko, 2012, What Can Survey Forecasts Tell Us About Information Rigidities?, *Journal of Political Economy* 120, 116-159.

Comin, Diego, and Thomas Philippon, 2005 The Rise in Firm-Level Volatility: Causes and Consequences. *NBER macroeconomics annual* 20, 167-201.

Franco, Francesco, and Thomas Philippon, 2007, Firms and Aggregate Dynamics, *Review of Economics and Statistics* 89, 9225-9225.

Galí, Jordi, 2015, *Monetary Policy, Inflation, and the Business Cycle, An Introduction to the New Keynesian Framework and Its Applications*, Second Edition, Princeton Press

Graham, Liam, and Stephen Wright, 2010, Information, Heterogeneity and Market Incompleteness, *Journal of Monetary Economics* 57, 164-174.

Greenwood, Jeremy, Zvi Hercowitz, and Gregory W. Huffman, 1988, Investment, Capacity Utilization, and the Real Business Cycle, *American Economic Review* 78, 402-17.

Hall, Robert E., 1978, Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence, *Journal of Political Economy* 86, 971–987.

Hansen, Lars P., and Thomas J. Sargent, 1980, Formulating and Estimating Dynamic Linear Rational Expectation Models, *Journal of Economic Dynamics and Control* 2, 7–46.

Harrison, Michael J., and David M. Kreps, 1978, Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations, *Quarterly Journal of Economics* 92, 323-336.

Hassan, Tarek, and Thomas M. Mertens, 2017, The Social Cost of Near-Rational Investment, *American Economic Review* 107, 1059-1103.

Huo, Zhen, and Naoki Takayama, 2018, Rational Expectations Models with Higher Order Beliefs, working paper, University of Minnesota.

Kasa, Kenneth, Todd B. Walker, and Charles H. Whiteman, 2014, Heterogeneous Beliefs and Tests of Present Value Models, *Review of Economic Studies* 81, 1137-1163.

King, Robert G., and Sergio T. Rebelo, 2000, Resuscitating Real Business Cycles, *NBER Working Papers*

Lindquist, Anders, and Giorgio Picci, 2015, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*, Springer-Verlag Berlin Heidelberg

Lucas, Robert, 1972, Expectations and the Neutrality of Money, *Journal of Economic Theory* 4, 103-124.

Makarov, Igor, and Oleg Rytchkov, 2012, Forecasting the Forecasts of Others: Implications for Asset Pricing, *Journal of Economic Theory* 147, 941–966.

Merton, Robert. C., 1969, Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case, *Review of Economics and Statistics* 51, 247-257.

Morris, Stephen, and Hyun S. Shin, 2002, Social Value of Public Information, *American Economic Review* 92, 1521-1534.

Rondina, Giacomo, and Todd B. Walker, 2015, Dispersed Information and Confounding Dynamics, working paper, Indiana University.

Rozanov, Yu A., 1967, *Stationary Random Processes*, San Francisco: Holden-Day.

Rudin, Walter, 1987, *Real and Complex Analysis*, McGraw-Hill Companies, Inc. 3 edn

Scheinkman, Jose A., and Wei Xiong, 2003, Overconfidence and Speculative Bubbles, *Journal of Political Economy* 111, 1183-1220.

Shiller, Robert, 1981, Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends? *American Economic Review* 71, 421-436.

Taub, Bart, 1989, Aggregate Fluctuations as An Information Transmission Mechanism, *Journal of Economic Dynamics and Control* 13, 113–150.

Townsend, Robert M., 1983, Forecasting the Forecasts of Others, *Journal of Political Economy* 91, 546–88.

Appendix

A Proofs of Results in Section 3

We consider a general utility function

$$\mathcal{E}_i\left[\sum_{t=0}^{\infty}\beta^t U(C_{it}, N_{it})\right],$$

where U is twice continuously differentiable and satisfy the usual concavity, monotonicity, and Inada conditions for consumption C_{it} and labor N_{it}. Then the optimality conditions from utility maximization give

$$\begin{aligned}
W_{it} &= -\frac{U_n(C_{it}, N_{it})}{U_c(C_{it}, N_{it})},\\
Q_t &= \mathcal{E}_{it}\left[M_{it+1}(Q_{t+1} + D_{t+1}\right],\ M_{it+1} = \frac{\beta U_c(C_{it+1}, N_{it+1})}{U_c(C_{it}, N_{it})}.
\end{aligned}$$

The (symmetric) deterministic steady state is characterized by the following nonlinear system

$$\begin{aligned}
Y_i &= C_i = C = Y = N^{\alpha},\\
W_i &= W = \left(1 - \frac{1}{\varsigma}\right)\alpha N^{\alpha-1},\\
D &= \left(1 - \left(1 - \frac{1}{\varsigma}\right)\alpha\right)N^{\alpha},\\
Q &= \frac{\beta}{1-\beta}D,\ S_i^h = 1,
\end{aligned}$$

and

$$-\frac{U_n(N^{\alpha}, N)}{U_c(N^{\alpha}, N)} = \alpha(1 - \frac{1}{\varsigma})N^{\alpha-1}. \tag{A.1}$$

Suppose that equation (A.1) has a unique solution $N > 0$. We then obtain a unique deterministic steady state.

Now we consider the log-linear approximation around the deterministic steady state. We use a lower case variable to denote its log deviation from the deterministic steady state. We derive

$$u_c(c_{it}, n_{it}) = -u_1 c_{it} + u_2 n_{it}; \quad u_n(c_{it}, n_{it}) = u_3 c_{it} + u_4 n_{it},$$

where u_1, u_2, u_3, and u_4 are functions of steady-state values as well as the preference parameters

$$\begin{aligned} u_1 &= \frac{-CU_{cc}}{U_c} > 0,\ u_2 = \frac{U_{cn}N}{U_c}, \\ u_3 &= \frac{CU_{nc}}{U_n},\ \ u_4 = \frac{U_{nn}N}{U_n}. \end{aligned}$$

The Euler equation can be log-linearized as

$$q_t = \mathbb{E}_{it}\left[\beta q_{t+1} + (1-\beta)\, d_{t+1}\right] + \mathbb{E}_{it}\left[m_{it+1}\right] + u_t + v_{it}. \tag{A.2}$$

where the stochastic discount factor has the general form

$$m_{it+1} = u_1(c_{it} - c_{it+1}) - u_2(n_{it} - n_{it+1}).$$

The wage rate satisfies

$$w_{it} = (u_3 + u_1)c_{it} + (u_4 - u_2)n_{it}. \tag{A.3}$$

Substituting this equation into the log-linearized budget constraint yields

$$\begin{aligned} Cc_{it} + Qs^h_{it+1} &= (Q+D)\, s^h_{it} + Dd_t + WN\,(n_{it} + w_{it}) \\ &= (Q+D)\, s^h_{it} + Dd_t + WN(1 + u_4 - u_2)n_{it} + WN(u_3 + u_1)c_{it}, \end{aligned} \tag{A.4}$$

which in turn implies

$$\begin{aligned} c_{it} &= -\frac{Q}{C - WN(u_3+u_1)} s^h_{it+1} + \left(\frac{Q+D}{C - WN(u_3+u_1)}\right) s^h_{it} \\ &\quad + \frac{D}{C - WN(u_3+u_1)} d_t + \frac{WN(1+u_4-u_2)}{C - WN(u_3+u_1)} n_{it}. \end{aligned}$$

Substituting this expression for c_{it} into the log-linearized SDF and Euler equation yields

$$\begin{aligned} \frac{u_1(2Q+D)}{C - WN(u_3+u_1)} s^h_{it+1} &= \frac{u_1(Q+D)}{C - WN(u_3+u_1)} s^h_{it} + \mathbb{E}_{it}\left[\frac{u_1 Q}{C - WN(u_3+u_1)} s^h_{it+2} + \Delta b_{it+1}\right] \\ &\quad + \mathbb{E}_{it}\left[\beta q_{t+1} + (1-\beta)\, d_{t+1}\right] - q_t. \end{aligned}$$

where $\Delta b_{it+1} \equiv b_{it} - b_{it+1}$ and

$$b_{it} \equiv \frac{u_1 D}{C - WN(u_3+u_1)} d_t + \left(u_1 \frac{WN(1+u_4-u_2)}{C - WN(u_3+u_1)} - u_2\right) n_{it}.$$

Define

$$\alpha_1 = \frac{u_1(2Q+D)}{C-WN(u_3+u_1)},\ \alpha_2 = \frac{u_1(Q+D)}{C-WN(u_3+u_1)},\ \alpha_3 = \frac{u_1 Q}{C-WN(u_3+u_1)},$$
$$\alpha_4 = \frac{u_1 D}{C-WN(u_3+u_1)},\ \alpha_5 = \left(u_1 \frac{WN(1+u_4-u_2)}{C-WN(u_3+u_1)} - u_2\right).$$

We then obtain equation (28) and can verify that

$$\alpha_1 = \alpha_2 + \alpha_3; \quad \lambda_s \equiv \frac{\alpha_2}{\alpha_1} = \frac{Q+D}{2Q+D} \in (1/2, 1).$$

Thus we have proven Lemma 1.

Log-linearizing (22), (18) and using (A.3) to eliminate w_{it} and n_{it}, we derive

$$\left[\frac{1}{\alpha}(u_4-u_2+1)-(1-\frac{1}{\varsigma})\right] y_{it} + (u_3+u_1)c_{it} = \left[\frac{1}{\alpha}(u_4-u_2+1)\right] a_{it} + \frac{1}{\varsigma}\mathbb{E}_{it}[y_t].$$

Aggregating this equation yields (25), where

$$\xi = \frac{\frac{1}{\alpha}(u_4-u_2+1)-(1-\frac{1}{\varsigma})+(u_3+u_1)}{\frac{1}{\alpha}(u_4-u_2+1)},$$

and

$$\theta = \frac{1}{\varsigma\left(\frac{1}{\alpha}(u_4-u_2+1)-(1-\frac{1}{\varsigma})+(u_3+u_1)\right)}.$$

To ensure a stationary solution, we need to impose assumptions on technology and utility such that $\theta \in (0,1)$.

For the utility function of Greenwood, Huffman, and Hercowitz (1988) used in our paper, we can simplify the computation significantly. In particular, we can derive the deterministic steady state in an explicit form:

$$N_i = N = \left(\alpha\left(1-\tfrac{1}{\varsigma}\right)\right)^{\frac{1}{\phi-\alpha+1}}; \quad Y_i = C_i = C = Y = \left(\alpha\left(1-\tfrac{1}{\varsigma}\right)\right)^{\frac{\alpha}{\phi-\alpha+1}},$$
$$D = \left(1-\left(1-\tfrac{1}{\varsigma}\right)\alpha\right)\left(\alpha\left(1-\tfrac{1}{\varsigma}\right)\right)^{\frac{\alpha}{\phi-\alpha+1}}; \quad W_i = W = \left(1-\tfrac{1}{\varsigma}\right)\alpha N^{\alpha-1}.$$

and $Q = \frac{\beta}{1-\beta}D$, $S_i^h = 1$. Given Assumption 1, all equilibrium variables are positive and $C - \frac{N^{1+\phi}}{1+\phi} > 0$. Log-linearizing equation (9) yields $w_{it} = \phi n_{it}$. We can also compute that

$$b_{it} = \frac{D}{C-\frac{N^{1+\phi}}{1+\phi}} d_t + \frac{\left[WN(1+\phi)-N^{\phi+1}\right]}{C-\frac{N^{1+\phi}}{1+\phi}} n_{it},$$

and

$$\begin{aligned}
\alpha_1 &= \frac{2Q+D}{C-\frac{N^{1+\phi}}{1+\phi}} > 0,\ \alpha_2 = \frac{Q+D}{C-\frac{N^{1+\phi}}{1+\phi}} > 0,\\
\alpha_3 &= \frac{Q}{C-\frac{N^{1+\phi}}{1+\phi}} > 0,\ \alpha_4 = \frac{D}{C-\frac{N^{1+\phi}}{1+\phi}} > 0,\\
\alpha_5 &= \frac{(1+\phi)WN - N^{\phi+1}}{C-\frac{N^{1+\phi}}{1+\phi}} = \frac{\phi N^{\phi+1}}{C-\frac{N^{1+\phi}}{1+\phi}} > 0.
\end{aligned}$$

Log-linearizing (18) yields

$$y_{it} = a_{it} + \alpha n_{it} \Longrightarrow n_{it} = \frac{1}{\alpha}\left(y_{it} - a_{it}\right).$$

Aggregating leads to $n_t = \frac{1}{\alpha}\left(y_t - a_t\right)$. Log-linearizing (24) yields $y_t = \alpha_6 d_t + \alpha_7 n_t$, where

$$\alpha_6 = \frac{D}{Y} > 0,\ \alpha_7 = \frac{(1+\phi)WN}{Y} > 0.$$

Using the above definitions of $\alpha_1, \alpha_2 ...,$ and α_7, we can easily establish Lemma 1.

B Proofs of Results in Section 4

Proof of Lemma 2: By the Wiener-Hopf prediction formula,

$$\mathbb{E}_{it}\left[y_t\right] = a_y^a(L)\, a_{it} + a_y^q(L) q_t = a_y^a(L)\left(a_t + \epsilon_{it}\right) + a_y^q(L) q_t,$$

where $a_y^a(z)$ and $a_y^q(z)$ can be computed using (S3.3). By the LLN (21),

$$\overline{\mathbb{E}}_t\left[y_t\right] = a_y^a(L)\left(a_t + \int_I \epsilon_{it} di\right) + a_y^q(L) q_t = a_y^a(L)\, a(L)\, \epsilon_{at} + a_y^q(L) q_t.$$

In the stationary equilibrium the equity price can be represented as

$$q_t = M_q^a(L)\epsilon_{at} + M_q^u(L)\epsilon_{ut},$$

where $M_q^a(z)$ and $M_q^u(z)$ are some analytic functions in $H^2(\mathbb{D})$. By the Parseval theorem,

$$\begin{aligned}
Var\left(\overline{\mathbb{E}}_t\left[y_t\right]\right) &= \left\|a_y^a(z)\, a(z) + a_y^q(z)\, M_q^a(z)\right\|_{\mathbf{H}^2}^2 \sigma_a^2 + \left\|a_y^q(z)\, M_q^u(z)\right\|_{\mathbf{H}^2}^2 \sigma_u^2\\
&< \left\|a_y^a(z)\, a(z) + a_y^q(z)\, M_q^a(z)\right\|_{\mathbf{H}^2}^2 \sigma_a^2 + \left\|a_y^q(z)\, M_q^u(z)\right\|_{\mathbf{H}^2}^2 \sigma_u^2 + \left\|a_y^a(z)\right\|_{\mathbf{H}^2}^2 \sigma_i^2\\
&= Var\left(\mathbb{E}_{it}\left[y_t\right]\right).
\end{aligned}$$

We can write $\mathbb{E}_{it}\left[y_t\right] + e_t = y_t$, where e_t is uncorrelated with $\mathbb{E}_{it}\left[y_t\right]$. Thus

$$Var\left(y_t\right) \geq Var\left(\mathbb{E}_{it}\left[y_t\right]\right).$$

Combining the two inequalities above gives us the desired result. Q.E.D.

Proof of Theorem 1: By equation (27),

$$Var(y_t) = Var\left(\frac{a_t}{\xi} + \theta\overline{\mathbb{E}}_t[y_t]\right).$$

Using the triangular inequality and Lemma 2, we have

$$\sqrt{Var(y_t)} \leq \sqrt{Var(a_t/\xi)} + \theta\sqrt{Var\left(\overline{\mathbb{E}}_t[y_t]\right)} < \frac{\|a(z)\|_{\mathbf{H}^2}\sigma_a}{\xi} + \theta\sqrt{Var(y_t)}.$$

Thus

$$\sqrt{Var(y_t)} < \frac{\|a(z)\|_{\mathbf{H}^2}\sigma_a}{(1-\theta)\xi}.$$

Using (35), we obtain the desired result. Q.E.D.

C Proofs of Results in Section 5

Proof of Lemma 3: By equation (28) and (39), we obtain

$$\alpha_1 s_{it+1}^h = \alpha_2 s_{it}^h - q_t + \chi_{it} + u_t + v_{it}. \tag{C.1}$$

Plugging equations (40), (41), and (43) into the equation above, we obtain

$$\begin{aligned}
& \alpha_1 M_s^i(L)\epsilon_{it} + \alpha_1 M_s^v(L)v_{it} \\
= \; & \alpha_2 L M_s^i(L)\epsilon_{it} + \alpha_2 L M_s^v(L)v_{it} + \pi_1(L)(\epsilon_{at} + \epsilon_{it}) \\
& + [\pi_2(L) - 1]\left(\frac{\pi_1(L)}{1-\pi_2(L)}\epsilon_{at} + \frac{u(L)}{1-\pi_2(L)}\epsilon_{ut}\right) \\
& + u(L)\epsilon_{ut} + v_{it}.
\end{aligned}$$

Matching coefficients on the two sides of the equation yields

$$\alpha_1 M_s^i(z) = \alpha_2 z M_s^i(z) + \pi_1(z); \quad \alpha_1 M_s^v(z) = \alpha_2 z M_s^v(z) + 1.$$

We then establish this lemma and obtain $M_s^v(z) = \frac{1}{\alpha_1 - \alpha_2 z}$. Q.E.D.

Proof of Theorem 2: Consider the equilibrium conjecture in (41). Given the assumption that $u(z) = \pi_1(z)$, it follows that

$$q_t = \frac{\pi_1(L)}{1-\pi_2(L)}\epsilon_{at} + \frac{\pi_1(L)}{1-\pi_2(L)}\epsilon_{ut}.$$

For q_t and u_t to be causal stationary processes, we need $\frac{\pi_1(z)}{1-\pi_2(z)}$ and $\pi_1(z)$ to be in the Hardy space $\mathbf{H}^2(\mathbb{D})$. We will verify this condition later.

Step 1. We start by deriving the Wold representation for the signal process $\{X_{it}\}$ given in (42). We compute the covariance generating function

$$S_x(z) = H(z)\Sigma_{\boldsymbol{\eta}} H(z^{-1})^\intercal = \begin{bmatrix} \sigma_a^2 + \sigma_i^2 & \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})}\sigma_a^2 \\ \frac{\pi_1(z)}{1-\pi_2(z)}\sigma_a^2 & \frac{\pi_1(z)\pi_1(z^{-1})}{(1-\pi_2(z))(1-\pi_2(z^{-1}))}(\sigma_a^2+\sigma_u^2) \end{bmatrix},$$

where $\Sigma_\eta = \begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_i^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{bmatrix}$ is the covariance matrix for the innovation vector $\eta_{it} = [\epsilon_{at}, \epsilon_{it}, \epsilon_{ut}]'$. We wish to derive the spectral factorization $S_x(z) = \Gamma(z)\Gamma(z^{-1})^\intercal$. Applying the triangular factorization method described in Appendix S3,[16] we obtain

$$\Gamma(z) = \begin{bmatrix} \sigma_e & \frac{\sigma_a^2}{\sigma_p} \\ 0 & \frac{\pi_1(z)}{1-\pi_2(z)}\sigma_p \end{bmatrix}, \quad \Gamma^{-1}(z) = \begin{bmatrix} \frac{1}{\sigma_e} & -\frac{\sigma_a^2}{\sigma_p^2\sigma_e}\frac{1-\pi_2(z)}{\pi_1(z)} \\ 0 & \frac{1-\pi_2(z)}{\pi_1(z)\sigma_p} \end{bmatrix},$$

where we define

$$\sigma_e^2 \equiv \sigma_i^2 + \frac{\sigma_a^2\sigma_u^2}{\sigma_a^2+\sigma_u^2}, \quad \sigma_p^2 \equiv \sigma_a^2 + \sigma_u^2.$$

Note that

$$\det \Gamma(z) = \sigma_p\sigma_e\frac{\pi_1(z)}{1-\pi_2(z)}.$$

By Theorem 4.6.11 in Lindquist and Picci (2015), $\Gamma(z)$ is a Wold spectral factor if and only if $\frac{\pi_1(z)}{1-\pi_2(z)}$ has no roots in the open unit disk. We shall make this assumption and then obtain the Wold representation $X_{it} = \Gamma(L)\, e_{it}$, where e_{it} is a two-dimensional Wold fundamental innovation vector with zero mean and identity covariance matrix.

Step 2. We next solve for the equilibrium quantities. We conjecture that $y_{it} = M_y(L)\eta_{it}$, where $M_y(z) = \left[M_y^a(z), M_y^i(z), M_y^u(z)\right]$ and $M_y^a(z)$, $M_y^i(z)$, and $M_y^u(z)$ are all in $\mathbf{H}^2(\mathbb{D})$. Aggregation leads to aggregate output $y_t = M_y(z)I_y\eta_{it}$, where $I_y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Using the Wiener-Hopf prediction formula, we derive that

$$\mathbb{E}_{it}[y_t] = \left[\psi_y(L)\right]_+ \Gamma^{-1}(L)X_{it},$$

where $[\cdot]_+$ is the annihilation operator and the z-transform of the operator ψ_y is

$$\psi_y(z) = S_{yx}(z)\left(\Gamma^{-1}(z^{-1})\right)^\intercal.$$

[16] Following Rondina and Walker (2015), we transform the lower-triangular matrix to the upper triangular form by right multiplication of an unitary matrix, which ease the algebra.

The cross-spectrum is given by

$$
\begin{aligned}
S_{yx}(z) &= M_y(z) I_y \Sigma_\eta H^\intercal(z^{-1}) \\
&= \left[M_y^a(z), 0, M_y^u(z)\right] \begin{bmatrix} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_i^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{bmatrix} \begin{bmatrix} 1 & \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})} \\ 1 & 0 \\ 0 & \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})} \end{bmatrix} \\
&= \left[M_y^a\left(z\right)\sigma_a^2, \quad \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})} \left(M_y^a\left(z\right)\sigma_a^2 + M_y^u\left(z\right)\sigma_u^2 \right) \right]
\end{aligned}
$$

Routine algebra reveals that

$$
\psi_y(z) = S_{yx}(z) \left(\Gamma^{-1}(z^{-1})\right)^\intercal = \left[(\frac{\sigma_a^2}{\sigma_e} - \frac{\sigma_a^4}{\sigma_p^2\sigma_e}) M_y^a(z) - \frac{\sigma_a^2\sigma_u^2}{\sigma_p^2\sigma_e} M_y^u(z), \quad \frac{M_y^a(z)\sigma_a^2 + M_y^u(z)\sigma_u^2}{\sigma_p} \right].
$$

Since $M_y^a\left(z\right), M_y^u\left(z\right) \in \mathbf{H}^2\left(\mathbb{D}\right)$, both components of $\psi_y(z)$ are in $\mathbf{H}^2\left(\mathbb{D}\right)$. Thus $\left[\psi_y(z)\right]_+ = \psi_y(z)$. In the innovation form, we have

$$
\begin{aligned}
\mathbb{E}_{it}[y_t] &= \left[\psi_y(L)\right]_+ \Gamma^{-1}(L) H(L) \eta_{it} \\
&= \left[(h_1+h_3) M_y^a(L) + (h_4-h_2) M_y^u(L),\ h_1 M_y^a(L) - h_2 M_y^u(L),\ h_3 M_y^a(L) + h_4 M_y^u(L) \right] \eta_{it},
\end{aligned}
$$

where we define

$$
\begin{aligned}
h_1 &\equiv \frac{\sigma_a^2}{\sigma_e^2} - \frac{\sigma_a^4}{\sigma_p^2\sigma_e^2},\ h_2 \equiv \frac{\sigma_a^2\sigma_u^2}{\sigma_p^2\sigma_e^2}, \\
h_3 &\equiv \frac{\sigma_a^2}{\sigma_p^2} + \frac{\sigma_a^6}{\sigma_p^4\sigma_e^2} - \frac{\sigma_a^4}{\sigma_p^2\sigma_e^2},\ h_4 \equiv \frac{\sigma_a^4\sigma_u^2}{\sigma_p^4\sigma_e^2} + \frac{\sigma_u^2}{\sigma_p^2}.
\end{aligned}
$$

Plugging $y_{it} = M_y(L)\eta_{it}$ and the preceding conditional expectation $\mathbb{E}_{it}[y_t]$ into (25) and matching coefficients, we obtain a system of linear equations

$$
\begin{aligned}
M_y^a(z) &= \frac{1}{\xi} + \theta \left[(h_1+h_3) M_y^a(z) + (h_4-h_2) M_y^u(z) \right], \\
M_y^u(z) &= \theta \left[h_3 M_y^a(z) + h_4 M_y^u(z) \right], \\
M_y^i(z) &= \frac{1}{\xi} + \theta \left[h_1 M_y^a(z) - h_2 M_y^u(z) \right],
\end{aligned}
$$

which yields the solution

$$
\begin{aligned}
M_y^a(z) &= \frac{1}{\xi m_1},\ M_y^u(z) = \frac{m_2}{\xi m_1}, \\
M_y^i(z) &= \frac{1}{\xi} + \theta \frac{h_1 - h_2 m_2}{\xi m_1},
\end{aligned}
$$

where we define

$$m_1 \equiv 1 - \frac{(h_4 - h_2)h_3\theta}{1-\theta h_4} - \theta(h_1+h_3),\ m_2 \equiv \frac{h_3\theta}{1-\theta h_4}.$$

The preceding solution is independent of z, confirming our previous conjecture.

Step 3. We proceed to the financial side of the model and compute the conditional expectations χ_{it} in (39). Using equation (43) and the Wiener-Hopf prediction formula, we compute the conditional expectation $\mathbb{E}_{it}\left[s^h_{it+2}\right] = \left[\psi_s\left(L\right)\right]_+ \Gamma^{-1}(L)X_{it}$. where the z-transform of the operator ψ_s is given by

$$\psi_s\left(z\right) = z^{-1}S_{sx}(z)\left(\Gamma^{-1}(z^{-1})\right)^{\intercal}$$

and the cross-spectrum is given by

$$\begin{aligned} S_{sx}(z) &= \left[0, M^i_s(z), 0\right] \left[\begin{array}{ccc} \sigma_a^2 & 0 & 0 \\ 0 & \sigma_i^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{array}\right] \begin{bmatrix} 1 & \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})} \\ 1 & 0 \\ 0 & \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})} \end{bmatrix} \\ &= \left[M^i_s\left(z\right)\sigma_i^2,\ \ 0\right]. \end{aligned}$$

Thus

$$\psi_s(z) = \frac{1}{z}\left[S_{sx}(z)\left(\Gamma^{-1}(z^{-1})\right)^{\intercal}\right] = \frac{1}{z}\left[\frac{\sigma_i^2}{\sigma_e}M^i_s(z),\ \ 0\right].$$

There is a pole at zero. Using a lemma in the Appendix A of Hansen and Sargent (1980) we can compute that

$$[\psi_s(z)]_+ = \psi_s(z) - \frac{\lim_{z\to 0} z\psi_s(z)}{z} = \frac{1}{z}\left[\frac{\sigma_i^2}{\sigma_e}(M^i_s(z) - M^i_s(0)),\ \ 0\right].$$

It follows that

$$\left[\psi_s\left(z\right)\right]_+ \Gamma^{-1}(z) = \left[\tau_1 \frac{M^i_s(z) - M^i_s(0)}{z}, \quad -\tau_2 \frac{1-\pi_2(z)}{\pi_1(z)} \frac{M^i_s(z) - M^i_s(0)}{z}\right],$$

where we define the signal-to-noise ratios

$$\tau_1 \equiv \frac{\sigma_i^2}{\sigma_e^2} \in (0,1)\,,\ \ \tau_2 \equiv \tau_1 \frac{\sigma_a^2}{\sigma_p^2} \in (0,1)\,. \tag{C.2}$$

By Lemmas 1 and 3, we derive

$$M^i_s(z) - M^i_s(0) = \frac{\pi_1\left(z\right)}{\alpha_1 - \alpha_2 z} - \frac{\pi_1\left(0\right)}{\alpha_1} = \frac{1}{\alpha_3}\left[\frac{\left(1-\lambda_s\right)\pi_1\left(z\right)}{1-\lambda_s z} - \left(1-\lambda_s\right)\pi_1\left(0\right)\right].$$

Thus

$$\mathbb{E}_{it}\left[s_{it+2}^{h}\right]=\frac{1}{\alpha_3 L}\left[\frac{(1-\lambda_s)\,\pi_1(L)}{1-\lambda_s L}-(1-\lambda_s)\,\pi_1(0)\right]\left[\tau_1,\ -\tau_2\frac{1-\pi_2(L)}{\pi_1(L)}\right]X_{it}. \quad \text{(C.3)}$$

Note that $\tau_2<\tau_1$ reflects the fact that equity prices do not fully aggregate information due to near-rational forecast errors.

Now we conjecture that $d_t = M_d(L)\,\eta_{it}$, $n_{it} = M_n(L)\,\eta_{it}$, and $b_{it} = M_b(L)\,\eta_{it}$, where

$$M_d(z)=\left[M_d^a(z),0,M_d^u(z)\right];\quad M_n(z)=\left[M_n^a(z),M_n^i(z),M_n^u(z)\right];\quad M_b(z)=\left[M_b^a(z),M_b^i(z),M_b^u(z)\right],$$

and each component of these vectors is in $\mathbf{H}^2(\mathbb{D})$. Plugging these equations and (36) into equations (30), (31), and $n_{it}=\frac{1}{\alpha}(y_{it}-a_{it})$, and matching coefficients, we can derive that

$$\begin{aligned}
M_d(z) &= \left[\frac{1}{\alpha_6}\left(1-\frac{\alpha_7}{\alpha}\right)M_y^a(z)+\frac{\alpha_7}{\alpha\alpha_6},\ 0,\ \frac{1}{\alpha_6}(1-\frac{\alpha_7}{\alpha})M_y^u(z)\right],\\
M_n(z) &= \frac{1}{\alpha}\left[M_y^a(z)-1,\ M_y^i(z)-1,\ M_y^u(z)\right],\\
M_b(z) &= \alpha_4 M_d(z)+\alpha_5 M_n(z).
\end{aligned}$$

Note that we have used the assumption that $a_{it}=\epsilon_{at}+\epsilon_{it}$. Since we have shown above that $M_y(z)$ is independent of z, $M_d(z)$, $M_n(z)$ and $M_b(z)$ are all independent of z.

Using the Wiener-Hopf prediction formula, we compute

$$\mathbb{E}_{it}\left[\Delta b_{it+1}\right]=\mathbb{E}_{it}\left[\left(1-L^{-1}\right)b_{it}\right]=\mathbb{E}_{it}\left[\left(1-L^{-1}\right)M_b(L)\,\eta_{it}\right]=\left[\psi_b(L)\right]_+\Gamma^{-1}(L)X_{it},$$

where the z-transform of the operator ψ_b is given by

$$\psi_b(z)=\frac{z-1}{z}S_{bx}(z)\left(\Gamma^{-1}(z^{-1})\right)^{\intercal},$$

and the cross-spectrum $S_{bx}(z)$ is given by

$$S_{bx}(z)=M_b(z)\Sigma_\eta H\left(z^{-1}\right)^{\intercal}=\left[M_b^a(z)\sigma_a^2+M_b^i(z)\sigma_i^2,\frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})}\left(M_b^a(z)\,\sigma_a^2+M_b^u(z)\,\sigma_u^2\right)\right].$$

It follows that

$$\begin{aligned}
\psi_b(z) &= \frac{z-1}{z}S_{bx}(z)\left(\Gamma^{-1}(z^{-1})\right)^{\intercal}\\
&= \frac{z-1}{z}\left[(\frac{\sigma_a^2}{\sigma_e}-\frac{\sigma_a^4}{\sigma_p^2\sigma_e})M_b^a(z)-\frac{\sigma_a^2\sigma_u^2}{\sigma_p^2\sigma_e}M_b^u(z)+\frac{\sigma_i^2}{\sigma_e}M_b^i(z),\quad \frac{M_b^a(z)\sigma_a^2+M_b^u(z)\sigma_u^2}{\sigma_p}\right]
\end{aligned}$$

The complex function $\psi_b(z)$ has a first-order pole at $z = 0$. Following Hansen and Sargent (1980), the annihilation operation is given by

$$[\psi_b(z)]_+ = \psi_b(z) - \frac{\lim_{z\to 0} z\psi_b(z)}{z}.$$

It follows immediately that

$$\begin{aligned}
[\psi_b(z)]_+ &= \psi_b(z) - \frac{(-1)}{z}\left[(\frac{\sigma_a^2}{\sigma_e} - \frac{\sigma_a^4}{\sigma_p^2\sigma_e})M_b^a(0) - \frac{\sigma_a^2\sigma_u^2}{\sigma_p^2\sigma_e}M_b^u(0) + \frac{\sigma_i^2}{\sigma_e}M_b^i(0), \quad \frac{M_b^a(0)\sigma_a^2 + M_b^u(0)\sigma_u^2}{\sigma_p}\right] \\
&= \psi_b(z) + \frac{1}{z}\left[h_1\sigma_e M_b^a(0) - h_2\sigma_e M_b^u(0) + \frac{\sigma_i^2}{\sigma_e}M_b^i(0), \quad \frac{M_b^a(0)\sigma_a^2 + M_b^u(0)\sigma_u^2}{\sigma_p}\right].
\end{aligned}$$

We can then derive that

$$\begin{aligned}
\mathbb{E}_{it}\left[\Delta b_{it+1}\right] &= \left[\psi_b\left(L\right)\right]_+ \Gamma^{-1}(L)X_{it} \\
&= \frac{1}{z}\left[G_b^{(1)}(L) - G_b^{(1)}(0), \quad \frac{1-\pi_2(L)}{\pi_1(L)}\left(G_b^{(2)}(L) - G_b^{(2)}(0)\right)\right]X_{it} \quad \text{(C.4)}
\end{aligned}$$

where we define the functions

$$\begin{aligned}
G_b^{(1)}(z) &\equiv (z-1)\left[h_1 M_b^a(z) - h_2 M_b^u(z) + \tau_1 M_b^i(z)\right], \\
G_b^{(2)}(z) &\equiv (z-1)\left[h_3 M_b^a(z) + h_4 M_b^u(z) - \tau_2 M_b^i(z)\right].
\end{aligned}$$

Using the same method, we can compute the conditional expectation of future dividends

$$\mathbb{E}_{it}\left[d_{t+1}\right] = \frac{1}{L}\left[G_d^{(1)}(L) - G_d^{(1)}(0), \quad \frac{1-\pi_2(L)}{\pi_1(L)}\left(G_d^{(2)}(L) - G_d^{(2)}(0)\right)\right]X_{it}, \qquad \text{(C.5)}$$

where we define the functions

$$G_d^{(1)}(z) \equiv h_1 M_d^a(z) - h_2 M_d^u(z); \quad G_d^{(2)}(z) \equiv h_3 M_d^a(z) + h_4 M_d^u(z).$$

Since $M_b\left(z\right), M_d\left(z\right)$, and $M_y\left(z\right)$ are constant independent of z, it follows from the previous construction that $G_d^{(1)}\left(z\right)$ and $G_d^{(2)}\left(z\right)$ are constant independent of z, but $G_b^{(1)}\left(z\right)$ and $G_b^{(2)}\left(z\right)$ are linear functions of z.

We then use the Wiener-Kolmogorov formula to compute $\mathbb{E}_{it}\left[q_{t+1}\right] = \left[\psi_q\left(L\right)\right]_+ \Gamma^{-1}\left(L\right)X_{it}$, where the z-transform of the operator ψ_q is given by

$$\psi_q\left(z\right) = \frac{1}{z}\left[0, \quad 1\right]S_x(z)\left(\Gamma^{-1}\left(z^{-1}\right)\right)^{\intercal} = \frac{1}{z}\left[0, \quad 1\right]\Gamma(z) = \frac{1}{z}\left[0, \quad \tfrac{\pi_1(z)}{1-\pi_2(z)}\sigma_p\right],$$

where the second equality follows from the previous definition of $S_x(z)$. Thus

$$\begin{aligned}
\left[\psi_q(z)\right]_+ \Gamma^{-1}(z) &= \left[\frac{1}{z}\left[0, \quad \tfrac{\pi_1(z)}{1-\pi_2(z)}\sigma_p\right]\right]_+ \Gamma^{-1}(z) \\
&= \frac{1}{z}\left[0, \quad \tfrac{\pi_1(z)}{1-\pi_2(z)}\sigma_p\right]\Gamma^{-1}(z) - \frac{1}{z}\left[0, \quad \tfrac{\pi_1(0)}{1-\pi_2(0)}\sigma_p\right]\Gamma^{-1}(z) \\
&= \left[0, \quad \tfrac{1}{z}\left(1 - \tfrac{1-\pi_2(z)}{\pi_1(z)}\tfrac{\pi_1(0)}{1-\pi_2(0)}\right)\right],
\end{aligned}$$

and

$$\mathbb{E}_{it}\left[q_{t+1}\right] = \left[0, \quad \tfrac{1}{z}\left(1 - \tfrac{1-\pi_2(z)}{\pi_1(z)}\tfrac{\pi_1(0)}{1-\pi_2(0)}\right)\right] X_{it}. \tag{C.6}$$

Step 4. Derive the solution for $\pi_1(z)$ and $\pi_2(z)$. Plugging the expressions for the conditional expectations (C.3), (C.4), (C.5), and (C.6) derived in Step 3 into equation (39), we obtain an expression for χ_{it}. Matching coefficients of $X_{it} = [a_{it}, q_t]'$ with those in (40), we construct the following equilibrium conditions:

$$\begin{aligned}
z\pi_1(z) &= \frac{1-\lambda_s}{1-\lambda_s z}\tau_1\pi_1(z) - (1-\lambda_s)\,\tau_1\pi_1(0) \\
&\quad + (1-\beta)\left[G_d^{(1)}(z) - G_d^{(1)}(0)\right] + G_b^{(1)}(z) - G_b^{(1)}(0),
\end{aligned} \tag{C.7}$$

and

$$\begin{aligned}
z\pi_2(z) = \frac{1-\pi_2(z)}{\pi_1(z)}\Bigg\{ &-\frac{1-\lambda_s}{1-\lambda_s z}\tau_2\pi_1(z) + (1-\lambda_s)\,\tau_2\pi_1(0) - \frac{\beta\pi_1(0)}{1-\pi_2(0)} \\
&+ (1-\beta)\left[G_d^{(2)}(z) - G_d^{(2)}(0)\right] + G_b^{(2)}(z) - G_b^{(2)}(0)\Bigg\} + \beta.
\end{aligned} \tag{C.8}$$

Simplifying equation (C.7) yields

$$\pi_1(z) = \frac{(1-\lambda_s z)\left[x(z) - (1-\lambda_s)\tau_1\pi_1(0)\right]}{P_1(z)}, \tag{C.9}$$

where we define the functions

$$P_1(z) \equiv -\lambda_s z^2 + z - (1-\lambda_s)\tau_1,$$

and

$$x(z) \equiv (1-\beta)\left[G_d^{(1)}(z) - G_d^{(1)}(0)\right] + G_b^{(1)}(z) - G_b^{(1)}(0). \tag{C.10}$$

By the analysis in Step 3, $x(z)$ is a linear function of z.

Since $\lambda_s \in (1/2, 1)$ by Lemma 1 and $\tau_1 \in (0,1)$, we have $P_1(0) = -(1-\lambda_s)\tau_1 < 0$, $P_1(1) = (1-\lambda_s)(1-\tau_1) > 0$, and $\lim_{z\to+\infty} P_1(z) = -\infty$. Thus $P_1(z) = 0$ has two real roots, denoted by $\gamma_1 \in (0,1)$ and $\gamma_2 > 1$. We can then write

$$\pi_1(z) = \frac{(1-\lambda_s z)}{-\lambda_s(z-\gamma_2)(z-\gamma_1)}\left[x(z) - (1-\lambda_s)\tau_1\pi_1(0)\right].$$

To remove the pole at γ_1, we set $\pi_1(0)$ such that

$$x(\gamma_1) - (1-\lambda_s)\tau_1\pi_1(0) = 0,$$

which implies that

$$\pi_1(0) = \frac{x(\gamma_1)}{(1-\lambda_s)\tau_1}.$$

We then collect terms and simplify expressions to derive

$$\pi_1(z) = \frac{(1-\lambda_s z)\left[x(z) - x(\gamma_1)\right]}{-\lambda_s(z-\gamma_1)(z-\gamma_2)}. \tag{C.11}$$

Since the pole $|\gamma_1| < 1$ is removed and $x(z)$ is a linear function of z, we deduce that $\pi_1(z) \in \mathbf{H}^2(\mathbb{D})$.

Next consider the equilibrium condition (C.8). It is straightforward to show that

$$\frac{\pi_1(z)}{1-\pi_2(z)} = \frac{\kappa(z) - \beta\pi_1(0)/(1-\pi_2(0))}{z-\beta}, \tag{C.12}$$

where we define the function

$$\begin{aligned}\kappa(z) \equiv{} & (1-\beta)[G_d^{(2)}(z) - G_d^{(2)}(0)] + G_b^{(2)}(z) - G_b^{(2)}(0) \\ & - \left[\frac{(1-\lambda_s)\tau_2}{1-\lambda_s z} - z\right]\pi_1(z) + (1-\lambda_s)\tau_2\pi_1(0).\end{aligned} \tag{C.13}$$

Since $\pi_1(z) \in \mathbf{H}^2(\mathbb{D})$, $\lambda_s \in (1/2, 1)$ by Lemma 1, $G_d^{(2)}(z)$ is a constant, and $G_b^{(2)}(z)$ is linear in z, it follows that $\kappa(z) \in \mathbf{H}^2(\mathbb{D})$.

As mentioned earlier, we need $\frac{\pi_1(z)}{1-\pi_2(z)}$ to be analytical in the unit disk. Thus we should remove the pole at $z = \beta$ by setting the constant $\pi_2(0)$ such that $\kappa(\beta) - \beta\pi_1(0)/(1-\pi_2(0)) = 0$. Solving this equation yields

$$\pi_2(0) = 1 - \frac{\pi_1(0)\beta}{\kappa(\beta)}.$$

We can then rewrite (C.12) as

$$\frac{\pi_1(z)}{1-\pi_2(z)} = \frac{\kappa(z) - \kappa(\beta)}{z-\beta}. \tag{C.14}$$

Since we have removed the pole at $z=\beta$ and $\kappa(z) \in \mathbf{H}^2(\mathbb{D})$, it follows that $\frac{\pi_1(z)}{1-\pi_2(z)} \in \mathbf{H}^2(\mathbb{D})$.

By our constructive proof above, we conclude that the equilibrium solution is characterized by unique rational functions of z in the frequency domain. As mentioned earlier to ensure the spectral factorization to be valid, we need to impose the assumption that the equation

$$\frac{\pi_1(z)}{1-\pi_2(z)} = \frac{\kappa(z)-\kappa(\beta)}{z-\beta} = 0$$

has no roots inside the open unit disk. The proof is then complete. Q.E.D.

Proof of Theorem 3: We first show that the denominator of the expression for $\pi_1(z)$ in (C.9) has a unit root as $\sigma_i \to \infty$. Consider the quadratic function,

$$P_1(z) \equiv -\lambda_s z^2 + z - (1-\lambda_s)\tau_1.$$

Since

$$\lim_{\sigma_i \to \infty} \tau_1 = \lim_{\sigma_i \to \infty} \frac{\sigma_i^2}{\sigma_e^2} = 1,$$

we have

$$\lim_{\sigma_i \to \infty} P_1(z) = -\lambda_s z^2 + z - (1-\lambda_s) = -\lambda_s (z-1)\left(z - \frac{1-\lambda_s}{\lambda_s}\right).$$

Since $\lambda_s \in (1/2, 1)$, the root $\frac{1-\lambda_s}{\lambda_s}$ is located inside the unit circle. We know that $P_1(z)$ has one root inside the unit circle and the other outside the unit circle. By the continuous dependence of roots on coefficients, the larger root γ_2 of $P_1(z)$ gradually converges to the unit root as $\sigma_i \to \infty$.

We next show that the numerator of $\pi_1(z)$ in (C.9) or (C.11) does have a zero at $z=1$ when $\sigma_i \to \infty$. By (C.11), it suffices to show that the analytic function $x(z) - x(\gamma_1)$ does not have a zero at $z=1$. Using the result derived in the proof of Theorem 2, we can show that $\lim_{\sigma_i \to \infty} h_1 = \lim_{\sigma_i \to \infty} h_2 = 0$, $\lim_{\sigma_i \to \infty} G_d^{(1)}(z) = 0$, and $\lim_{\sigma_i \to \infty} G_b^{(1)}(z) = \frac{\alpha_5}{\alpha}\left(\frac{1}{\xi} - 1\right)(z-1)$. It follows from (C.10) that $\lim_{\sigma_i \to \infty} x(z) = \frac{\alpha_5(1-\xi)}{\alpha\xi} z$. Therefore,

$$\lim_{\sigma_i \to \infty} [x(1) - x(\gamma_1)] = \frac{\alpha_5(1-\xi)}{\alpha\xi}(1 - \lim_{\sigma_i \to \infty} \gamma_1) = \frac{\alpha_5(1-\xi)}{\alpha\xi}\left(1 - \frac{1-\lambda_s}{\lambda_s}\right).$$

By Assumption 1, we know that $\xi \equiv \frac{1+\phi-\alpha(1-1/\varsigma)}{1+\phi} \in (0,1)$, as $\phi > 0$ and $\varsigma > 1$. It follows from $\lambda_s \in (\frac{1}{2}, 1)$ that

$$\lim_{\sigma_i \to \infty} [x(1) - x(\gamma_1)] \neq 0.$$

Hence, $\pi_1(z)$ does not converge to zero at $z = 1$ when $\sigma_i \to \infty$, but it has a pole at $z = \lim_{\sigma_i \to \infty} \gamma_2 = 1$. Since $\pi_1(z)$ is rational in the frequency domain, a pole at the unit circle is sufficient to ensure that

$$\lim_{\sigma_i \to \infty} \left|\left|\pi_1(z)\right|\right|_{\mathbf{H}^2} \to \infty.$$

Now we wish to show that the analytic function $\frac{\pi_1(z)}{1-\pi_2(z)}$ has no zero at $z = 1$ as $\sigma_i \to \infty$. We only need to consider the equation $\kappa(z) - \kappa(\beta) = 0$ by (C.14). We can rewrite the expression for $\kappa(z)$ in (C.13) as

$$\kappa(z) = A(z) - \left[\frac{(1-\lambda_s)\tau_2}{1-\lambda_s z} - z\right]\pi_1(z) + (1-\lambda_s)\tau_2\pi_1(0),$$

where $A(z)$ is a linear function of z. Plugging (C.11) into this equation, we can derive

$$\begin{aligned}\kappa(z) - \kappa(\beta) \quad &= \quad A(z) + (1-\lambda_s)\tau_2\pi_1(0) - \kappa(\beta) \\ &\quad + \frac{[x(z) - x(\gamma_1)]\left[(1-\lambda_s)\tau_2 - (1-\lambda_s z)z\right]}{\lambda_s(z-\gamma_1)(z-\gamma_2)}.\end{aligned}$$

The linear function $A(z)$ is bounded in the closed unit disk, $\sup_{|z|\leq 1}|A(z)| < \infty$. Moreover, we know that $\pi_1(z)$ is analytic and rational inside the open unit disk $|z| < 1$, even when $\sigma_i \to \infty$. Thus $\pi_1(0)$ and $\kappa(\beta)$ are finite for $0 < \beta < 1$. It follows that the expression on the first line of the right-hand side of the equation above is bounded at $z = 1$ when $\sigma_i \to \infty$.

Consider the expression on the second line of the equation above. The denominator converges to zero at $z = 1$ as $\sigma_i \to \infty$. For the numerator, we have

$$\begin{aligned}&\lim_{\sigma_i \to \infty} [x(z) - x(\gamma_1)]\left[(1-\lambda_s)\tau_2 - (1-\lambda_s z)z\right]\big|_{z=1} \\ &= \left[\frac{\alpha_5(1-\xi)}{\alpha\xi}\left(1 - \frac{1-\lambda_s}{\lambda_s}\right)\right](1-\lambda_s)\left(\frac{\sigma_a^2}{\sigma_p^2} - 1\right) \neq 0,\end{aligned}$$

where we have used the previous definition of σ_e^2 to derive

$$\lim_{\sigma_i \to \infty} \tau_2 = \lim_{\sigma_i \to \infty} \frac{\sigma_a^2\sigma_i^2}{\sigma_e^2\sigma_p^2} = \frac{\sigma_a^2}{\sigma_p^2} \in (0,1).$$

We conclude that $\kappa(1) - \kappa(\beta)$ converges to infinity as $\sigma_i \to \infty$.

Therefore, $\frac{\pi_1(z)}{1-\pi_2(z)}$ does not have a zero at $z = 1$ when $\sigma_i \to \infty$, but it has a pole at $z = \lim_{\sigma_i \to \infty} \gamma_2 = 1$. This implies that the rational function $\frac{\pi_1(z)}{1-\pi_2(z)}$ will have infinite norm at the limit,

$$\lim_{\sigma_i \to \infty} \left|\left|\frac{\pi_1(z)}{1-\pi_2(z)}\right|\right|_{\mathbf{H}^2} \to \infty.$$

This completes the proof. Q.E.D.

Online Supplementary Appendix

Jianjun Miao Jieran Wu Eric Young

S1 Analysis in Section 6

S1.1 Equilibrium System

As in Section 5, we derive the equilibrium system in four steps.

Step 1. Derive the Wold representation for the signal system under Assumption 3. Given the AR(1) processes for a_t and u_t, the signal representation follows

$$X_{it} = H(L)\eta_{it} \equiv \begin{bmatrix} \frac{1}{1-\rho_a L} & 1 & 0 \\ \frac{\pi_1(L)}{\left(1-\pi_2(L)\right)\left(1-\rho_a L\right)} & 0 & \frac{1}{\left(1-\pi_2(L)\right)\left(1-\rho_u L\right)} \end{bmatrix} \begin{bmatrix} \epsilon_{at} \\ \epsilon_{it} \\ \epsilon_{ut} \end{bmatrix}, \tag{S1.1}$$

and so the spectral density for the signal is

$$S_x(z) = H(z)\Sigma_{\boldsymbol{\eta}} H(z^{-1})^\intercal = \begin{bmatrix} \frac{\sigma_a^2}{(1-\rho_a z)(1-\rho_a z^{-1})} + \sigma_i^2 & \frac{\pi_1(z^{-1})}{1-\pi_2(z^{-1})} \frac{1}{(1-\rho_a z)(1-\rho_a z^{-1})} \sigma_a^2 \\ \frac{\pi_1(z)}{1-\pi_2(z)} \frac{1}{(1-\rho_a z)(1-\rho_a z^{-1})} \sigma_a^2 & \frac{\frac{\pi_1(z)\pi_1(z^{-1})}{(1-\rho_a z)(1-\rho_a z^{-1})}\sigma_a^2 + \frac{\sigma_u^2}{(1-\rho_u z)(1-\rho_u z^{-1})}}{(1-\pi_2(z))(1-\pi_2(z^{-1}))} \end{bmatrix}.$$

Using the method presented in Appendix S3, we can first factorize the spectral density in a lower triangular form

$$\tilde{\Gamma}(z) = \begin{bmatrix} \sigma_w \frac{z-\lambda_w}{1-\rho_a z} & 0 \\ \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z) z}{(1-\pi_2(z))(1-\lambda_w z)(1-\rho_a z)} & \frac{1}{\sigma_w} \frac{\tilde{\pi}_1(z)}{1-\pi_2(z)} \frac{1-\rho_a z}{1-\lambda_w z} \end{bmatrix},$$

where the constants $\lambda_w \in (0,1)$ and σ_w are determined by the univariate spectral factorization of the first signal a_{it} in the frequency domain,

$$\sigma_w^2 \frac{(1-\lambda_w z)(1-\lambda_w z^{-1})}{(1-\rho_a z)(1-\rho_a z^{-1})} = \frac{\sigma_a^2}{(1-\rho_a z)(1-\rho_a z^{-1})} + \sigma_i^2.$$

It follows that

$$\sigma_w^2(1-\lambda_w z)(1-\lambda_w z^{-1}) = \sigma_a^2 + \sigma_i^2(1-\rho_a z)(1-\rho_a z^{-1}).$$

Matching coefficients on the two sides of the equality yields

$$\lambda_w = \frac{1}{2\rho_a}\left[\left(1+\tau+\rho_a^2\right) - \sqrt{\tau^2+2\tau+2\tau\rho_a^2+1-2\rho_a^2+\rho_a^4}\right],$$

and $\sigma_w^2 = \frac{\rho_a \sigma_i^2}{\lambda_w}$. Here $\tau \equiv \sigma_a^2/\sigma_i^2 \in (0,\infty)$ denotes the relative volatility of the aggregate shock to the idiosyncratic shock. It is easy to verify that $0 < \lambda_w < \rho_a < 1$ and $\lim_{\sigma_i\to\infty}\lambda_w = \rho_a$.

Define the function $\tilde{\pi}_1(z)$ by the following equation

$$\tilde{\pi}_1(z)\tilde{\pi}_1(z^{-1}) = \frac{\pi_1(z)\pi_1(z^{-1})\sigma_a^2\sigma_i^2}{(1-\rho_a z)(1-\rho_a z^{-1})} + \frac{(1-\lambda_w z)(1-\lambda_w z^{-1})\sigma_u^2\sigma_w^2}{(1-\rho_a z)(1-\rho_a z^{-1})(1-\rho_u z)(1-\rho_u z^{-1})}. \tag{S1.2}$$

A stationary equilibrium requires that the endogenous function $\pi_1 \in \mathbf{H}^2(\mathbb{D})$. It is then clear that the right-hand side of equation (S1.2) is a well-defined spectral density supported by a stationary process. Then by the Paley-Wiener Theorem (e.g. Lindquist and Picci, 2015, Theorem 4.4.1), there exists a Wold spectral factor $\tilde{\pi}_1(z) \in \mathbf{H}^2(\mathbb{D})$ that satisfies the factorization (S1.2). Using a similar argument, we can show that the function $\frac{\tilde{\pi}_1(z)}{1-\pi_2(z)} \in \mathbf{H}^2(\mathbb{D})$. Hence, the matrix $\tilde{\Gamma}(z)$ is a valid spectral factor in $\mathbf{H}^2(\mathbb{D})$ that satisfies $S_x(z) = \tilde{\Gamma}(z)\,\tilde{\Gamma}^\intercal(z^{-1})$. The determinant of $\tilde{\Gamma}(z)$ is given by

$$\det\tilde{\Gamma}(z) = \frac{\tilde{\pi}_1(z)}{1-\pi_2(z)}\frac{z-\lambda_w}{1-\lambda_w z}.$$

As in Section 5, we restrict our attention to the equilibrium such that $\frac{\tilde{\pi}_1(z)}{1-\pi_2(z)}$ has no roots in the open unit disk. To derive the wold fundamental representation, we need to remove the root at $z=\lambda_w \in (0,1)$. Using the Blaschke matrix $B(z)$ in Step 2 of Appendix S3, we set

$$\Gamma(z) = \tilde{\Gamma}(z)V^{-1}B(z),$$

where

$$V = \begin{bmatrix} \sqrt{\frac{h^2}{1+h^2}} & \sqrt{\frac{1}{1+h^2}} \\ \sqrt{\frac{1}{1+h^2}} & -\sqrt{\frac{h^2}{1+h^2}} \end{bmatrix} = \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{bmatrix},\ B(z) = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1-\lambda_w z}{z-\lambda_w} \end{bmatrix}.$$

Here the constant

$$h \equiv \frac{\pi_1(\lambda_w)\lambda_w\sigma_a^2}{\tilde{\pi}_1(\lambda_w)(1-\rho_a\lambda_w)^2}$$

is endogenous and will be determined in equilibrium. The unitary matrix V is symmetric and satisfies $V = V^{\intercal} = V^{-1}$, and $\det V = -1$. We then obtain the Wold fundamental matrix

$$\Gamma(z) = \begin{bmatrix} \sigma_w \frac{z-\lambda_w}{1-\rho_a z} V_{11} & \sigma_w \frac{1-\lambda_w z}{1-\rho_a z} V_{12} \\ \\ \Gamma_\pi^{(1)}(z) & \Gamma_\pi^{(2)}(z) \end{bmatrix},$$

where we define

$$\Gamma_\pi^{(1)}(z) \equiv \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z) z}{(1-\pi_2(z))(1-\lambda_w z)(1-\rho_a z)} V_{11} + \frac{1}{\sigma_w} \frac{\tilde{\pi}_1(z)}{1-\pi_2(z)} \frac{1-\rho_a z}{1-\lambda_w z} V_{12},$$
$$\Gamma_\pi^{(2)}(z) \equiv \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z) z}{(1-\pi_2(z))(z-\lambda_w)(1-\rho_a z)} V_{12} + \frac{1}{\sigma_w} \frac{\tilde{\pi}_1(z)}{1-\pi_2(z)} \frac{1-\rho_a z}{z-\lambda_w} V_{22}.$$

We compute that

$$\Gamma^{-1}(z) = \begin{bmatrix} G_1(z) \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z)}{\tilde{\pi}_1(z)} + G_2(z) \frac{1}{\sigma_w} & -\frac{1-\pi_2(z)}{\tilde{\pi}_1(z)} \sigma_w G_3(z) \\ \\ -\left[G_4(z) \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z)}{\tilde{\pi}_1(z)} + G_5(z) \frac{1}{\sigma_w} \right] & \frac{1-\pi_2(z)}{\tilde{\pi}_1(z)} \sigma_w G_6(z) \end{bmatrix},$$

where we define

$$G_1(z) = -V_{12} \frac{z}{(z-\lambda_w)(1-\rho_a z)}, \quad G_2(z) = -V_{22} \frac{1-\rho_a z}{z-\lambda_w},$$
$$G_3(z) = -V_{12} \frac{1-\lambda_w z}{1-\rho_a z}, \quad G_4(z) = -V_{11} \frac{z}{(1-\lambda_w z)(1-\rho_a z)},$$
$$G_5(z) = -V_{12} \frac{1-\rho_a z}{1-\lambda_w z}, \quad G_6(z) = -V_{11} \frac{z-\lambda_w}{1-\rho_a z}.$$

Note that all $G_1(z), ..., G_6(z)$ are independent of the endogenous price signal except for the constant in V. We also define the following functions that will be repeatedly used later:

$$\Gamma_I^{(1)}(z) = G_1(z) \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z)}{\tilde{\pi}_1(z)} + G_2(z) \frac{1}{\sigma_w}, \quad \Gamma_I^{(3)}(z) \equiv \sigma_w G_3(z) \frac{\pi_1(z)}{\tilde{\pi}_1(z)}$$
$$\Gamma_I^{(2)}(z) = G_4(z) \frac{\sigma_a^2}{\sigma_w} \frac{\pi_1(z)}{\tilde{\pi}_1(z)} + G_5(z) \frac{1}{\sigma_w}, \quad \Gamma_I^{(4)}(z) \equiv \sigma_w G_6(z) \frac{\pi_1(z)}{\tilde{\pi}_1(z)}.$$

By the Paley-Wiener Theorem and the fact that $\tilde{\pi}_1(z)$ is analytic in the open unit disk and Wold fundamental, these functions are analytic in the open unit disk.[17]

Step 2. Solve for the equilibrium quantities. We conjecture that $y_{it} = M_y(L)\eta_{it}$, where $M_y(z) = \left[M_y^a(z), M_y^i(z), M_y^u(z)\right]$ and $M_y^a(z)$, $M_y^i(z)$, and $M_y^u(z)$ are all in

[17] Sayed and Kailath (2001) summarized the property of the Wold fundamental matrix implied by the Paley-Wiener theorem.

$\mathbf{H}^2(\mathbb{D})$. Aggregation leads to aggregate output $y_t = M_y(z) I_y \eta_{it}$, where I_y is defined earlier. Using the Wiener-Hopf prediction formula, we derive that

$$\mathbb{E}_{it}[y_t] = \left[\psi_y^{(1)}(L) \;\; \psi_y^{(2)}(L)\right]_+ \Gamma^{-1}(L) H(L) \eta_{it},$$

in terms of innovations, where the z-transform of the operator $\psi_y = \left[\psi_y^{(1)} \;\; \psi_y^{(2)}\right]$ is given by

$$\psi_y(z) = z^{-1} S_{yx}(z) \left(\Gamma^{-1}(z^{-1})\right)^\intercal. \tag{S1.3}$$

The annihilation is given by $\left[\psi_y^{(1)}(z)\right]_+ = \psi_y^{(1)}(z) - P_y^{(1)}(z)$ and $\left[\psi_y^{(2)}(z)\right]_+ = \psi_y^{(2)}(z) - P_y^{(2)}(z)$, where $P_y^{(1)}(z)$ and $P_y^{(2)}(z)$ denote the negative powers of z in the Laurent series expansions of $\psi_y^{(1)}(z)$ and $\psi_y^{(2)}(z)$, respectively. There are no explicit formulas for $P_y^{(1)}(z)$ and $P_y^{(2)}(z)$ in general.

Using (S1.3), $y_t = M_y(z) I_y \eta_{it}$, and the cross-spectrum

$$S_{yx} = M_y(z) I_y \Sigma_\eta H^\intercal(z^{-1}) = \left[M_y^a, 0, M_y^u\right] \begin{bmatrix} \sigma_a^2 & & \\ & \sigma_u^2 & \\ & & \sigma_u^2 \end{bmatrix} \begin{bmatrix} \frac{1}{1-\rho_a z^{-1}} & \frac{\pi_1(z^{-1})}{(1-\pi_2(z^{-1}))(1-\rho_a z^{-1})} \\ 1 & 0 \\ 0 & \frac{1}{(1-\pi_2(z^{-1}))(1-\rho_u z^{-1})} \end{bmatrix},$$

we can derive

$$\psi_y^{(1)}(z) = M_y^a(z) \sigma_a^2 A_n^{(1)}(z) - M_y^u(z) \sigma_u^2 A_n^{(2)}(z); \quad \psi_y^{(2)}(z) = -M_y^a(z) \sigma_a^2 A_n^{(3)}(z) + M_y^u(z) \sigma_u^2 A_n^{(4)}(z),$$

where we define

$$A_n^{(1)}(z) = \frac{1}{1-\rho_a z^{-1}} \left[\Gamma_I^{(1)}(z^{-1}) - \Gamma_I^{(3)}(z^{-1})\right]; \quad A_n^{(2)}(z) = \frac{1}{1-\rho_u z^{-1}} \frac{1}{\pi_1(z^{-1})} \Gamma_I^{(3)}(z^{-1}),$$
$$A_n^{(3)}(z) = \frac{1}{1-\rho_a z^{-1}} \left[\Gamma_I^{(2)}(z^{-1}) - \Gamma_I^{(4)}(z^{-1})\right]; \quad A_n^{(4)}(z) = \frac{1}{1-\rho_u z^{-1}} \frac{1}{\pi_1(z^{-1})} \Gamma_I^{(4)}(z^{-1}).$$

Substituting the preceding expression for $\mathbb{E}_{it}[y_t]$ into (25) and matching coefficients for η_{it}, we obtain

$$M_y^a(z) = \frac{1}{\xi} \frac{1}{1-\rho_a z} + \frac{1}{1-\rho_a z} \left[G_y^{(1)}(z) - A_y^{(1)}(z) + G_y^{(2)}(z) - A_y^{(2)}(z)\right] \theta, \tag{S1.4}$$

$$M_y^i(z) = \frac{1}{\xi} + \left[G_y^{(1)}(z) - A_y^{(1)}(z)\right] \theta, \tag{S1.5}$$

$$M_y^u(z) = \frac{1}{1-\rho_u z} \frac{\theta}{\pi_1(z)} \left[G_y^{(2)}(z) - A_y^{(2)}(z)\right], \tag{S1.6}$$

where we define

$$G_y^{(1)}(z) = \psi_y^{(1)}(z) \Gamma_I^{(1)}(z) - \psi_y^{(2)} \Gamma_I^{(2)}(z); \qquad A_y^{(1)}(z) = P_y^{(1)}(z) \Gamma_I^{(1)}(z) - P_y^{(2)} \Gamma_I^{(2)}(z),$$
$$G_y^{(2)}(z) = \psi_y^{(2)}(z) \Gamma_I^{(4)}(z) - \psi_y^{(1)}(z) \Gamma_I^{(3)}(z); \quad A_y^{(2)}(z) = P_y^{(2)}(z) \Gamma_I^{(4)}(z) - P_y^{(1)} \Gamma_I^{(3)}(z).$$

Here $\Gamma_I^{(1)}(z), ..., \Gamma_I^{(4)}(z)$ are defined earlier.

Using equations (S1.4) and (S1.6) and the definition of $G_y^{(1)}(z)$ and $G_y^{(2)}(z)$, we can derive that

$$\begin{bmatrix} Q_1(z) & Q_2(z) \\ Q_3(z) & Q_4(z) \end{bmatrix} \begin{bmatrix} M_y^a(z) \\ M_y^u(z) \end{bmatrix} = \begin{bmatrix} \frac{1}{\xi} - A_y^{(1)}(z)\theta - A_y^{(2)}(z)\theta \\ -A_y^{(2)}(z)\theta \end{bmatrix}, \tag{S1.7}$$

where we define

$$\begin{aligned} Q_1(z) &= (1-\rho_a z) - \theta\sigma_a^2 H_a(z); \quad Q_2(z) = \theta\sigma_u^2 H_u(z), \\ Q_3(z) &= \theta\sigma_a^2 H_d(z); \quad Q_4(z) = (1-\rho_u z)\pi_1(z) - \theta\sigma_u^2 H_c(z), \end{aligned}$$

and

$$\begin{aligned} H_a(z) &= A_n^{(1)}(z)\left(\Gamma_I^{(1)}(z) - \Gamma_I^{(3)}(z)\right) + A_n^{(3)}(z)\left(\Gamma_I^{(2)}(z) - \Gamma_I^{(4)}(z)\right), \\ H_u(z) &= A_n^{(2)}(z)\left(\Gamma_I^{(1)}(z) - \Gamma_I^{(3)}(z)\right) + A_n^{(4)}(z)\left(\Gamma_I^{(2)}(z) - \Gamma_I^{(4)}(z)\right), \\ H_c(z) &= A_n^{(4)}(z)\Gamma_I^{(4)}(z) + A_n^{(2)}\Gamma_I^{(3)}(z), \\ H_d(z) &= A_n^{(3)}(z)\Gamma_I^{(4)}(z) + A_n^{(1)}\Gamma_I^{(3)}(z). \end{aligned}$$

Once $\pi_1(z)$ and $\pi_2(z)$ are known, we can use the system (S1.7) to determine $M_y^a(z)$ and $M_y^u(z)$. Equation (S1.5) then determines $M_y^i(z)$.

As in the proof of Theorem 2, we deduce that $d_t = M_d(L)\eta_{it}$, $n_{it} = M_n(L)\eta_{it}$, and $b_{it} = M_b(L)\eta_{it}$, where

$$\begin{aligned} M_d(z) &= \left[\frac{1}{\alpha_6}\left(1-\frac{\alpha_7}{\alpha}\right)M_y^a(z) + \frac{\alpha_7}{\alpha\alpha_6}\frac{1}{1-\rho_a z},\ 0,\ \frac{1}{\alpha_6}(1-\frac{\alpha_7}{\alpha})M_y^u(z)\right], && \text{(S1.8)} \\ M_n(z) &= \frac{1}{\alpha}\left[M_y^a(z) - \frac{1}{1-\rho_a z},\ M_y^i(z) - 1,\ M_y^u(z)\right], && \text{(S1.9)} \\ M_b(z) &= \alpha_4 M_d(z) + \alpha_5 M_n(z). && \text{(S1.10)} \end{aligned}$$

Each component of these vectors is in $\mathbf{H}^2(\mathbb{D})$.

Step 3. We proceed to the financial side of the model. We need to compute several conditional expectations for χ_{it} in (39). First, we use the Wiener-Hopf formula to derive

$$\alpha_3 \mathbb{E}_{it}\left[s_{it+2}^h\right] = \alpha_3 \left[\psi_s(L)\right]_+ \Gamma^{-1}(L) X_{it},$$

where the z-transform of the operator ψ_s is given by $\psi_s(z) = z^{-1}S_{sx}(z)\left(\Gamma^{-1}(z^{-1})\right)^{\intercal}$, and

$$\alpha_3\left[\psi_s^{(1)}(z)\right]_+ = \alpha_3\psi_s^{(1)}(z) - P_s^{(1)}(z); \quad \alpha_3\left[\psi_s^{(2)}(z)\right]_+ = \alpha_3\psi_s^{(2)}(z) - P_s^{(2)}(z).$$

Here $P_s^{(1)}(z)$ and $P_s^{(2)}(z)$ denote the negative powers of z in the Laurent series expansions of $\alpha_3\psi_s^{(1)}(z)$ and $\alpha_3\psi_s^{(2)}(z)$, respectively. It follows that

$$\alpha_3\left[\psi_s^{(1)}(z),\ \ \psi_s^{(2)}(z)\right]_+\Gamma^{-1}(z)=\left[G_s^{(1)}(z)-A_s^{(1)}(z),\ \ \frac{1-\pi_2(z)}{\pi_1(z)}\left(G_s^{(2)}(z)-A_s^{(2)}(z)\right)\right],$$

where

$$\begin{aligned}G_s^{(1)}(z)&=\sigma_i^2z^{-1}\alpha_3M_s^i(z)\left[\Gamma_I^{(1)}(z)\Gamma_I^{(1)}(z^{-1})+\Gamma_I^{(2)}(z)\Gamma_I^{(2)}(z^{-1})\right],\\ G_s^{(2)}(z)&=\sigma_i^2z^{-1}\alpha_3M_s^i(z)\left[-\Gamma_I^{(3)}(z)\Gamma_I^{(1)}(z^{-1})-\Gamma_I^{(4)}(z)\Gamma_I^{(2)}(z^{-1})\right],\end{aligned}$$

and

$$A_s^{(1)}(z)=P_s^{(1)}(z)\Gamma_I^{(1)}(z)-P_s^{(2)}(z)\Gamma_I^{(2)}(z);\quad A_s^{(2)}(z)=P_s^{(2)}(z)\Gamma_I^{(4)}(z)-P_s^{(1)}(z)\Gamma_I^{(3)}(z).$$

It is easy to verify that Lemma 3 continues to hold, which implies

$$\begin{aligned}G_s^{(1)}(z)&=\sigma_i^2\frac{1-\lambda_s}{z(1-\lambda_sz)}\pi_1(z)\left[\Gamma_I^{(1)}(z)\Gamma_I^{(1)}(z^{-1})+\Gamma_I^{(2)}(z)\Gamma_I^{(2)}(z^{-1})\right],\\ G_s^{(2)}(z)&=\sigma_i^2\frac{1-\lambda_s}{z(1-\lambda_sz)}\pi_1(z)\left[-\Gamma_I^{(3)}(z)\Gamma_I^{(1)}(z^{-1})-\Gamma_I^{(4)}(z)\Gamma_I^{(2)}(z^{-1})\right].\end{aligned}$$

Second, the Wiener-Hopf formula gives

$$\mathbb{E}_{it}\left[q_{t+1}\right]=\left[\psi_q\left(L\right)\right]_+\Gamma^{-1}(L)X_{it},$$

where the z-transform of the operator ψ_q is given by

$$\psi_q\left(z\right)=\frac{1}{z}\left[0,\ \ 1\right]S_x(z)\left(\Gamma^{-1}\left(z^{-1}\right)\right)^{\intercal}=\frac{1}{z}\left[0,\ \ 1\right]\Gamma(z)=z^{-1}\left[\Gamma_\pi^{(1)}(z)\ \ \ \Gamma_\pi^{(2)}(z)\right],$$

where $\Gamma_\pi^{(1)}(z)$ and $\Gamma_\pi^{(2)}(z)$ are defined earlier. Since $z=0$ is the only inside pole of $\psi_q(z)$, it follows from the lemma in Appendix A of Hansen and Sargent (1980) that

$$\left[\psi_q\left(L\right)\right]_+\Gamma^{-1}(z)=z^{-1}\left[0\ \ 1\right]-P_q\left(z\right)\Gamma^{-1}(z),$$

where

$$P_q\left(z\right)=z^{-1}\left[\tfrac{1}{\sigma_w}\tfrac{\tilde{\pi}_1(0)}{1-\pi_2(0)}V_{12},\ \ \tfrac{1}{\sigma_w}\tfrac{\tilde{\pi}_1(0)}{1-\pi_2(0)}\left(-\tfrac{1}{\lambda_w}\right)V_{22}\right].$$

Thus

$$\begin{aligned}\mathbb{E}_{it}\left[q_{t+1}\right]\ =\ &-z^{-1}\frac{1}{\sigma_w}\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}\left[V_{12}\Gamma_I^{(1)}(z)+\frac{1}{\lambda_w}V_{22}\Gamma_I^{(2)}(z)\right]a_{it}\\ &+z^{-1}\frac{1}{\sigma_w}\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}\left[V_{12}\Gamma_I^{(3)}(z)+\frac{1}{\lambda_w}V_{22}\Gamma_I^{(4)}(z)\right]\frac{1-\pi_2(z)}{\pi_1(z)}q_t.\end{aligned}$$

Third, the Wiener-Hopf formula gives

$$\mathbb{E}_{it}\left[d_{t+1}\right] = \left[\psi_d\left(L\right)\right]_+ \Gamma^{-1}(L) X_{it},$$

where the z-transform of the operator ψ_d is given by

$$\psi_d\left(z\right) = \left[\psi_d^{(1)}(z),\;\; \psi_d^{(2)}(z)\right] = z^{-1} S_{dx}(z) \left(\Gamma^{-1}(z^{-1})\right)^{\intercal},$$

and $\left[\psi_d^{(1)}(z)\right]_+ = \psi_d^{(1)}(z) - P_d^{(1)}(z)$, $\left[\psi_d^{(2)}(z)\right]_+ = \psi_s^{(2)}(z) - P_d^{(2)}(z)$. Here $P_d^{(1)}(z)$ and $P_d^{(2)}(z)$ denote the negative powers of z in the Laurent series expansions of $\psi_d^{(1)}(z)$ and $\psi_d^{(2)}(z)$, respectively. As in Step 2 we can compute that

$$\begin{aligned}
\psi_d^{(1)}(z) &= z^{-1}\left[M_d^a(z) A_n^{(1)}(z)\sigma_a^2 - M_d^u(z) A_n^{(2)}(z)\sigma_u^2\right], \\
\psi_d^{(2)}(z) &= z^{-1}\left[-M_d^a(z) A_n^{(3)}(z)\sigma_a^2 + M_d^u(z) A_n^{(4)}(z)\sigma_u^2\right].
\end{aligned}$$

It follows that

$$\mathbb{E}_{it}\left[d_{t+1}\right] = \left[G_d^{(1)}(L) - A_d^{(1)}(L),\;\; \frac{1-\pi_2(z)}{\pi_1(z)}\left(G_d^{(2)}(L) - A_d^{(2)}(L)\right)\right] X_{it},$$

where

$$G_d^{(1)}(z) = \psi_d^{(1)}(z)\Gamma_I^{(1)}(z) - \psi_d^{(2)}\Gamma_I^{(2)}(z); \quad G_d^{(2)}(z) = \psi_d^{(2)}(z)\Gamma_I^{(4)}(z) - \psi_d^{(1)}(z)\Gamma_I^{(3)}(z),$$

and

$$A_d^{(1)}(z) = P_d^{(1)}(z)\Gamma_I^{(1)}(z) - P_d^{(2)}(z)\Gamma_I^{(2)}(z); \quad A_d^{(2)}(z) = P_d^{(2)}(z)\Gamma_I^{(4)}(z) - P_d^{(1)}(z)\Gamma_I^{(3)}(z).$$

Finally, the Wiener-Hopf formula gives

$$\mathbb{E}_{it}\left[\Delta b_{it+1}\right] = \left[\psi_b\left(L\right)\right]_+ \Gamma^{-1}(L) X_{it},$$

where the z-transform of the operator ψ_b is given by

$$\psi_b\left(z\right) = \left[\psi_b^{(1)}(z),\;\; \psi_b^{(2)}(z)\right] = z^{-1}(z-1) S_{bx}(z) \left(\Gamma^{-1}(z^{-1})\right)^{\intercal},$$

and $\left[\psi_b^{(1)}(z)\right]_+ = \psi_b^{(1)}(z) - P_b^{(1)}(z)$, $\left[\psi_b^{(2)}(z)\right]_+ = \psi_b^{(2)}(z) - P_b^{(2)}(z)$. Here $P_b^{(1)}(z)$ and $P_b^{(2)}(z)$ denote the negative powers of z in the Laurent series expansions of $\psi_b^{(1)}(z)$ and $\psi_b^{(2)}(z)$, respectively. It follows that

$$\left[\psi_b^{(1)}(z),\;\; \psi_b^{(2)}(z)\right]_+ \Gamma^{-1}(z) = \left[G_b^{(1)}(z) - A_b^{(1)}(z),\;\; \frac{1-\pi_2(z)}{\pi_1(z)}\left(G_b^{(2)}(z) - A_b^{(2)}(z)\right)\right],$$

where

$$G_b^{(1)}(z) = \psi_b^{(1)}(z)\Gamma_I^{(1)}(z) - \psi_b^{(2)}\Gamma_I^{(2)}(z); \quad G_b^{(2)}(z) = \psi_b^{(2)}(z)\Gamma_I^{(4)}(z) - \psi_b^{(1)}(z)\Gamma_I^{(3)}(z),$$

and

$$A_b^{(1)}(z) = P_b^{(1)}(z)\Gamma_I^{(1)}(z) - P_b^{(2)}(z)\Gamma_I^{(2)}(z); \quad A_b^{(2)}(z) = P_b^{(2)}(z)\Gamma_I^{(4)}(z) - P_b^{(1)}(z)\Gamma_I^{(3)}(z).$$

As in Step 2 we can also derive that

$$\begin{aligned}
\psi_b^{(1)}(z) &= z^{-1}(z-1)\left[M_b^a(z)A_n^{(1)}(z)\sigma_a^2 - M_b^u(z)A_n^{(2)}(z)\sigma_u^2 + \Gamma_I^{(1)}(z^{-1})M_b^i(z)\sigma_i^2\right], \\
\psi_b^{(2)}(z) &= z^{-1}(z-1)\left[-M_b^a(z)A_n^{(3)}(z)\sigma_a^2 + M_b^u(z)A_n^{(4)}(z)\sigma_u^2 - \Gamma_I^{(2)}(z^{-1})M_b^i(z)\sigma_i^2\right].
\end{aligned}$$

Step 4. Derive the equilibrium system for $\pi_1(z)$ and $\pi_2(z)$. By Step 3 we obtain an expression for χ_{it}. Matching coefficients of $X_{it} = [a_{it}, q_t]^\intercal$ with those in (40), we obtain the following equilibrium conditions for $\pi_1(z)$ and $\pi_2(z)$:

$$\pi_1(z) = \frac{(1-\lambda_s)}{z(1-\lambda_s z)}\left[\Gamma_I^{(1)}(z)\Gamma_I^{(1)}(z^{-1}) + \Gamma_I^{(2)}(z)\Gamma_I^{(2)}(z^{-1})\right]\sigma_i^2\pi_1(z) - A_s^{(1)}(z) + \frac{R^{(1)}(z)}{z(1-\lambda_s z)}, \tag{S1.11}$$

and

$$\begin{aligned}
\pi_2(z) \quad = \quad & \frac{1-\pi_2(z)}{z(1-\lambda_s z)\pi_1(z)}\Big\{(\lambda_s - 1)\left[\Gamma_I^{(1)}(z^{-1})\Gamma_I^{(3)}(z) + \Gamma_I^{(2)}(z^{-1})\Gamma_I^{(4)}(z)\right]\sigma_i^2\pi_1(z) \\
& -z(1-\lambda_s z)A_s^{(2)}(z) + R^{(2)}(z)\Big\} + z^{-1}\beta,
\end{aligned} \tag{S1.12}$$

where $R^{(1)}(z)$ and $R^{(2)}(z)$ are defined as

$$\begin{aligned}
R^{(1)}(z) = &\Bigg\{ -\beta\frac{1}{\sigma_w}\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}z^{-1}\left(V_{12}\Gamma_I^{(1)}(z) + \frac{1}{\lambda_w}V_{22}\Gamma_I^{(2)}(z)\right) \\
& + (1-\beta)\left[G_d^{(1)}(z) - A_d^{(1)}(z)\right] + \left[G_b^{(1)}(z) - A_b^{(1)}(z)\right]\Bigg\} z(1-\lambda_s z)
\end{aligned}$$

and

$$\begin{aligned}
R^{(2)}(z) = &\Bigg\{ \beta\frac{1}{\sigma_w}\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}z^{-1}\left(V_{12}\Gamma_I^{(3)}(z) + \frac{1}{\lambda_w}V_{22}\Gamma_I^{(4)}(z)\right) \\
& + (1-\beta)\left[G_d^{(1)}(z) - A_d^{(1)}(z)\right] + \left[G_b^{(2)}(z) - A_b^{(2)}(z)\right]\Bigg\} z(1-\lambda_s z).
\end{aligned}$$

Define an operator $\mathcal{T}$ that maps the vector of functions $[\pi_1(z), \pi_2(z)]$ to the vector of functions that are equal to the expressions on the right-hand sides of equations (S1.11) and (S1.12). Since the signal system contains endogenous prices, many variables in these expressions depend on $[\pi_1(z), \pi_2(z)]$ in a complicated way. Thus the operator $\mathcal{T}$ is nonlinear in general. The equilibrium functions $\pi_1(z)$ and $\pi_2(z)$ correspond to the fixed point of $\mathcal{T}$ in $\mathbf{H}^2(\mathbb{D})$. Moreover, we use (S1.12) to derive that

$$\begin{aligned}\frac{\pi_1(z)}{1-\pi_2(z)} \quad = \quad & \frac{1}{(1-\lambda_s z)(z-\beta)}\left\{-z(1-\lambda_s z)A_s^{(2)}(z)+R^{(2)}(z)\right. \\ & \left.+\left[z(1-\lambda_s z)-(1-\lambda_s)\left(\Gamma_I^{(1)}(z^{-1})\Gamma_I^{(3)}(z)+\Gamma_I^{(2)}(z^{-1})\Gamma_I^{(4)}(z)\right)\sigma_i^2\right]\pi_1(z)\right\}.\end{aligned} \tag{S1.13}$$

We also have to ensure that $\frac{\pi_1(z)}{1-\pi_2(z)} \in \mathbf{H}^2(\mathbb{D})$ in equilibrium. Note that our triangular spectral factorization method also sheds light on the rationale behind the non-linearity and the non-rational representation of the equilibrium. Specifically, the non–linearity arises from the first–step of the spectral factorization in which a new function $\tilde{\pi}_1(z)$ is created and the integrity of the original function $\frac{\pi_1(L)}{\left(1-\pi_2(L)\right)}$ cannot be preserved. By comparison, Assumption 2 in Section 5 leads to a spectral factorization with no additional endogenous function. It also preserves the integrity of the original functions $\frac{\pi_1(L)}{\left(1-\pi_2(L)\right)}$ as a whole. Similar argument also applies to Kasa, Walker, and Whiteman (2014), as their signal system is square so that factorization does not need the first step, avoiding the complication.

S1.2 Numerical Methods

The equilibrium is characterized by the fixed point of the operator $\mathcal{T}$. Due to the endogeneity of the price signal, this operator is nonlinear and thus the model does not admit a solution in the form of rational functions. We now approximate the true model solution, which is in the form of MA(∞), by finite-order ARMA(p,q) processes in the time domain or by rational functions in the frequency domain. Rational functions also allow us to evaluate the annihilation operator tractably using the lemma in Appendix A of Hansen and Sargent (1980). The numerical method involves the following steps.

Step 1. We begin by an initial guess for $\pi_1(z)$ in the form of an irreducible rational function:

$$\pi_1(z) = \sigma_\pi \frac{\prod_{i=1}^{q}(1+\theta_i z)}{\prod_{j=1}^{p}\left(1-\rho_j z\right)}, \tag{S1.14}$$

where p and q are the orders of the ARMA representation and σ_π, θ_i, and $\left|\rho_j\right| < 1$ are constants. Given the initial guess, we solve for the canonical factorization equation

(S1.2) to obtain

$$\tilde{\pi}_1(z) = \sigma_{\tilde{\pi}} \frac{\prod_{i=1}^{m+1} \left(1 + \hat{\theta}_i z\right)}{(1-\rho_a z)(1-\rho_u z)\prod_{j=1}^{p}\left(1-\rho_j z\right)}, \tag{S1.15}$$

where $m = \max(p, q)$ and $\sigma_{\tilde{\pi}}$ and $\hat{\theta}_i$ are determined by the factorization:

$$\begin{aligned}\sigma_{\tilde{\pi}}^2 \prod_{i=1}^{m+1} \left(1+\hat{\theta}_i z\right)\left(1+\hat{\theta}_i z^{-1}\right) =& \sigma_a^2\sigma_i^2\sigma_\pi^2 \prod_{i=1}^{q} (1+\theta_i z)\left(1+\theta_i z^{-1}\right)(1-\rho_u z)(1-\rho_u z^{-1}) \\ &+ \sigma_u^2\sigma_w^2 (1-\lambda_w z)(1-\lambda_w z^{-1}) \prod_{j=1}^{p}\left(1-\rho_j z\right)\left(1-\rho_j z^{-1}\right).\end{aligned} \tag{S1.16}$$

In particular, set $\left|\hat{\theta}_i\right| < 1, \ \forall i = 1, 2, ...m+1$.

In addition, we take an initial guess for the constant $\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}$.

Step 2. Solve for the decision rules for quantities on the real side of the economy. We use (S1.7) to derive $M_y^a(z)$ and $M_y^u(z)$. We need to compute $P_y^{(1)}(z)$ and $P_y^{(2)}(z)$ by using the lemma in Hansen and Sargent (1980). Given the guess for $\pi_1(z)$ in (S1.14), (S1.15), and the expressions for $\psi_y^{(1)}(z)$ and $\psi_y^{(2)}(z)$ derived in Step 2 of Section S1.1, we deduce that $-\hat{\theta}_1, ...,$ and $-\hat{\theta}_{m+1}$ are the poles of $\psi_y^{(1)}(z)$ and $\psi_y^{(2)}(z)$ that are inside the unit disk. Thus we have

$$\begin{aligned} P_y^{(1)}(z) &= \sum_{k=1}^{m+1} \frac{\psi_{k,y}}{z+\hat{\theta}_k}, \ P_y^{(2)}(z) = \sum_{k=1}^{m+1} f_k \frac{\psi_{k,y}}{z+\hat{\theta}_k}, \\ f_k &\equiv \frac{V_{11}}{V_{12}} \frac{1+\lambda_w\hat{\theta}_k}{\hat{\theta}_k+\lambda_w}, \ k = 1, 2, ...m+1, \end{aligned}$$

where each $\psi_{k,y}$ is a constant defined as

$$\psi_{k,y} = \lim_{z\to-\hat{\theta}_k} (z+\hat{\theta}_k)\left[M_y^a(z)\sigma_a^2 A_n^{(1)}(z) - M_y^u(z)\sigma_u^2 A_n^{(2)}(z)\right],$$

provided that all poles $\left\{-\hat{\theta}_k\right\}_{k=1}^{m+1}$ inside the unit disk are distinct. No constant $\psi_{k,y}$ can be solved numerically using the preceding formula because $M_y^a(z)$ and $M_y^u(z)$ are unknown functions to be determined. We will use the method below to determine all $\psi_{k,y}$.

Plugging the guess for $\pi_1(z)$ and the expressions above for $P_y^{(1)}(z)$ and $P_y^{(2)}(z)$ (taking all unknown constant $\psi_{k,y}$ as given) into (S1.7), we obtain the following linear system:

$$\begin{bmatrix} Q_1(z) & Q_2(z) \\ \tilde{Q}_3(z) & \tilde{Q}_4(z) \end{bmatrix} \begin{bmatrix} M_y^a(z) \\ M_y^u(z) \end{bmatrix} = \begin{bmatrix} \frac{1}{\xi} - A_y^{(1)}(z)\theta - A_y^{(2)}(z)\theta \\ -\prod_{j=1}^{p}\left(1-\rho_j z\right) A_y^{(2)}(z)\theta \end{bmatrix} \equiv \begin{bmatrix} C_y^{(1)}(z) \\ C_y^{(2)}(z) \end{bmatrix}, \qquad \text{(S1.17)}$$

where

$$\tilde{Q}_3(z) = \theta\sigma_a^2 \prod_{j=1}^{p}\left(1-\rho_j z\right) H_d(z),$$

$$\tilde{Q}_4(z) = \prod_{i=1}^{q}\left(1+\theta_i z\right)\left(1-\rho_u z\right)\sigma_\pi - \theta\sigma_u^2 \prod_{j=1}^{p}\left(1-\rho_j z\right) H_c(z).$$

Solving this linear system yields

$$\begin{bmatrix} M_y^a(z) \\ M_y^u(z) \end{bmatrix} = \frac{1}{D_1^2(z)\left[\tilde{Q}_4(z)Q_1(z) - Q_2(z)\tilde{Q}_3(z)\right]} \begin{bmatrix} D_1(z)\tilde{Q}_4(z)C_y^{(1)}(z) - D_1(z)Q_2(z)C_y^{(2)}(z) \\ -D_1(z)\tilde{Q}_3(z)C_y^{(1)}(z) + D_1(z)Q_1(z)C_y^{(2)}(z) \end{bmatrix},$$

where we define

$$D_1(z) = \prod_{i=1}^{m+1}\left(1+\hat{\theta}_i z\right)\left(z+\hat{\theta}_i\right).$$

We can verify that the above solutions for $M_y^a(z)$ and $M_y^u(z)$ are irreducible rational functions. That is, the numerator and denominator are pure polynomial functions.

The denominator function $D_y(z) \equiv D_1^2(z)\left[\tilde{Q}_4(z)Q_1(z) - Q_2(z)\tilde{Q}_3(z)\right]$ determines the existence and uniqueness of a stationary equilibrium. The necessary condition for the existence requires that $D_y(z)$ has precisely $m+1$ roots inside the open unit disk. We verify this condition in every iteration in our numerical computations. Let $\{z_j\}_{j=1}^{m+1}$ denote all the inside roots of $D_y(z)$. To pin down the vector of constants $\psi_y = \left[\psi_{1,y}, ..., \psi_{m+1,y}\right]^\intercal$, we use the following system of $m+1$ equations:

$$D_1(z_j)\tilde{Q}_4(z_j)C_y^{(1)}(z_j) - D_1(z_j)Q_2(z_j)C_y^{(2)}(z_j) = 0, \; j = 1, 2, ...m+1,$$

which gives a linear system for ψ_y :

$$A^c\psi_y = C^c,$$

where A^c is an $(m+1)\times(m+1)$ matrix of constants, and C^c is an $(m+1)$ dimensional vector of constants. We derive this system by substituting $P_y^{(1)}(z)$ and $P_y^{(2)}(z)$ (which depend on ψ_y) into $A_y^{(i)}(z)$ and $C_y^{(i)}(z)$, $i = 1, 2$. For simplicity, we omit the detailed algebra here. The idea is that the solution for ψ_y must remove the poles of $D_y(z)$ inside the open unit disk so that the solutions for $M_y^a(z)$ and $M_y^u(z)$ are analytic inside the open unit disk. If the matrix A^c is invertible, the solution is unique. We verify this condition in every iteration of our numerical computations. Given the solutions for $M_y^a(z)$ and $M_y^u(z)$, we solve for $M_y^i(z)$ using (S1.5). We can also solve for $M_b(z)$, $M_n(z)$, and $M_d(z)$ using the formulas derived in Step 2 of Section S1.1.

Step 3. We compute all annihilated functions of negative powers of z on the financial side of the model using the Hansen-Sargent lemma. Let $\{z_k\}_{k=1}^{m+2} = \left\{0, -\hat{\theta}_1, ..., -\hat{\theta}_{m+1}\right\}$ denote the set of poles inside the unit disk. Provided that all poles are distinct, we have

$$P_s^{(1)}(z) = \sum_{k=1}^{m+2} \frac{\psi_{k,s}^{(1)}}{z - z_k},\ P_s^{(2)}(z) = -\sum_{k=1}^{m+2} \frac{\psi_{k,s}^{(2)}}{z - z_k},$$
$$P_d^{(1)}(z) = \sum_{k=1}^{m+2} \frac{\psi_{k,d}^{(1)}}{z - z_k},\ P_d^{(2)}(z) = \sum_{k=1}^{m+2} \frac{\psi_{k,d}^{(2)}}{z - z_k},$$
$$P_b^{(1)}(z) = \sum_{k=1}^{m+2} \frac{\psi_{k,b}^{(1)}}{z - z_k},\ P_b^{(2)}(z) = \sum_{k=1}^{m+2} \frac{\psi_{k,b}^{(2)}}{z - z_k},$$

where the constants are given by

$$\psi_{k,s}^{(1)} = \lim_{z \to z_k} (z - z_k) \left[z^{-1}\alpha_3 M_s^i(z)\Gamma_I^{(1)}(z^{-1})\right] \sigma_i^2; \quad \psi_{k,s}^{(2)} = \lim_{z \to z_k} (z - z_k) \left[z^{-1}\alpha_3 M_s^i(z)\Gamma_I^{(2)}(z^{-1})\right] \sigma_i^2.$$

and

$$\psi_{k,d}^{(1)} = \lim_{z \to z_k} (z - z_k) z^{-1} \left[M_d^a(z) A_n^{(1)}(z)\sigma_a^2 - M_d^u(z) A_n^{(2)}(z)\sigma_u^2\right],$$
$$\psi_{k,d}^{(2)} = \lim_{z \to z_k} (z - z_k) z^{-1} \left[M_d^u(z) A_n^{(4)}(z)\sigma_u^2 - M_d^a(z) A_n^{(3)}(z)\sigma_a^2\right].$$

and

$$\psi_{k,b}^{(1)} = \lim_{z \to z_k} (z - z_k)(z-1) z^{-1} \left[M_b^a(z) A_n^{(1)}(z)\sigma_a^2 - M_b^u(z) A_n^{(2)}(z)\sigma_u^2 + M_b^i(z)\Gamma_I^{(1)}(z^{-1})\sigma_i^2\right],$$
$$\psi_{k,b}^{(2)} = \lim_{z \to z_k} (z - z_k)(z-1) z^{-1} \left[M_b^u(z) A_n^{(4)}(z)\sigma_u^2 - M_b^a(z) A_n^{(3)}(z)\sigma_a^2 - M_b^i(z)\Gamma_I^{(2)}(z^{-1})\sigma_i^2\right].$$

Given the guess of $\pi_1(z)$ in (S1.14) and the solutions for $M_y(z), M_d(z), M_n(z)$, and $M_b(z)$ in the previous step, we can compute the constants $\psi_{k,d}^{(1)}$, $\psi_{k,d}^{(2)}$, $\psi_{k,b}^{(1)}$, and $\psi_{k,b}^{(2)}$ for $k = 1, 2, ..., m+2$. The other constants $\psi_{k,s}^{(1)}$ and $\psi_{k,s}^{(2)}$ will be solved in the next step. We cannot use the formulas above to determine $\psi_{k,s}^{(1)}$ and $\psi_{k,s}^{(2)}$ because $M_s^i(z)$ is unknown function to be determined in equilibrium. We can verify that

$$\psi^{(2)}_{k,s} = h_k \psi^{(1)}_{k,s},$$

$$h_k = \begin{cases} \frac{V_{12}}{V_{22}} \frac{1-\lambda_w z_k}{z_k - \lambda_w}, & \text{if } z_k = 0, \\ \\ \frac{V_{11}}{V_{12}} \frac{1-\lambda_w z_k}{z_k - \lambda_w}, & else. \end{cases}$$

Thus we only need to solve for $\psi^{(1)}_{k,s}$, $k = 1, ..., m+2$.

Step 4. Solve for the update of $\pi_1(z)$ and $\pi_2(z)$ using equations (S1.11) and (S1.12). Given the guess for $\pi_1(z)$ in (S1.14), we can verify that $R^{(1)}(z)$ is an analytic rational function. Let $R^{(1)}_D(z)$ denote the denominator polynomial function of $R^{(1)}(z)$ in its irreducible form. Since $R^{(1)}(z)$ is analytic, $R^{(1)}_D(z) \neq 0$ inside the open unit disk. We can write

$$R^{(1)}_D(z) = R^{(1)}_D(0) \prod_{i=1}^{g} (1 + z_i z),$$

where g denotes the degree of $R^{(1)}_D(z)$ and $-z_i^{-1}, ..., -z_g^{-1}$ are the g roots of $R^{(1)}_D(z)$ that are outside the open unit disk. Using the definition of the unitary matrix V, we can show that the denominator of the rational function $z(1-\lambda_s z) - (1-\lambda_s)\sigma_i^2 \Big[\Gamma^{(1)}_I(z)\Gamma^{(1)}_I(z^{-1})$ $+ \Gamma^{(2)}_I(z)\Gamma^{(2)}_I(z^{-1})\Big]$ in the irreducible form is given by

$$D_1(z) = \prod_{k=1}^{m+1} \left(1 + \hat{\theta}_k z\right)\left(z + \hat{\theta}_k\right).$$

Notice that some factors in $D_1(z)$ and $R^{(1)}_D(z)$ may be identical. We define $D_2(z)$ as their least common multiple.

We now rewrite (S1.11) as

$$\pi_1(z) = \frac{D_2(z)\left[R^{(1)}(z) - z(1-\lambda_s z)A^{(1)}_s(z)\right]}{D_2(z)\Big[z(1-\lambda_s z) - (1-\lambda_s)\sigma_i^2\left[\Gamma^{(1)}_I(z)\Gamma^{(1)}_I(z^{-1}) + \Gamma^{(2)}_I(z)\Gamma^{(2)}_I(z^{-1})\right]\Big]}, \tag{S1.18}$$

where both the numerator and the denominator are pure polynomial functions. Let $\pi^D_1(z)$ denote the denominator function. The existence and uniqueness of a stationary equilibrium solution for $\pi_1(z)$ is determined by the roots of $\pi^D_1(z)$. More specifically, to determine the $m+2$ dimensional vector of unknown constants $\psi_s = \left[\psi^{(1)}_{1,s}, ..., \psi^{(1)}_{m+2,s}\right]^\intercal$, we need $\pi^D_1(z)$ to have precisely $m+2$ distinct roots inside the open unit disk. We verify this condition in every iteration of the numerical computation. Without risk of confusion, let $\{\hat{z}_k\}_{k=1}^{m+2}$ denote the set of distinct roots of $\pi_1^D(z)$ that are inside the open unit disk.

We then pin down ψ_s by removing the poles $\{\hat{z}_k\}_{k=1}^{m+2}$ and evaluating the numerator polynomial

$$D_2(\hat{z}_k)\left[R^{(1)}(\hat{z}_k) - \hat{z}_k(1-\lambda_s\hat{z}_k)A_s^{(1)}(\hat{z}_k)\right] = 0, \quad \forall k = 1, 2, ...m+2,$$

which leads to the linear system

$$A^\pi \psi_s = C^\pi,$$

where we have used the definition of $A_s^{(1)}(z)$ and the expression of $P_s^{(1)}(z)$ derived in Step 2. We deduce that A^π is an $(m+2)\times(m+2)$ matrix with elements given by

$$A^\pi(k,i) = \frac{\Gamma_I^{(1)}(\hat{z}_k)D_2(\hat{z}_k)}{\hat{z}_k - z_i} + \frac{\Gamma_I^{(2)}(\hat{z}_k)D_2(\hat{z}_k)}{\hat{z}_k - z_i}h_i,$$

for $k = 1, 2, ...m+2$ and $i = 1, 2, ...m+2$, and $z_i \in \left\{0, -\hat{\theta}_1, ..., -\hat{\theta}_{m+1}\right\}$. The k^{th} element of $(m+2)\times 1$ vector C^π is given by

$$C^\pi(k) = R^{(1)}(\hat{z}_k)D_2(\hat{z}_k), \quad \forall k = 1, 2, ..., m+2.$$

If A^π is full rank, the solution is indeed unique. Again, we verify this condition in every iteration.

Once determining ψ_s, we update the guess for $\pi_1(z)$ using the solution in (S1.18). Given this solution for $\pi_1(z)$, we use (S1.13) to solve for $\frac{\pi_1(z)}{1-\pi_2(z)}$. Observe that the numerator on the right-hand side of (S1.13) is analytic inside the open unit disk, but we still need to remove the pole at $z = \beta$. We set the constant $\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}$ to remove this pole. That is,

$$\phi(\beta)\pi_1(\beta) - \beta(1-\lambda_s\beta)A_s^{(2)}(\beta) + R^{(2)}(\beta) = 0,$$

where

$$\phi(z) = z(1-\lambda_s z) - (1-\lambda_s)\left[\Gamma_I^{(1)}(z^{-1})\Gamma_I^{(3)}(z) + \Gamma_I^{(2)}(z^{-1})\Gamma_I^{(4)}(z)\right]\sigma_i^2.$$

This leads to the following solution for the constant

$$\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)} = \frac{\sigma_w}{\beta(1-\lambda_s\beta)\left[V_{12}\Gamma_I^{(3)}(\beta) + \frac{1}{\lambda_w}V_{22}\Gamma_I^{(4)}(\beta)\right]}\times$$

$$\left\{\beta(1-\lambda_s\beta)\left(A_s^{(2)}(\beta)-(1-\beta)\left[G_d^{(1)}(\beta)-A_d^{(1)}(\beta)\right]-\left[G_b^{(2)}(\beta)-A_b^{(2)}(\beta)\right]\right)-\phi(\beta)\pi_1(\beta)\right\}. \tag{S1.19}$$

We use this solution to update the initial guess for $\frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}$. Finally, we iterate until convergence.

In summary, we employ the following iterative algorithm to solve the model.

Algorithm 1 Numerical Approximation of Equilibrium

Step 0. Begin with a guess for p, q, σ_π, $\pi_c \equiv \frac{\tilde{\pi}_1(0)}{1-\pi_2(0)}$, $\{\theta_i\}_{i=1}^q$, $\{\rho_j\}_{j=1}^p$ with $|\rho_j| < 1, \forall j$.

Step 1: Set $m = \max\{p,q\}$ and compute $\sigma_{\tilde{\pi}}$ and $\{\hat{\theta_i}\}_{i=1}^m$ using (S1.16).

Step 2: Solve for the functions $M_y(z)$, $M_d(z)$, $M_b(z)$, and $M_n(z)$.

Step 3: Let $\pi_1^A(z)$ and π_c^+ be the expressions on the right-hand sides of (S1.18) and (S1.19), respectively.

Step 4: Update the initial guess using

$$\pi_1^+(z) = \sigma_\pi^+ \frac{\prod_{i=1}^q \left(1+\theta_i^+ z\right)}{\prod_{j=1}^p \left(1-\rho_j^+ z\right)},$$

where $\sigma_\pi^+, \theta_i^+, \rho_j^+$ are the solution to the problem

$$\min_{\sigma_\pi,\theta_i,\rho_j} \sum_{n=1}^N \left|\pi_1^+(n) - \pi_1^A(n)\right|^2,$$

where $\pi_1^+(n)$ and $\pi_1^A(n)$ are the coefficients of the moving average expansion of $\pi_1^+(z)$ and $\pi_1^A(z)$, with $N = 70$.

Step 5: Iterate Steps 0–4 until $\max\left\{\left|\rho_j^+ - \rho_j\right|, \left|\theta_i^+ - \theta_i\right|, \left|\sigma_\pi^+ - \sigma_\pi\right|\right\} < 10^{-3}$; $\forall i, j$

Step 6: Compute $\epsilon = \max\left\{\left|\left|\pi_1^+(z) - \pi_1^A(z)\right|\right|_{\mathbf{H}^2}, \left|\pi_c^+ - \pi_c\right|\right\}$; if $\epsilon < 10^{-5}$, stop; otherwise, set $p := p+1$, $q := q+1$ and repeat Steps 0-5.

S1.3 Macro-Financial Disconnection

In the extension of Section 6.2, information is segregated between groups as agents in one group receive no signal about the other group's shocks. Let $\mathcal{I}_{it}^p$ and $\mathcal{I}_t^n$ denote the information set for agents in participating island i and any non-participating island, respectively. Then conditional expectations of the other group's shocks are equal to their unconditional mean, i.e.,

$$\begin{aligned} &\mathbb{E}_j\left[F(L)\epsilon_{at}^p \mid \mathcal{I}_t^n\right] = 0; \quad \mathbb{E}_j\left[F(L)\epsilon_{it} \mid \mathcal{I}_t^n\right] = 0; \quad \forall i \in I_p,\ j \in I_n \\ &\mathbb{E}_i\left[F(L)\epsilon_{at}^n \mid \mathcal{I}_{it}^p\right] = 0; \quad \mathbb{E}_i\left[F(L)\epsilon_{jt} \mid \mathcal{I}_{it}^p\right] = 0; \quad \forall i \in I_p,\ j \in I_n \end{aligned} \tag{S1.20}$$

for any square–summable lag polynomial $F(L)$. Then we can use (32) to characterize the equity market equilibrium[18]

$$q_t = \int_{i \in I_p} \mathbb{E}_i \left[\alpha_3 s_{it+2}^h + \Delta b_{it+1} \mid \mathcal{I}_{it}^p \right] di + \int_{i \in I_p} \mathbb{E}_i \left[\beta q_{t+1} + (1-\beta) d_{t+1} \mid \mathcal{I}_{it}^p \right] di + u_t. \tag{S1.21}$$

Given property (S1.20) and our information structure, (S1.21) implies that we can focus on the equilibrium in which the equity price are driven by

$$q_t = M_q^p(L)\epsilon_{at}^p + M_q^u(L)\epsilon_{ut}$$

which resembles (41). Intuitively, the stock price does not respond to fluctuations of non-participants' TFP shocks. Moreover, the information structure and the dynamic interactions between shareholding choices s_{it}^h and q_t, (43) and (44), remain the same as in the basic model. Therefore, the unit root result in the equity price volatility is still valid, although the quantitative outcome depends on the participation measure κ and the modified real equilibrium.

Next, we characterize the log-linearized equilibrium in the real economy,

$$y_{it} = \frac{1}{\xi}\left(a_t^p + \epsilon_{it}\right) + \theta \mathbb{E}_i \left[\kappa y_t^p + (1-\kappa) y_t^n \mid \mathcal{I}_{it}^p \right]; \;\; \forall i \in I_p \tag{S1.22}$$

$$y_{jt} = \frac{1}{\xi}\left(a_t^n + \epsilon_{jt}\right) + \theta \mathbb{E}_j \left[\kappa y_t^p + (1-\kappa) y_t^n \mid \mathcal{I}_t^n \right]; \;\; \forall j \in I_n \tag{S1.23}$$

where $y_t = \kappa y_t^p + (1-\kappa) y_t^n$, and $y_t^p = \frac{1}{\kappa}\int_{i \in I_p} y_{it} di$, $y_t^n = \frac{1}{1-\kappa}\int_{j \in I_n} y_{jt} dj$ are log-linearized group aggregates. We conjecture the "segregated" equilibrium decision rules follow

$$\begin{aligned} y_{it} &= M_y^p(L)\epsilon_{at}^p + M_y^{i,p}(L)\epsilon_{it} + M_y^u(L)\epsilon_{ut}; \;\; \forall i \in I_p \\ y_{jt} &= M_y^n(L)\epsilon_{at}^n + M_y^{j,n}(L)\epsilon_{jt}; \;\; \forall j \in I_n \end{aligned}$$

We then use (S1.20) and the fact that $\mathbb{E}_j \left[\epsilon_{ut} \mid \mathcal{I}_t^n \right] = 0$ to get,

$$y_{it} = \frac{1}{\xi}\left(a_t^p + \epsilon_{it}\right) + \theta \kappa \mathbb{E}_i \left[y_t^p \mid \mathcal{I}_{it}^p \right]; \;\; \forall i \in I_p \tag{S1.24}$$

$$y_{jt} = \frac{1}{\xi}\left(a_t^n + \epsilon_{jt}\right) + \theta (1-\kappa) y_t^n; \;\; \forall j \in I_n \tag{S1.25}$$

Note that (S1.24) resembles (25), which leads to the decision rule for total output on participating islands,

$$y_t^p = \frac{1}{\kappa}\int_{i \in I_p} y_{it} di = M_y^p(L)\epsilon_{at}^p + M_y^u(L)\epsilon_{ut}$$

[18]The log-linearized coefficients determined by the steady state will be different from the basic model. In particular, the production side remains the same, while a re-distriburion of consumption occurs between participating and non-participating islands.

Meanwhile, aggregating (S1.25) produces a simple solution for y_t^n,

$$y_t^n = \frac{1}{1-\kappa}\int_{j\in I_n} M_y^n(L)\epsilon_{at}^n dj = \frac{1}{\xi\big[1-\theta(1-\kappa)\big](1-\rho_a L)}\epsilon_{at}^n$$

It is then more transparent to write the equilibrium aggregate output as

$$y_t = \kappa\left[M_y^p(L)\epsilon_{at}^p + M_y^u(L)\epsilon_{ut}\right] + (1-\kappa)\frac{1}{\xi\big[1-\theta(1-\kappa)\big](1-\rho_a L)}\epsilon_{at}^n$$

Since $M_y^p(L)$ and $M_y^u(L)$ are determined by (S1.24), which is almost equivalent to the equilibrium condition in the basic model except for the appearance of κ parameter, the solution for $M_y^p(L)$ and $M_y^u(L)$ is invariant up to the changes in κ and the steady-state coefficients.

S2 Frequency Domain Methods

In this section we introduce some mathematical background for the frequency domain methods. We study casual covariance stationary real-valued equilibrium processes that have an MA(∞) representation. For example, the aggregate output process in the model of Section 3 can be written as

$$y_t = \sum_{j=0}^{\infty} M_j \varepsilon_{a,t-j}, \tag{S2.1}$$

where $\{M_j\}_{j=0}^{\infty}$ is square summable, i.e., $\sum_{j=0}^{\infty}|M_j|^2 < \infty$. Solving for the infinite sequence of $\{M_j\}_{j=0}^{\infty}$ is a daunting task. The idea of the frequency domain method is to transform this problem into an equivalent problem of solving for an analytical function in the Hardy space. To define this space, we recall that $\mathbb{C}$ denotes the complex plan, $\mathbb{T}$ denotes the unit circle, and $\mathbb{D}$ denotes the open unit disk.

Definition S1 *The Hardy space $\mathbf{H}^2(\mathbb{D})$ is the class of analytical functions g in the unit disk $\mathbb{D}$ satisfying*

$$\left\{\frac{1}{2\pi}\sup_{0\le r<1}\int_{-\pi}^{\pi}\left|g\left(re^{i\omega}\right)\right|^2 d\omega\right\}^{1/2} < \infty.$$

It can be verified that the expression on the preceding inequality defines a norm on $\mathbf{H}^2(\mathbb{D})$, denoted as $\|g\|_{\mathbf{H}^2}$. The Hardy space can also be viewed as a certain closed vector subspace of the complex L^2 space for the unit circle $\mathbb{T}$. This connection is provided by the fact that the radial limit

$$\tilde{g}\left(e^{i\omega}\right) = \lim_{r\uparrow 1} g\left(re^{i\omega}\right)$$

exists for almost all $\omega \in [-\pi, \pi]$. The function $\tilde{g}$ belongs to the space $L^2(\mathbb{T})$ of functions $f : \mathbb{T} \to \mathbb{C}$ with the inner product

$$< f_1, f_2 >= \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(e^{i\omega}\right) \overline{f_2\left(e^{i\omega}\right)} d\omega, \quad f_1, f_2 \in L^2(\mathbb{T}).$$

Then we have

$$\|g\|_{\mathbf{H}^2} = \|\tilde{g}\|_{L^2} = \lim_{r\uparrow 1} \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left|g\left(re^{i\omega}\right)\right|^2 d\omega \right\}^{1/2} < \infty.$$

Denote by $\mathbf{H}^2(\mathbb{T})$ the vector subspace of $L^2(\mathbb{T})$ consisting of all limit functions $\tilde{g}$, when g varies in $\mathbf{H}^2(\mathbb{D})$.

Theorem S1 *(Katznelson 1976) $f \in \mathbf{H}^2(\mathbb{T})$ if and only if $f \in L^2(\mathbb{T})$ and $\hat{f}_n = 0$ for all $n < 0$, where $\hat{f}_n$ is the Fourier coefficient of a function f integrable on the unit circle,*

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f\left(e^{i\omega}\right) e^{-i\omega n} d\omega, \ n = 0, \pm 1, \pm 2, \ldots.$$

Suppose that $\tilde{g} \in \mathbf{H}^2(\mathbb{T})$ and $\tilde{g}$ has Fourier coefficients $\{a_n\}$ with $a_n = 0$ for all $n < 0$. We define

$$g(z) = \sum_{n=0}^{\infty} a_n z^n, \quad |z| < 1.$$

The following theorem ensures $g \in \mathbf{H}^2(\mathbb{D})$. Thus we have a bijection between $\mathbf{H}^2(\mathbb{D})$ and $\mathbf{H}^2(\mathbb{T})$.

Theorem S2 *If $f(z)$ is an analytic function in $\mathbb{D}$ and its Laurent expansion is*

$$f(z) = \sum_{n=0}^{\infty} b_n z^n,$$

then $f \in \mathbf{H}^2(\mathbb{D})$ if and only if $\{b_n\}_{n=0}^{\infty}$ is square summable, i.e., $\sum_{n=0}^{\infty} |b_n|^2 < \infty$. When this condition is satisfied

$$\sum_{n=0}^{\infty} |b_n|^2 = \|f\|_{\mathbf{H}^2}.$$

We call the map from the sequence $\{b_n\}_{n=0}^{\infty}$ to $f(z)$ a z-transform. Theorem S2 also allows us to give an equivalent definition of the Hardy space $\mathbf{H}^2(\mathbb{D})$ as the class of analytical functions $f : \mathbb{D} \to \mathbb{C}$, which are the z-transforms of some square summable sequences. Thus solving for $\{M_j\}_{j=0}^{\infty}$ in (S2.1) is equivalent to solving for a function $M(z)$ in the hardy space $\mathbf{H}^2(\mathbb{D})$. In particular, we can write $y_t = M(L)\epsilon_{at}$, where $M(z) \in \mathbf{H}^2(\mathbb{D})$ is the object we will solve for. We can use Theorem S2 to compute the variance of y_t easily because

$$Var(y_t) = \sigma_a^2 \sum_{j=0}^{\infty} M_j^2 = \sigma_a^2 \|M(z)\|_{\mathbf{H}^2}.$$

Finally, a rational function $f(z) \in \mathbf{H}^2(\mathbb{D})$ if and only if $f(z)$ is analytic in the closed unit disk. In particular, poles are not allowed on the unit circle.

S3 Computing Expectations in the Frequency Domain

We present our approach in a general framework. Suppose that the signal is an ℓ-dimensional variable X_t, defined in terms of infinite-order moving average processes.[19] Let $\mathbb{C}$ denote the complex plane, $\mathbb{T}$ denote the unit circle $\{z \in \mathbb{C} : |z| = 1\}$, and $\mathbb{D}$ denote the open unit disk $\{z \in \mathbb{C} : |z| < 1\}$.

Definition S2 *(signal representation) The $\ell-$dimensional real-valued signal process $\{X_t\}$ is linearly regular and admits representation*

$$\underset{\ell\times 1}{X_t} = \underset{\ell\times k}{H(L)}\, \underset{k\times 1}{\eta_t}, \quad \ell \le k,$$

where L denotes the lag operator, $\{\eta_t\}$ represents structural Gaussian innovations with mean zero and covariance matrix Σ_η, and $H(z)$ is an $\ell \times k$ matrix analytic function defined on the open unit disk $\mathbb{D}$ in the matrix-valued Hardy space $\mathbf{H}^2(\mathbb{D})$.[20]

We call $H(\cdot)$ the signal matrix or the transfer function as in the mathematics literature. To simplify the signal extraction problem, it is useful to assume a maximal rank condition for the signal process so that no redundant information is contained in X_t.

[19] We can extend the definition to contain information about future innovations (e.g. Bacchetta and Wincoop, 2008).

[20] See the Appendix S2 for the definition of the Hardy space. This definition can be easily extended to matrix cases, see Lindquist and Picci (2015), Appendix B.2

Assumption 4 *The $\ell-$dimensional signal process X_t has maximal rank, i.e. the rank of its associated spectral density $f_x(\omega)$ equals its dimension:*

$$rank\,(f_x(\omega)) = \ell$$

for almost all $\omega \in [-\pi, \pi]$.

An methodological contribution of our paper is that we study a non-square signal representation in that $\ell < k$. The existing literature focuses on the case of square signal representations with $\ell = k$ (e.g., Kasa, Walker, and Whiteman (2014), and Rondina and Walker (2015)). To use the Wiener-Hopf prediction formula, we need the Wold fundamental representation for the signal process. For the case of non-square signal representation, finding the Wold representation is non-trivial. We use spectral factorization techniques to solve this problem.

S3.1 A Two-Step Spectral Factorization Procedure

Our goal is to find a Wold representation for $\{X_t\}$. We are looking for an outer analytic matrix function $\Gamma(\cdot)$ in the Hardy space $\mathbf{H}^2(\mathbb{D})$ such that[21]

$$\underset{\ell\times 1}{X_t} = \underset{\ell\times\ell}{\Gamma(L)}\,\underset{\ell\times 1}{e_t}, \quad f_x(\omega) = \Gamma\left(e^{-i\omega}\right)\Gamma^*\left(e^{-i\omega}\right), \ \omega \in [-\pi, \pi], \tag{S3.1}$$

where asterisk denotes the conjugate transpose, $\{e_t\}$ is some mutually uncorrelated Wold (fundamental) innovation process with mean zero and an identity covariance matrix, f_x is the spectral density, and $\Gamma(\cdot)$ is an outer analytic function.[22]

For the square signal case with $\ell = k$, we can directly apply the Beurling-Blaschke factorization method to derive the Wold representation as in Kasa, Walker, and Whiteman (2014) and Rondina and Walker (2015). However, this method does not apply to the non-square case with $\ell < k$. We propose a two-step spectral factorization procedure. In step 1 we apply the convolution theorem to find the spectral density $f_x(\omega)$ of the signal process $\{X_t\}$. Then we use the Rozanov (1967) theorem to find a lower

[21] $\Gamma(z)$ is also called "canonical" or "fundamental" spectral factor. We refer readers to Lindquist and Picci (2015), Chapter 4 for characterizations of outer functions. One prominent feature of outer functions is that they cannot have zeros inside the unit disk. Note that Lindquist and Picci (2015) use the engineering definition of $z = e^{i\omega}$ so that the analytic region is reversed comparing with this paper, but all analytic results remain valid.

[22] Note that the Wold fundamental innovations can have non-diagonal, non-normalized covariance matrices. Using the unitary eigen-decomposition of the covariance matrix, we can obtain the orthonormal Wold representations with an identity covariance matrix.

triangular decomposition of $f_x(\omega)$. In step 2 we apply the Beurling-Blaschke factorization method to the lower triangular matrix. Due to the length constraints, we omit the algebraic derivations in this section. These details are contained in the Appendix S4.

Before describing the two-step procedure, we start with the following well-known result in Time Series, whose proof is omitted for brevity.

Lemma S1 *Suppose that X_t is the vector of signals defined in Definition S2 and Assumption 4 holds. Moreover, the transfer function $H(z)$ is a non-square matrix function with dimension $k > \ell$. Then the spectral density $f_x(\omega)$ is an $\ell \times \ell$ matrix function defined on $[-\pi, \pi]$ and*

$$f_x(\omega) = H\left(e^{-i\omega}\right) \Sigma_\eta H^*\left(e^{-i\omega}\right) = H(z) \Sigma_\eta H\left(z^{-1}\right)^\intercal, \quad z = e^{-i\omega},$$

where the superscript $\intercal$ denotes the transpose of a matrix. Furthermore, $f_x(\omega)$ is a Hermitian normal matrix that is non-negative definite for almost all $\omega \in [-\pi, \pi]$. If we extend the definition of z to the entire complex plane $\mathbb{C}$, then the autocovariance generating function is given by $S_x(z) = H(z) \Sigma_\eta H(z^{-1})^\intercal$, but without the Hermitian non-negativeness property for general $z \in \mathbb{C}$.

Lemma S1 allows us to transform the non-square signal transfer matrix function into the square spectral density matrix $f_x(\omega)$. Based on this lemma, the first step of the spectral factorization procedure is to decompose $f_x(\omega)$ into triangular matrix functions using Rozanov's (1967) analytical method.

Step 1. Given an $\ell \times \ell$ spectral density matrix $f_x(\omega)$ with full rank almost everywhere, construct an $\ell \times \ell$ lower triangular matrix function $\widetilde{\Gamma}(e^{-i\omega})$ such that

$$f_x(\omega) = \widetilde{\Gamma}\left(e^{-i\omega}\right) \widetilde{\Gamma}^*\left(e^{-i\omega}\right),$$

where

$$\widetilde{\Gamma}(z) = \begin{bmatrix} \widetilde{\Gamma}_{11}(z) & 0 & ... & 0 \\ \widetilde{\Gamma}_{21}(z) & \widetilde{\Gamma}_{22}(z) & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\Gamma}_{\ell 1}(z) & \widetilde{\Gamma}_{\ell 2}(z) & ... & \widetilde{\Gamma}_{\ell\ell}(z) \end{bmatrix}.$$

If $f_x(\omega)$ is rational, then all elements of the matrix function are rational and analytic in the closed unit disk $\mathbb{T} \cup \mathbb{D}$ and hence in the $\mathbf{H}^2(\mathbb{D})$ space. Moreover, $\widetilde{\Gamma}(e^{-i\omega})$ has full rank in $\mathbb{D}$ except for at most a finite number of points.

If the determinant of the analytic matrix $\widetilde{\Gamma}(z)$ vanishes at finitely many points inside the unit disk, it is not a Wold spectral factor. Without loss of generality, let $\{z_1, z_2, ...z_n\}$ be the finite set of distinct points such that $\det\left(\widetilde{\Gamma}(z_j)\right) = 0$, $|z_j| < 1$, $j \in \{1, 2, ...n\}$. Let $\overline{z}_j$ denote the conjugate of z_j. We assume that all zeros are of order 1 (this property is generic).

The second step of our spectral factorization method employs a multivariate version of the Beurling-Blaschke factorization theorem to remove any zeros inside the unit disk.

Step 2. The Wold spectral factor $\Gamma(z)$ can be obtained by the factorization for Hardy space functions as

$$\Gamma(z) = \widetilde{\Gamma}(z) \prod_{j=1}^{n} V_j^{-1} B_j(z),$$

where the $\ell \times \ell$ Blaschke matrices $B_j(z)$ are (inverse) inner matrix functions of the form

$$B_j(z) = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & \frac{1-\bar{z}_j z}{z-z_j} \end{bmatrix},$$

and the constant unitary matrix V_j is given by the singular value decomposition of $\widetilde{\Gamma}(z)$ evaluated at the zeros

$$\widetilde{\Gamma}(z_j) = U_j D V_j,$$

where D is a diagonal matrix containing the singular values.

The constant unitary matrices V_j remove the unwelcome poles brought in by the Blaschke factors. There are different ways of computing these matrices, and we use the eigen-decomposition method. In particular, the orthonormal column vectors of V_j can be directly picked from normalized linear independent eigenvectors of the Hermitian matrix $G_j(z_j) = \widetilde{\Gamma}^*(z_j)\widetilde{\Gamma}(z_j)$, which are automatically pairwise-orthogonal for distinct eigenvalues. For more complicated systems, the eigenvectors can be found easily using symbolic toolboxes in Matlab or Mathematica.

S3.2 Wiener-Hopf Prediction Formula

Using the Wold representation for the signal process, we can compute the conditional expectations given the history of signals. Since in our model agents need to perform optimal linear filtering to estimate unobserved shocks, we use the Wiener-Hopf prediction formula, a generalization of the Wiener-Kolmogorov forecasting formula.

Consider any random vector Θ_t satisfying $\Theta_t = G(L)\,\eta_t$, where $G(z)$ is a matrix analytic function in some matrix-valued Hardy space, we wish to compute the conditional expectation $\mathbb{E}\left[L^m\Theta_t\right|\{X_{t-n}\}_{n=0}^{\infty}]$ given the history of signals $\{X_{t-n}\}_{n=0}^{\infty}$, where m is any integer. The Wiener-Hopf prediction formula gives

$$\mathbb{E}\left[L^m\Theta_t\right|\{X_{t-n}\}_{n=0}^{\infty}] = \Xi(L)\,X_t, \tag{S3.2}$$

where the analytic matrix function $\Xi(z)$ is given by

$$\Xi(z) = \left[z^m S_{\Theta x}(z)\left(\Gamma^{-1}\left(z^{-1}\right)\right)^{\intercal}\right]_+ \Gamma^{-1}(z). \tag{S3.3}$$

Here $\Gamma(z)$ is the Wold spectral factor derived in the previous subsection and $S_{\Theta x}(z) = G(z)\,\Sigma_\eta H(1/z)^{\intercal}$ is the covariance generating function. The annihilation operator $[\cdot]_+$ is linear and is used to remove the principal part of the Laurent series expansion of the analytic functions around a common region of convergence.[23] This formula reduces to the Wiener-Kolmogorov formula when $\Theta_t = X_t$ so that $\Xi(z) = [z^m\Gamma(z)]_+\,\Gamma^{-1}(z)$. If the forecast objects follow geometrically discounted processes, the formula reduces to the Hansen-Sargent optimal prediction formula.

S4 Algebraic Derivation on Spectral Factorization in Appendix S3

Derivations in Step 1: Since $f_x(\omega)$ is rational, it has a constant, maximal rank of ℓ except at a finite number of points on the unit circle $\mathbb{T}$. To develop the triangular factorization of the spectral density, we need the following lemma from Rozanov (1967) on rational functions.

Lemma S2 *Every non-negative (real) rational function $f(\omega)$ of $e^{-i\omega}$ can be represented in the form*

$$f(\omega) = \frac{|P(e^{-i\omega})|^2}{|Q(e^{-i\omega})|^2} = \frac{P(e^{-i\omega})\overline{P(e^{-i\omega})}}{Q(e^{-i\omega})\overline{Q(e^{-i\omega})}} = \frac{P(z)\overline{P(z)}}{Q(z)\overline{Q(z)}}$$

for $z \in \mathbb{T}$. The polynomial functions $P(z)$ and $Q(z)$ have no zeros in the open unit disk. If f satisfies

$$f(\omega) = f(-\omega)$$

Then the coefficients of $P(z)$ and $Q(z)$ can be chosen all real.

[23] See Kailath, Sayed, and Hassibi (2000) for a textbook proof of the Wiener-Hopf prediction formula. Hansen and Sargent (1980) provide a practical method of computing the annihilation operator using elementary complex analysis.

Proof: See Rozanov (1967), Lemma 10.1. $\square$

If we extend $f(z)$ to be a complex function in the entire complex plane, the preceding lemma implies that it can be factorized in a "symmetric" way such that if λ_i is a root for $f(z)$, so is the conjugate inverse $1/\overline{\lambda_i}$.

Now consider the $\ell \times \ell$ spectral density matrix $f_x(\omega)$, by definition it is Hermitian, normal, and non-negative definite for almost all ω. For simplicity, we drop the x subscript and write the f matrix as

$$f(\omega) = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1\ell} \\ f_{21} & f_{22} & \cdots & f_{2\ell} \\ \vdots & \vdots & \ddots & \vdots \\ f_{\ell 1} & f_{\ell 2} & \cdots & f_{\ell\ell} \end{bmatrix}.$$

Using the Sylvester's criterion for the non-negative definite matrix, define the family of leading principal minors as $M_j(\omega)$, $j = 1, 2, ...\ell$. By definition, $M_j(\omega) \geq 0$ a.e., and $M_1(\omega) = f_{11} \geq 0$ a.e.

Next we implement elementary row operations on the matrix. Adding to the rth row ($r = 2, 3, ...\ell$) the first row, multiplied by $-\frac{f_{r1}}{f_{11}}$, yielding

$$f(\omega) = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1\ell} \\ 0 & f_{22} - f_{12}\frac{f_{21}}{f_{11}} & \cdots & f_{2\ell} - f_{1\ell}\frac{f_{21}}{f_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & f_{\ell 2} - f_{12}\frac{f_{\ell 1}}{f_{11}} & \cdots & f_{\ell\ell} - f_{1\ell}\frac{f_{\ell 1}}{f_{11}} \end{bmatrix}.$$

Similarly, adding to the jth column ($j = 2, 3, ...\ell$) from the first column multiplied by $-\frac{f_{1j}}{f_{11}}$, we have

$$f^{(2)}(\omega) = \begin{bmatrix} f_{11} & 0 & \cdots & 0 \\ 0 & f_{22} - f_{12}\frac{f_{21}}{f_{11}} & \cdots & f_{2\ell} - f_{1\ell}\frac{f_{21}}{f_{11}} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & f_{\ell 2} - f_{12}\frac{f_{\ell 1}}{f_{11}} & \cdots & f_{\ell\ell} - f_{1\ell}\frac{f_{\ell 1}}{f_{11}} \end{bmatrix} = \begin{bmatrix} f_{11} & \mathbf{0} \\ \mathbf{0} & g^{(2)} \end{bmatrix},$$

where the elements of matrix $g^{(2)} = [g_{rj}^{(2)}]$ have the form $g_{rj}^{(2)} = f_{rj} - \frac{f_{r1}f_{1j}}{f_{11}}$.

Notice that the diagonal element $g_{22}^{(2)}$ satisfies $g_{22}^{(2)}(\omega) = \frac{M_2(\omega)}{M_1(\omega)}$ a.e. If we denote $g^{(1)} = f^{(1)} = f$, then $f^{(2)}$ is obtained by using the row-column transformations on $f^{(1)}$. Now consider the matrix $g^{(2)}$,

$$g^{(2)} = \begin{bmatrix} f_{22} - f_{12}\frac{f_{21}}{f_{11}} & \cdots & f_{2\ell} - f_{1\ell}\frac{f_{21}}{f_{11}} \\ \vdots & \ddots & \vdots \\ f_{\ell 2} - f_{12}\frac{f_{\ell 1}}{f_{11}} & \cdots & f_{\ell\ell} - f_{1\ell}\frac{f_{\ell 1}}{f_{11}} \end{bmatrix}.$$

we apply the same transformation for $g^{(2)}$ to eliminate its first row and column except the leading coefficient, yielding

$$g^{(2)} = \begin{bmatrix} f_{22} - f_{12}\frac{f_{21}}{f_{11}} & \mathbf{0} \\ \\ \mathbf{0} & g^{(3)} \end{bmatrix}$$

it is easy to verify that $g_{33}^{(3)}(\omega) = \frac{M_3(\omega)}{M_2(\omega)}$. We then arrive at a new $\ell \times \ell$ matrix as

$$f^{(3)}(\omega) = \begin{bmatrix} f_{11} & 0 & \mathbf{0} \\ \\ 0 & f_{22} - f_{12}\frac{f_{21}}{f_{11}} & \mathbf{0} \\ \\ \mathbf{0} & \mathbf{0} & g^{(3)} \end{bmatrix}.$$

Continue this process until we reach a diagonal matrix $f^{(\ell)}(\omega)$, admitting the following form

$$f^{(\ell)}(\omega) = \begin{bmatrix} h_{11} & & & \\ & h_{22} & & \\ & & \ddots & \\ & & & h_{\ell\ell} \end{bmatrix}.$$

It is easy to see that the diagonal elements are

$$h_{11}(\omega) = M_1(\omega); \qquad h_{rr}(\omega) = \frac{M_r(\omega)}{M_{r-1}(\omega)}, \quad r = 2, 3, ...\ell.$$

It follows that $f(\omega)$ admits the following LDU-like decomposition.

The spectral density $f_x(\omega)$ can be decomposed as $f_x = g\ f^{(\ell)}\ g^*$, where the matrix function $g(\omega)$ is lower triangular with diagonal elements equal to one,

$$g(\omega) = \begin{bmatrix} 1 & 0 & ... & 0 \\ g_{21} & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{\ell 1} & g_{\ell 2} & ... & 1 \end{bmatrix}.$$

The off-diagonal non-zero elements are defined as $g_{rj} = \frac{g_{rj}^{(j)}}{h_{jj}}$, $r > j$; where $g_{rl}^{(l)}$ is determined by the recursion

$$g_{rj}^{(1)} = f_{rj}; \qquad g_{rj}^{(i)} = g_{rj}^{(i-1)} - \frac{g_{r,i-1}^{(i-1)} - g_{i-1,j}^{(i-1)}}{g_{i-1,i-1}^{(i-1)}}, \quad i = 2, 3, ...j.$$

Since the element of $f_x(\omega)$ are rational functions, the matrix transformation implies that elements of g and $f^{(\ell)}$ are rational as well. Next we define $g_{rj}(\omega) = \frac{P_{rj}(z)}{Q_{rj}(z)}$, where $z = e^{-i\omega}$. We extend the definition of z to the entire complex plane, and fix a column $j \in \{1, 2, ..., \ell\}$. Let $\alpha_p^{(j)}, p = 1, 2, ...$, denote the roots of the set of polynomials $\{Q_{rj}(z) : r = 1, ..., \ell\}$ that are located inside the unit circle, counting multiplicities. Define

$$c_j(z) = \prod_p (z - \alpha_p^{(j)}), \quad D_j(z) = \frac{h_{jj}(z)}{|c_j(z)|^2}.$$

Note that $D_j(z)$ is non-negative by construction. We can use Lemma S2 to decompose $D_j(z)$ as

$$D_j(z) = \left| \frac{\Phi_j(z)}{\Psi_j(z)} \right|^2 = \frac{\Phi_j(z)\Phi_j(\frac{1}{z})}{\Psi_j(z)\Psi_j(\frac{1}{z})}$$

on the unit circle, where we can choose $\Phi_j(z)$ and $\Psi_j(z)$ such that they have no zeros inside the unit disk (when extending the definition of z to the entire complex plane). The second equality follows from the real-coefficients assumption. If the polynomials have complex-valued coefficients, we need to conjugate the coefficients accordingly.

Now set

$$\widetilde{\Gamma}_{rj}(z) = g_{rj}(z)c_j(z)\frac{\Phi_j(z)}{\Psi_j(z)}, \quad r = 1, ..., \ell,$$

where $z = e^{-i\omega}$. Continuing this construction for all columns of g , we obtain the desired matrix $\widetilde{\Gamma}(z)$ such that $f_x(\omega) = \widetilde{\Gamma}(e^{-i\omega})\widetilde{\Gamma}^*(e^{-i\omega})$, where all elements of the matrix function

$$\widetilde{\Gamma}(z) = \begin{bmatrix} \widetilde{\Gamma}_{11}(z) & 0 & \dots & 0 \\ \widetilde{\Gamma}_{21}(z) & \widetilde{\Gamma}_{22}(z) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\Gamma}_{\ell 1}(z) & \widetilde{\Gamma}_{\ell 2}(z) & \dots & \widetilde{\Gamma}_{\ell\ell}(z) \end{bmatrix}$$

are analytic in the closed unit disk and hence in the $\mathbf{H}^2(\mathbb{D})$ space.

Derivations of Step 2: In step 1, we obtain

$$f_x(\omega) = \widetilde{\Gamma}\left(e^{-i\omega}\right)\widetilde{\Gamma}^*\left(e^{-i\omega}\right).$$

The Beurling-Blaschke factorization theorem states that every $\widetilde{\Gamma}(z) \in \mathbf{H}^2(\mathbb{D})$ can be written in the form

$$\widetilde{\Gamma}(z) = \Gamma(z)Q(z), \tag{S4.1}$$

where $Q(z)$ is an $\ell \times \ell$ matrix inner function. The proof of this theorem can be found in Rudin (1987), Theorem 17.17. The matrix generalization of this theorem can be found in Lindquist and Picci (2015), Theorem 4.6.5-4.6.8. The factorization is unique up to constant unitary matrices.[24] Since $\widetilde{\Gamma}(z)$ is rational, the outer function $\Gamma(z)$ is also rational as well. A rational outer function is completely characterized by the location of its zeros. That is, a rational function $\Gamma(z)$ is an outer function if and only if $\det(\Gamma(z)) \neq 0, \forall |z| < 1$. Hence, the inner function $Q(z)$ can be reduced to the Blaschke matrices satisfying

$$Q(z) = \prod_{j=1}^{n} \widetilde{B_j}(z) V_j, \tag{S4.2}$$

where $\widetilde{B_j}$ satisfies

$$\widetilde{B_j}(z) = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & \frac{z-z_j}{1-\bar{z}_j z} \end{bmatrix} = B_j^{-1}(z),$$

and z_j are zeros of $\det(Q(z))$ or $\det\left(\widetilde{\Gamma}(z)\right)$ satisfying $|z_j| < 1$. Here V_j are constant unitary matrices. In other words, the singular part of the rational inner function is absent (see Rudin (1987), Theorem 17.9 and Lindquist and Picci (2015), Theorem 4.6.11). Compared with the general definition of the Blaschke factors, we implicitly assume there are no zeros at $z = 0$ and omit the norm terms $\frac{\bar{z}_j}{|z_j|}$ since finite Blaschke products have no convergence issues. Combining (S4.1) and (S4.2), we have

$$\Gamma(z) = \widetilde{\Gamma}(z) \prod_{j=1}^{n} V_j^{-1} \left[\widetilde{B_j}(z)\right]^{-1} = \widetilde{\Gamma}(z) \prod_{j=1}^{n} V_j^{-1} B_j(z).$$

Note that the Blacheke-inner function satisfies $Q(z)Q^*(z) = I, \ \ \forall |z| = 1$, on the unit circle. The spectral density is preserved under the factorization

$$\Gamma(z)\Gamma^*(z) = \widetilde{\Gamma}(z) \prod_{j=1}^{n} V_j^{-1} B_j(z) \prod_{j=1}^{n} B_j^*(z) (V_j^{-1})^* \widetilde{\Gamma}^*(z) = f_x(\omega),$$

[24] The conditional uniqueness corresponds only to orthonormal Wold innovations. In fact, given a Wold representation $X_t = \Gamma(L)v_t$, the transformation $X_t = \Gamma(L)\Sigma\Sigma^{-1}v_t$ is also Wold fundamental provided that the constant matrix Σ is invertible. In this case, the Wiener-Hopf formula will be modified to contain Σ.

where $z = e^{-i\omega}$. Moreover, all zeros inside the unit disk are removed because

$$\begin{aligned}\det(\Gamma(z)) &= \det(\widetilde{\Gamma}(z)) \prod_{j=1}^{n} \det(V_j^{-1}) \prod_{j=1}^{n} \frac{1-\bar{z}_j z}{z - z_j} \\ &= \Upsilon(z) \prod_{j=1}^{n} (z - z_j) \prod_{j=1}^{n} \det(V_j^{-1}) \prod_{j=1}^{n} \frac{1-\bar{z}_j z}{z - z_j} \\ &= \Upsilon(z) \prod_{j=1}^{n} \det(V_j^{-1}) \prod_{j=1}^{n} (1 - \bar{z}_j z) \\ &\neq 0 \quad \forall |z| < 1\end{aligned}$$

where $\Upsilon(z) = \frac{\det(\widetilde{\Gamma}(z))}{\prod_{j=1}^{n}(z-z_j)}$ has no zeros inside the unit disk by construction. Unfortunately, the right multiplication of the Blaschke matrices also brought poles ($z = z_j$) for the element in the $\widetilde{\Gamma}(z)$ matrix that has no inside zeros. In order to maintain the analyticity inside the unit disk so that $\Gamma(z) \in \mathbf{H}^2_{\ell\times\ell}(\mathbb{D})$, we need to get rid of these by-product poles. We remove these poles inside the unit disk by setting appropriate constant unitary matrices V_j.

In practice, V_j can be obtained by the singular value decomposition in a sequential procedure. For $j = 1$, we have

$$\Gamma_1(z) = \widetilde{\Gamma}(z)\, V_1^{-1} B_1(z)$$

Without the constant unitary matrix V_1, the matrix transformation

$$\widetilde{\Gamma}(z)\, B_1(z) = \begin{bmatrix} \widetilde{\Gamma}_{11}(z) & 0 & ... & 0 \\ \widetilde{\Gamma}_{21}(z) & \widetilde{\Gamma}_{22}(z) & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\Gamma}_{\ell 1}(z) & \widetilde{\Gamma}_{\ell 2}(z) & ... & \widetilde{\Gamma}_{\ell\ell}(z) \end{bmatrix} \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & \frac{1-\bar{z}_1 z}{z - z_1} \end{bmatrix}.$$

It is clear the potential poles can only appear in the last column, if we assume that $\widetilde{\Gamma}_{\ell\ell}(z)$ has no zeros at $z = z_1$. To remove this pole, we follows Rozanov (1967) by employing the singular value decomposition (SVD) for $\widetilde{\Gamma}(z)$ at $z = z_1$

$$\widetilde{\Gamma}(z_1) = U_1 D_1 V_1 = U_1 \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & 0 \end{bmatrix} V_1.$$

By definition, the unitary matrices U_1 and V_1 are given by the (unitary) eigen-decomposition,

$$G_(z_1) = \widetilde{\Gamma}(z_1)\, \widetilde{\Gamma}^*(z_1) = U_1 \bar{D}_1 U_1^*; \quad \hat{G}_(z_1) = \widetilde{\Gamma}^*(z_1)\, \widetilde{\Gamma}(z_1) = V_1 \hat{D}_1 V_1^*.$$

Such decomposition always exists as $G_(z_1)$ and $\hat{G}_(z_1)$ are Hermitian and non-negative definite by construction. The diagonal matrices $\bar{D}_1$ and $\hat{D}_1$ contains eigenvalues of $G_(z_1)$ and $\hat{G}_(z_1)$, which are not necessarily distinct. The diagonal matrix D_1 in the SVD contains the singular values of $\widetilde{\Gamma}(z)$. The non-zero singular values $\{\lambda_1, \lambda_2, ...\lambda_p\}$ are the square root of the non-zero eigenvalues of $G_(z_1)$ and $\hat{G}_(z_1)$, which are not necessarily distinct. Since we know that $\det(\widetilde{\Gamma}(z_1)) = 0$,

$$\det(G_(z_1)) = \det\left(\widetilde{\Gamma}(z_1)\right)\det\left(\widetilde{\Gamma}(z_1)^*\right) = 0.$$

Therefore, there exists at least one singular value in D_1 that is zero, i.e. $p < d$.[25] Now evaluate $\Gamma_1(z)$ at $z = z_1$,

$$\begin{aligned}\Gamma_1(z_1) &= \widetilde{\Gamma}(z_1)\, V_1^{-1} B_1(z_1) = U_1 D_1 V_1 V_1^{-1} B_1(z_1) \\ &= U_1 \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & \frac{1-\bar{z}_1 z_1}{z_1 - z_1} \end{bmatrix}.\end{aligned}$$

Since the last column of D_1 are identically zero, the pole at $\frac{1-\bar{z}_1 z_1}{z_1 - z_1}$ vanishes at $z = z_1$. In other words, $\Gamma_1^{(i,j)}(z_1) < \infty$ are all well-defined without poles. On the other hand, condition (S4.3) ensures that zeros at $z = z_1$ is removed as well.

Now consider the second step $j = 2$,

$$\Gamma_2(z) = \Gamma_1(z)\, V_2^{-1} B_2(z).$$

Without the constant unitary matrix V_2,

$$\Gamma_1(z)\, B_2(z) = \Gamma_1(z) \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & \frac{1-\bar{z}_2 z}{z - z_2} \end{bmatrix}$$

would have poles in the last column. Note that $\Gamma_1(z)$ is no longer lower triangular after the first step transformation. To remove these poles at $z = z_2$, we employ the SVD again,

$$\begin{aligned}\Gamma_2(z_2) &= \Gamma_1(z_2)\, V_2^{-1} B_1(z_2) = U_2 D_2 V_2 V_2^{-1} B_2(z_2) \\ &= U_2 \begin{bmatrix} \widetilde{\lambda}_1 & 0 & ... & 0 \\ 0 & \widetilde{\lambda}_2 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & ..... & \frac{1-\bar{z}_2 z_2}{z_2 - z_2} \end{bmatrix},\end{aligned}$$

[25] The rank loss generally depends on the multiplicity of zeros in $\det(\widetilde{\Gamma}(z_1))$.

where $\{\widetilde{\lambda}_1, \widetilde{\lambda}_2, ...\widetilde{\lambda}_{\widetilde{p}}\}$ are the non-zero singular values. Again, there exists at least one zero in the diagonal of D_2 matrix ($\widetilde{p} < d$), since $\det(\Gamma_1(z_2)) = 0$. Arranging the zeros in the last positions of the diagonal, it follows immediately that $\Gamma_2^{(i,j)}(z_1) < \infty$ are all well-defined without poles, since the last column of D_2 are identically zero and the poles introduced by $\frac{1-\bar{z}_2 z_2}{z_2 - z_2}$ vanishes.

Continue this sequential procedure for all z_j, it follows that $\Gamma(z)$ is analytic (component wise) at $z = \{z_1, z_2, ...z_n\}$ inside the unit disk. By (S4.3), we conclude that $\Gamma(z)$ is indeed Wold (outer) spectral factor. The underlying construction can be trivially extended to the case with higher-order zeros, see Rozanov (1967), p47. In particular, the location of the Blaschke factor $\frac{1-\bar{z}_j z}{z - z_j}$ (along the diagonal) is inconsequential, as long as we put the zero in the corresponding diagonal position of D_j.

A Working Example of 2×3 Signal System To illustrate the use of our method, we consider an alternative specification of 2×3 signal system. Let the signal representation be

$$X_{it} = H(L)\eta_{it} \equiv \begin{bmatrix} \frac{1}{1-\rho_a L} & 1 & 0 \\ & & \\ F(L) & 0 & F(L) \end{bmatrix} \begin{bmatrix} \epsilon_{at} \\ \epsilon_{it} \\ \epsilon_{ut} \end{bmatrix},$$

where $F(z)$ is some an outer function in $\mathbf{H}^2(\mathbb{D})$.

Step 1: The spectral density $f_x(\omega)$ is given by

$$f_x(\omega) \equiv \begin{bmatrix} \frac{1}{(1-\rho_a z)(1-\rho_a z^{-1})}\sigma_a^2 + \sigma_i^2 & \frac{F(z^{-1})}{(1-\rho_a z)}\sigma_a^2 \\ & \\ \frac{F(z)}{(1-\rho_a z^{-1})}\sigma_a^2 & F(z)F(z^{-1})[\sigma_a^2 + \sigma_u^2] \end{bmatrix},$$

where $z = e^{-i\omega}$. The leading principal minors are given by

$$M_1(\omega) = f_{11}(\omega) = \frac{(1-\lambda_w z)(1-\lambda_w z^{-1})}{(1-\rho_a z)(1-\rho_a z^{-1})}\sigma_w^2;$$

$$M_2(\omega) = \det(f_x(\omega)) = \frac{F(z)F(z^{-1})}{(1-\rho_a z)(1-\rho_a z^{-1})}\left[\sigma_g^2(1-\lambda_w)(1-\lambda_w z^{-1}) - \sigma_a^4\right],$$

where we define $\sigma_p^2 = \sigma_a^2 + \sigma_u^2$ and $\sigma_g^2 = \sigma_w^2 \sigma_p^2$, $\lambda_w \in (0,1)$. Using Lemma S2,

$$\sigma_g^2(1-\lambda_w)(1-\lambda_w z^{-1}) - \sigma_a^4 = \sigma_h^2(1-\lambda_h z)(1-\lambda_h z^{-1}).$$

The new parameters σ_h and λ_h satisfy $\lambda_h = \frac{\lambda_w \sigma_g^2}{\sigma_h^2}$ and $\sigma_h^2(1+\lambda_h^2) = \sigma_g^2(1+\lambda_w^2) - \sigma_a^4$. In particular, we can pick a real $\lambda_h \in (0,1)$. Then the spectral density admits the following decomposition,

$$f_x(\omega) = \begin{bmatrix} 1 & 0 \\ g_{21}(\omega) & 1 \end{bmatrix} \begin{bmatrix} h_{11}(\omega) & 0 \\ 0 & h_{22}(\omega) \end{bmatrix} \begin{bmatrix} 1 & g_{21}^*(\omega) \\ 0 & 1 \end{bmatrix}.$$

The diagonal elements h_{11} and h_{22} are given by

$$h_{11}(\omega) = M_1(\omega); \quad h_{22}(\omega) = \frac{M_2(\omega)}{M_1(\omega)}.$$

In addition, we use the recursion formula to get $g_{21}(\omega) = \frac{g_{21}^{(1)}}{h_1 1} = \frac{f_{21}}{h_{11}}$. Therefore,

$$g_{21}(\omega) = \frac{\sigma_a^2}{\sigma_w^2} \frac{F(z)(1-\rho_a z)}{(1-\lambda_w z)(1-\lambda_w z^{-1})}.$$

Now fix the first column $j=1$, we know the only inside pole is at $z = \lambda_w$ in g_{21}. This implies

$$C_1(z) = (z-\lambda_w); \quad D_1(z) = \frac{h_{11}(z)}{|C_1(z)|^2} = \left|\frac{\Phi_1(z)}{\Psi_1(z)}\right|^2.$$

Hence $\frac{\Phi_1(z)}{\Psi_1(z)} = \frac{\sigma_w}{1-\rho_a z}$. This in turn implies

$$\tilde{\Gamma}_{11}(z) = g_{11} C_1(z) \frac{\Phi_1(z)}{\Psi_1(z)} = \sigma_w \frac{z-\lambda_w}{1-\rho_a z}; \quad \tilde{\Gamma}_{21}(z) = g_{21} C_1(z) \frac{\Phi_1(z)}{\Psi_1(z)} = \frac{\sigma_a^2}{\sigma_w} \frac{F(z) z}{(1-\lambda_w z)}.$$

We repeat this procedure for the second column. Notice that the second column of g are constants, therefore, $C_2(z) = 1$ and $\frac{\Phi_2(z)}{\Psi_2(z)} = \frac{\sigma_h}{\sigma_w} \frac{F(z)(1-\lambda_h z)}{(1-\lambda_w z)}$. In the end, we obtain the lower-triangular matrix

$$\tilde{\Gamma}(z) = \begin{bmatrix} \sigma_w \frac{z-\lambda_w}{1-\rho_a z} & 0 \\ \frac{\sigma_a^2}{\sigma_w} \frac{F(z) z}{(1-\lambda_w z)} & \frac{\sigma_h}{\sigma_w} \frac{F(z)(1-\lambda_h z)}{(1-\lambda_w z)} \end{bmatrix}.$$

clearly, $\tilde{\Gamma}(z) \in \mathbf{H}^2_{2\times 2}(\mathbb{D})$.

Step 2: We remove the inside zeros at $z = \lambda_w$ to achieve the Wold fundamental representation. Using the Blaschke factorization, we have $\Gamma(z) = \tilde{\Gamma}(z) V_1^{-1} B(z)$, where $B(z) = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1-\lambda_w z}{z-\lambda_w} \end{bmatrix}$ and V_1 satisfies the unitary eigen-decomposition of $\hat{G}(\lambda_w) = \widetilde{\Gamma}^*(\lambda_w) \widetilde{\Gamma}(\lambda_w) = V_1 \hat{D}_1 V_1^*$. It is easy to check that eigenvalues of Hermitian matrix $\hat{G}(\lambda_w)$ are distinct. Therefore, we can pick two eigenvectors from the two eigenvalues, which are necessarily orthogonal by the spectral theorem. Normalizing these two eigenvectors yields the unitary matrix as desired,

$$V_1 = \begin{bmatrix} \sqrt{\frac{h^2}{1+h^2}} & \sqrt{\frac{1}{1+h^2}} \\ \sqrt{\frac{1}{1+h^2}} & -\sqrt{\frac{h^2}{1+h^2}} \end{bmatrix},$$

where $h = \frac{\sigma_a^2}{\sigma_h} \frac{\lambda_w}{(1-\lambda_h \lambda_w)}$. The resulting matrix $\Gamma(z)$ is the Wold fundamental matrix

$$\Gamma(z) = \begin{bmatrix} \sigma_w \frac{z-\lambda_w}{1-\rho_a z} V_1^{(11)} & \sigma_w \frac{1-\lambda_w z}{1-\rho_a z} V_1^{(12)} \\ F(z) \frac{\sigma_h V_1^{(12)}}{\sigma_w} & F(z) \frac{V_1^{(12)} \sigma_a^2}{\sigma_w (1-\lambda_h \lambda_w)} \end{bmatrix}.$$

Finally, we can transform $\Gamma(z)$ into an upper triangular form by right multiplication of another unitary matrix V_2,

$$V_2 = \begin{bmatrix} \sqrt{\frac{1}{1+x^2}} & \sqrt{\frac{x^2}{1+x^2}} \\ -\sqrt{\frac{x^2}{1+x^2}} & \sqrt{\frac{1}{1+x^2}} \end{bmatrix},$$

where $x = \frac{\sigma_h(1-\lambda_h\lambda_w)}{\sigma_a^2}$. After some algebraic simplifications, we obtain

$$\Gamma(z) = \begin{bmatrix} \frac{\sigma_h}{\sigma_p}\frac{1-\lambda_h z}{1-\lambda_w z} & \frac{\sigma_a^2}{\sigma_p}\frac{1}{1-\rho_a z} \\ 0 & F(z)\sigma_p \end{bmatrix}.$$

Since we assume $F(z)$ is outer, i.e. has no roots in the open unit disk, $\Gamma(z)$ is the Wold representation.

References

Bacchetta, Philippe, and Eric van Wincoop, 2008, Higher Order Expectations in Asset Pricing, *Journal of Money, Credit and Banking* 40, 837-866.

Brockwell, Peter J., and Richard A. Davis, 2002, *Time Series: Theory and Methods*, Springer Series in Statistics, 2 edn

Hansen, Lars P., and Thomas J. Sargent, 1980, Formulating and Estimating Dynamic Linear Rational Expectation Models, *Journal of Economic Dynamics and Control* 2, 7–46.

Kailath, Thomas, Ali H. Sayed, and Babak Hassibi, 2000, *Linear Estimation*, Prentice Hall, Inc.

Kailath, Thomas, and Ali H. Sayed, 2001, A Survey of Spectral Factorization Methods, *Numerical Linear Algebra with Applications* 08, 467-496

Kasa, Kenneth, Todd B. Walker, and Charles H. Whiteman, 2014, Heterogeneous Beliefs and Tests of Present Value Models, *Review of Economic Studies* 81, 1137-1163.

Lindquist, Anders, and Giorgio Picci, 2015, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification*, Springer-Verlag Berlin Heidelberg

Rondina, Giacomo, and Todd B. Walker, 2015, Dispersed Information and Confounding Dynamics, working paper, Indiana University.

Rozanov, Yu A., 1967, *Stationary Random Processes*, San Francisco: Holden-Day.

Rudin, Walter, 1987, *Real and Complex Analysis*, McGraw-Hill Companies, Inc. 3 edn