A Single-Judge Solution to Beauty Contests*

Zhen Huo
Yale University

Marcelo Pedroni
University of Amsterdam

Tuesday 21st May, 2019

Abstract

We show that the equilibrium policy rule in beauty-contest models is equivalent to that of a single agent's forecast of the economic fundamental. This forecast is conditional on a modified information process, which simply discounts the precision of idiosyncratic shocks by the degree of strategic complementarity. The result holds for any linear Gaussian signal process (static or persistent, stationary or non-stationary, exogenous or endogenous), and also extends to network games. Theoretically, this result provides a sharp characterization of the equilibrium and its properties under dynamic information. Practically, it provides a straightforward method to solve models with complicated information structures.

Keywords: Dispersed Information, Beauty-Contest Model, Higher-Order Beliefs
JEL classifications: C72, D80, E30

*We are thankful to Darij Grinberg and Qirui Li for their help with some of the proofs. We are also grateful for the comments of the editor and four referees. The paper benefits from discussions with George-Marios Angeletos, Adrien Auclert, Jess Benhabib, Dirk Bergemann, Alex Clymo, Xavier Gabaix, Tibor Heumann, Simas Kučinskas, Jennifer La'O, Thomas Lubik, Giuseppe Moscarini, Diego Perez, Monika Piazzesi, Thomas Sargent, Martin Schneider, Tony Smith, Venky Venkateswaran, Aleh Tsyvinski, and Muhamet Yildiz. We would also like to thank the seminar participants at NYU, Yale, Stanford, University of Hong Kong, University of Amsterdam, Barcelona Summer Forum EGF, SWUFE, SED, and SEM Conference. This paper was previously circulated under the title "Infinite Higher-Order Beliefs and First-Order Beliefs: An Equivalence Result".

1 Introduction

"It is not a case of choosing those that, to the best of one's judgment, are really the prettiest, nor even those that average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practice the fourth, fifth and higher degrees." — Keynes (1936)

Many economic problems with information frictions can be formalized as a beauty-contest problem. The effects of monetary policy (Woodford, 2002), sentiments driven business cycle fluctuations (Angeletos and La'O, 2010; Benhabib, Wang, and Wen, 2015), and the optimal degree of transparency in central bank communication (Morris and Shin, 2002) are some examples. In these beauty-contest models, a common feature is that agents care about certain pay-off relevant economic fundamentals, and due to strategic complementarity,[1] about other agents' actions and beliefs, as well.

As originally emphasized by Keynes (1936), in these beauty-contest models with asymmetric information, the dependence of an agent's action on the actions, and therefore, beliefs of others, makes it necessary to infer others' beliefs about others' beliefs, and so on. These higher-order beliefs introduce rich dynamics that help account for certain salient features of the data, but they are typically difficult to characterize. In this paper, we show that, in a large class of models, in fact, agents' equilibrium behavior can be equivalently represented by the solution to a single-agent problem. In this alternative problem, agents only need to forecast economic fundamentals but are equipped with less precise private information. In Keynes' language, our results imply that the nature of the optimal strategy, in a sense, is still to choose the prettiest based on the best of one's judgment while downplaying any idiosyncrasies in one's taste.

The rationale for this equivalence can be understood in a simple static setting. Imagine a single-agent problem of forecasting an economic fundamental having received a public and a private signal about it. Bayesian inference implies that the relative weights put on the public and private signals should be proportional to their precision. In contrast, an agent in a beauty contest, due to strategic complementarity, also needs to forecast others' actions. The public signal is more useful for that purpose. As a result, the optimal strategy in a beauty contest is to assign additional weight to the public signal or discount the private signal. This policy rule—underweighting private signals—turns out to be exactly equivalent to that of a single-agent problem with the precision of the private signal reduced by a fraction equal to the degree of strategic complementarity.

Main result In this paper, we formalize this equivalence and generalize it to models with dynamic information. We first define an *α-modified signal process*, where α parameterizes the strength of

[1] We formulate our argument based on the case with strategic complementarity most of the time, but our results extend to the case with strategic substitutability.

strategic complementarity among agents' actions. In this modified signal process, the precision of all idiosyncratic shocks is multiplied by $1-\alpha$, while those of common shocks remain the same. Here, the key distinction is between idiosyncratic and common shocks, rather than between private and public signals.[2] We then define an auxiliary single-agent problem, in which the objective is to simply forecast the economic fundamental using the α-modified signals while completely disregarding others' actions.

Our main result (Theorem 1) establishes that the policy rule in the equilibrium of the beauty-contest model is the same as in the corresponding single-agent problem with α-modified signals. This result yields a sharp characterization of the equilibrium. It bypasses the complexity of higher-order expectations and shows that the optimal policy rule can be understood via a modified forecast of the exogenous economic fundamental. It also avoids the fixed-point which results from the interaction between rational agents and provides a new window to understand the underlying general-equilibrium (GE) feedback effects. This result does not imply that higher-order beliefs are not important, but it does imply that the computation of equilibrium can be greatly simplified.

The very nature of non-cooperative solution concepts in game theory is to transform a collective decision problem into a family of single-player decision problems. What is special about our solution is the simple form it takes and how this form is preserved across a class of models. Importantly, the equivalence is proved under a general information structure that allows for an arbitrary number of Gaussian shocks and any linear signal process—static or persistent, stationary or non-stationary, exogenous or endogenous. For all different signal processes, the modification necessary to yield the equivalence is always the same: discounting the precision of idiosyncratic shocks by a fraction α. A similar modification also applies to models in which agents choose multiple correlated actions each period.

This generality makes our method particularly useful for applied work where the fundamental is often persistent and signals are correlated over time. In this case, one needs to find a set of sufficient statistics or state variables that can summarize the past information. In contrast to complete-information models in which the right state variables are easy to obtain (such as capital and TFP in real-business-cycle models), with dispersed information this is usually not the case. When signals permit a state-space representation, the single-agent solution yields a formula for the policy rule that is ready for computation. It also implies that, even though recording all higher-order expectations would require an infinite number of state variables, solving for the optimal action only requires a finite number of state variables.

Another immediate implication of the main result is that the aggregate outcome in a beauty-contest

[2] Relative to the previous two-signal example, a signal can contain multiple common and idiosyncratic shocks in general, and the α-modified signal process determines the exact extent to which agents should discount the signal.

model is equivalent to the average forecast in the single-agent problem. It follows that, in particular, two determinants of the aggregate outcome—the strength of strategic complementarity and the degree of information frictions—cannot be separately identified solely by aggregate time series. When lacking evidence on one of them, the fact that the single-agent model yields more dispersed forecast errors and actions can help achieve identification when cross-sectional moments are available. With independent micro evidence on both, our solution method can simply be used to quantify the role of information frictions.

Applications The single-agent solution provides an alternative perspective to understand properties of beauty-contest models, extends existing results to general information structures, and expands the type of questions one can explore.

We first revisit two important themes in the dispersed information literature: inertia and GE attenuation. With dynamic information, the aggregate outcome displays additional inertia relative to its complete information counterpart as it inherits the persistence of all higher-order expectations (Woodford, 2002; Nimark, 2017; Angeletos and Huo, 2018). Through the lens of the single-agent solution, the aggregate outcome is anchored by past outcomes as forecasts are anchored by priors. We contribute to show that the strength of the linkages to past outcomes is pinned down by the Kalman gain matrix associated with the α-modified signals and we provide its exact formula.

The GE feedback underlying beauty-contest models is captured by the strategic complementarity, and incomplete information tends to attenuate its effects (Angeletos and Lian, 2016a). Our single-agent solution contributes to the literature by providing an alternative interpretation and a generalization of existing comparative statics. With what is known as an independent-value best response, we show that the GE effects are captured by the first-order expectation of the fundamental with α-modified signals. It is then immediate that the GE effects are attenuated when information is incomplete, since the expectation of the fundamental moves less than the fundamental itself.[3] Moreover, a larger degree of complementarity implies a noisier signal in the single-agent solution, and therefore a stronger attenuation effect. These results hold for a general class of information processes.[4]

We also utilize the fact that the single-agent solution can deal with nonstationary processes to demonstrate that the results on the social value of public information in Morris and Shin (2002) can be reversed in transition. In a static setting, Morris and Shin (2002) prove that increasing the precision of public information can be detrimental to social welfare. We first extend this finding to dynamic information when comparing steady-state welfare. The result is actually strengthened in

[3]For a common-value best response, the attenuation effects imply that the volatility of aggregate outcome is always bounded above by that of the average expectation of fundamental.

[4]In Angeletos and Lian (2016b), similar comparative statics are established for a two-signal static environment. Also, differently from Bergemann and Morris (2013) we fix the underlying information structure when varying the degree of strategic complementarity.

the sense that a more persistent fundamental leads to a wider range of public information precision levels for which more precise information can be detrimental. However, we also show that comparing steady states can be misleading since welfare can be improved for quite a long time during the transition before decreasing towards a lower level in the new steady state.[5]

Extensions We explore two extensions of our results. First, we consider a network game with incomplete information, in which agents have heterogeneous payoff structures and information structures. The equilibrium actions turn out to be a weighted average of forecasts of fundamentals conditional on different modified signals. When information is persistent, our characterization of the equilibrium in terms of forecasts of fundamentals remains convenient and intuitive. It allows for a recursive representation and avoids the burden of carrying an infinite number of signals, which is new to the existing literature.

Second, we consider beauty contest models with intertemporal strategic complementarities. Under a stylized information structure with forward-looking complementarity (Allen, Morris, and Shin, 2006; Angeletos and Lian, 2018; Angeletos and Huo, 2018), we show that our single-agent solution still works, though the modification to the signal process is more involved. Because of the forward-looking behavior, the required discounting of private signals is not simply the degree of strategic complementarity anymore but also hinges on the details of the information structure. More generally, when an individual's payoff depends on future and past aggregate actions in an arbitrary way, we show that our single-agent solution may not hold, drawing a limit to the applicability of our results.

Related Literature Our results complement a large literature on dispersed information and imperfect coordination. The classical papers that explore the general properties of beauty-contest models include Morris and Shin (2002), Angeletos and Pavan (2007), Bergemann and Morris (2013), and Angeletos and Lian (2016b). These papers focus on a static setting. The single-agent solution contributes to this literature by providing a novel equilibrium characterization, which accommodates essentially arbitrary information structures under Gaussian shocks. This generality becomes particularly useful for models with dynamic information.

In terms of the applications in macroeconomics, our paper is closely related to Woodford (2002) and Angeletos and La'O (2010), which study models with dispersed and dynamic information. These papers numerically explore the persistence of the aggregate outcome under a particular signal structure. We offer an exact formula that characterizes the entire dynamics of the aggregate outcome for general information structures. With this formula, the comparative statics results, and the single-agent perspective, our work complements the large literature that studies how information frictions shape

[5] Amador and Weill (2012) also considers the transition dynamics. In their setting, the fundamental is fixed over time and the welfare decreasing is due to that the informativeness of endogenous signal becomes worse, while we allow the fundamental to fluctuate over time and we consider only exogenous information as in Morris and Shin (2002).

business cycles and interact with monetary policy (Hellwig, 2004; Lorenzoni, 2009; Venkateswaran, 2014; Drenik and Perez, 2015).[6]

Our contribution to the literature on network games with incomplete information (De Martí and Zenou, 2015; Bergemann, Heumann, and Morris, 2017) is to provide an equilibrium characterization that extends to environments with dynamic information and are not subject to the curse of dimensionality. The extension of our results to forward-looking complementarity is related to Nimark (2017), Huo and Takayama (2017b), and Angeletos and Huo (2018). These papers, however, do not recast the policy rule as a single-agent solution.

We make two contributions to the literature on computing dispersed-information models: (1) our solution allows one to explore environments with non-stationary processes, which encompasses models with time-varying parameters and regime switching, for instance; (2) for models with exogenous information, our method delivers the solution without the need to solve a fixed-point problem—the policy rule can be obtained from a straightforward linear forecast. The guess-and-verify approach used in Woodford (2002) and Angeletos and La'O (2010) is useful for some relatively low-dimensional signal processes, while the approach in this paper eliminates the need to guess the solution altogether. The method in Huo and Takayama (2017b) can solve a larger class of models, but it relies on frequency-domain techniques that may involve solving for the roots of high-order polynomials. This paper stays within the time domain where easy and robust algorithms such as the Kalman filter can be applied. The single-agent solution method is also useful for solving models with endogenous information, which complements existing methods (Sargent, 1991; Lorenzoni, 2009; Nimark, 2017) with the advantage that our solution requires a smaller number of state variables.

Finally, there is a large literature that emphasizes the importance of higher-order beliefs (e.g. Rubinstein (1989), Carlsson and van Damme (1993), Kajii and Morris (1997), Weinstein and Yildiz (2007a), and Ely and Pęski (2011)). These results can be unsettling since it is hard to imagine that one could observe an individual's infinite belief hierarchy, yet it could be that a model's prediction can change if beliefs of arbitrarily high order are modified. Thus, it is relevant to point out that the beauty contest models studied in this paper satisfy Weinstein and Yildiz (2007b)'s criterion of 'global stability under uncertainty' which implies that beliefs of sufficiently high order have a vanishing effect on equilibrium outcomes. This does not mean, however, that higher-order beliefs are not important—they can still have quantitatively and qualitatively large effects.

The rest of the paper is organized as follows. Section 2 explains the single-agent solution in a simple static model. Section 3 sets up the environment and defines equilibrium. Section 4 proves the main result and variations of it. Section 5 provides several applications of the single-agent solution. Section 6 shows how it applies to models with endogenous information. Section 7 discusses the extensions

[6] This is a very partial list. We refer to Angeletos and Lian (2016b) for a much more comprehensive review.

to network games and to models with dynamic complementarity. Section 8 concludes.

2 A Motivating Example

To illustrate the basic idea and motivate our general result, we start from the static beauty-contest model considered by Morris and Shin (2002). The economy consists of a continuum of agents, indexed by i, with payoff functions given by

$$U(y_i, y, \theta) = -(1-\alpha)(y_i - \theta)^2 - \alpha(y_i - y)^2.$$

Agent i's best response, y_i, is a weighted average of her forecast of an exogenous fundamental, θ, and the aggregate action, y. That is,

$$y_i = (1-\alpha)\,\mathbb{E}_i[\theta] + \alpha\,\mathbb{E}_i[y], \qquad \text{with} \qquad y = \int y_i. \tag{2.1}$$

The parameter α, which we assume satisfies $\alpha \in (-1, 1)$,[7] determines the degree of strategic complementarity ($\alpha > 0$) or substitutability ($\alpha < 0$) between agents' actions. The operator $\mathbb{E}_i[\cdot]$ denotes the expectation conditional on agent i's information set, which consists of a public signal z and a private signal x_i about the exogenous fundamental θ,

$$z = \theta + \varepsilon, \quad \text{and} \quad x_i = \theta + \nu_i, \tag{2.2}$$

where $\varepsilon \sim \mathcal{N}(0, \tau_\varepsilon^{-1})$, and $\nu_i \sim \mathcal{N}(0, \tau_\nu^{-1})$ stand for public and private noise, and τ_ε and τ_ν are their respective precisions. We assume that agents have a common prior about the fundamental θ, that is, $\theta \sim \mathcal{N}(0, \tau_\theta^{-1})$.[8]

A useful way to understand the logic behind agents' actions is to examine their higher-order expectations. Define higher-order expectations recursively in the following way:

$$\overline{\mathbb{E}}^0[\theta] \equiv \theta, \qquad \text{and} \qquad \overline{\mathbb{E}}^k[\theta] \equiv \int \mathbb{E}_i\left[\overline{\mathbb{E}}^{k-1}[\theta]\right].$$

By consecutive substitution of the equations in (2.1), we can write agent i's action as a weighted

[7] With $\alpha > 1$, there could be multiple equilibria if the action is bounded. By assuming $\alpha \in (-1, 1)$, we can guarantee the existence of a unique equilibrium that can be represented by the sum of infinite higher-order expectations, which satisfies the 'global stability under uncertainty' condition provided by Weinstein and Yildiz (2007b).

[8] Morris and Shin (2002) assume that agents have an improper prior about θ. This change is mostly immaterial, but connects better to our subsequent dynamic analysis.

average of her higher-order expectations,[9]

$$y_i = (1-\alpha)\sum_{k=0}^{\infty}\alpha^k\ \mathbb{E}_i\left[\overline{\mathbb{E}}^k[\theta]\right].$$

One can immediately see that an individual agent's optimal action depends not only on her own assessment of the fundamental, but also on her expectation about the average assessment of the fundamental, and so on. This is the original idea put forward by Keynes about the nature of beauty contests. Under the normality assumptions made above, all the higher-order expectations can be written as linear combinations of the signals. It is, then, straightforward to define the Bayesian-Nash equilibrium of this economy:

Definition 1. *Given the signal process (2.2), a linear Bayesian-Nash equilibrium is a policy rule* $\mathbf{h} = \{h_1, h_2\} \in \mathbb{R}^2$*, such that*

$$y_i = h_1 z + h_2 x_i$$

satisfies the equations in (2.1).

The method of undetermined coefficients yields the following equilibrium policy rule:

$$y_i = \frac{\tau_\varepsilon}{\tau_\theta + \tau_\varepsilon + (1-\alpha)\tau_\nu} z + \frac{(1-\alpha)\tau_\nu}{\tau_\theta + \tau_\varepsilon + (1-\alpha)\tau_\nu} x_i. \tag{2.3}$$

This policy rule has the property that agents put relatively more weight on the public signal when the degree of strategic complementarity, α, is higher. The reason is that the public signal is more informative about the aggregate action and has an additional coordination role. A higher α implies that agents have a stronger incentive to be closer to the aggregate action, and therefore rely less on private signals. So far, this solution and its interpretation is reminiscent of Morris and Shin (2002). Now, we offer an alternative perspective.

A Single-Agent Solution For comparison, the forecast about the fundamental itself is

$$\mathbb{E}_i[\theta] = \frac{\tau_\varepsilon}{\tau_\theta + \tau_\varepsilon + \tau_\nu} z + \frac{\tau_\nu}{\tau_\theta + \tau_\varepsilon + \tau_\nu} x_i, \tag{2.4}$$

where the weights on the public and private signals are proportional to their relative precisions.

[9] To see this, note that

$$\begin{aligned} y_i =& (1-\alpha)\mathbb{E}_i[\theta_t] + \alpha\mathbb{E}_i[y] = (1-\alpha)\mathbb{E}_i[\theta] + \alpha\mathbb{E}_i\left[\int (1-\alpha)\mathbb{E}_j[\theta] + \alpha\mathbb{E}_j[y]\right] = \\ =& (1-\alpha)\mathbb{E}_i[\theta] + \alpha(1-\alpha)\mathbb{E}_i\left[\overline{\mathbb{E}}^1[\theta]\right] + \alpha^2\mathbb{E}_i\left[\int \mathbb{E}_j[y]\right] = \dots = (1-\alpha)\sum_{k=0}^{\infty}\alpha^k\ \mathbb{E}_i\left[\overline{\mathbb{E}}^k[\theta_t]\right]. \end{aligned}$$

The equilibrium policy rule (2.3) resembles the simple forecast rule (2.4) of the economic fundamental. They differ only with respect to the precision of the private signal noise. Define the following modified private signal

$$\widetilde{x}_i = \theta + \widetilde{\nu}_i, \quad \widetilde{\nu}_i \sim \mathcal{N}(0, ((1-\alpha)\tau_\nu)^{-1}),$$

with the precision of the private signal noise discounted by the degree of strategic complementarity α, while the public signal remains unchanged. Let $\widetilde{\mathbb{E}}_i[\cdot]$ denote the expectation conditional on the modified signals.

The coefficients of the equilibrium policy rule, $\{h_1, h_2\}$, can be obtained immediately by solving the problem of forecasting θ, given the modified signals, which leads to

$$\widetilde{\mathbb{E}}_i[\theta] = \frac{\tau_\varepsilon}{\tau_\theta + \tau_\varepsilon + (1-\alpha)\tau_\nu} z + \frac{(1-\alpha)\tau_\nu}{\tau_\theta + \tau_\varepsilon + (1-\alpha)\tau_\nu} \widetilde{x}_i = h_1 z + h_2 \widetilde{x}_i.$$

Hence, the weights the agents put on signals in equilibrium are exactly the same as the weights in a simple forecast of the fundamental with adjusted precision of private signals.

To see where the discounting comes from, consider the case where $\tau_\varepsilon = 0$, so that agents disregard the public signal. This specification delivers a particularly simple structure for higher-order expectations,

$$\overline{\mathbb{E}}^1[\theta] = \lambda\theta, \quad \text{and} \quad \overline{\mathbb{E}}^k[\theta] = \lambda \overline{\mathbb{E}}^{k-1}[\theta]. \tag{2.5}$$

where $\lambda \equiv \frac{\tau_\nu}{\tau_\theta + \tau_\nu}$. Then, the aggregate action can be expressed as

$$y = (1-\alpha) \sum_{k=0}^{\infty} \alpha^k \, \overline{\mathbb{E}}^{k+1}[\theta] = (1-\alpha)\lambda \sum_{k=0}^{\infty} (\alpha\lambda)^k \, \theta \tag{2.6}$$

Next, consider a variant representative-agent economy where agents share the same first-order expectation, $\overline{\mathbb{E}}[\theta] = \widetilde{\lambda}\theta$, and all agents' beliefs are common knowledge. In this economy, all higher-order expectations collapse to the first-order expectation, $\overline{\mathbb{E}}^k[\theta] = \overline{\mathbb{E}}[\theta]$ for all k, and the aggregate action is akin to a simple forecasting problem. The higher-order expectation representation still applies, so

$$y = (1-\alpha) \sum_{k=0}^{\infty} \alpha^k \, \overline{\mathbb{E}}^{k+1}[\theta] = (1-\alpha)\widetilde{\lambda} \sum_{k=0}^{\infty} \alpha^k \, \theta. \tag{2.7}$$

In equation (2.6), the lack of common knowledge leads to the additional discounting $(\alpha\lambda)^k$, compared with α^k in equation (2.7). Note that by setting $\widetilde{\lambda}$ to a lower value than λ such that

$$\widetilde{\lambda}\sum_{k=0}^{\infty}\alpha^k = \lambda\sum_{k=0}^{\infty}(\alpha\lambda)^k,$$

the aggregate action in these two economies are equivalent. It is easy to verify that the precise way to lower $\widetilde{\lambda}$ is to discount the private signal's precision by a fraction of α.[10]

Several remarks are in order. First, the fact that the equilibrium can be expressed by a particular first-order expectation is not surprising on its own. In fact, in equilibrium, agents' actions are always a first-order expectation about $(1-\alpha)\theta+\alpha y$. What is less obvious about our single-agent solution is that the equilibrium outcome is the same as the first-order expectation about the fundamental itself, and the required modification of the information structure is simple, known *ex ante*, and independent of the details of the signal process.

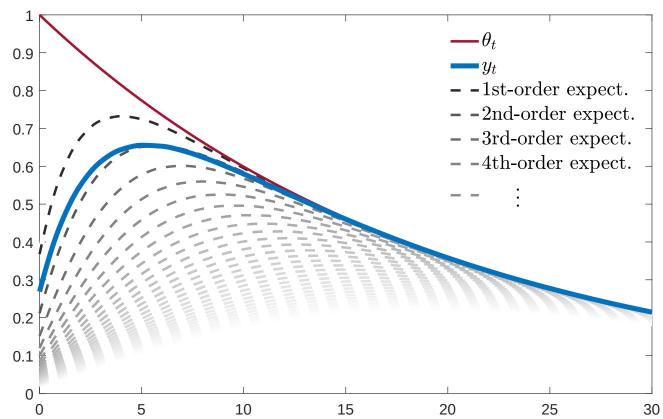


FIGURE 1: Impulse Response of Higher-Order Expectations with Persistent Information

Parameters: $\tau_\eta = 1$, $\rho = 0.95$, $\tau_\nu = 0.25$, and $\alpha = 0.8$.

To appreciate this point, notice that the proportionality between higher-order expectations in equation (2.5) is only valid for this very particular information structure. When information is persistent, higher-order expectations typically have complex dynamics. For example, consider a dynamic version of the setup above in which the fundamental θ_t follows an AR(1) process,[11]

$$\theta_t = \rho\theta_t + \eta_t, \qquad \eta_t \sim \mathcal{N}(0,\tau_\eta^{-1}). \tag{2.8}$$

[10]To be explicit, it follows that:

$$\widetilde{\lambda}\frac{1}{1-\alpha} = \lambda\frac{1}{1-\alpha\lambda} \Rightarrow \widetilde{\lambda} = \frac{\tau_\nu}{\tau_\theta+\tau_\nu}\frac{1-\alpha}{1-\alpha\frac{\tau_\nu}{\tau_\theta+\tau_\nu}} \Rightarrow \widetilde{\lambda} = \frac{(1-\alpha)\tau_\nu}{\tau_\theta+(1-\alpha)\tau_\nu}.$$

[11]The exact setup is explained in Example 1.

Figure 1 shows the impulse responses of higher-order expectations to the fundamental shock, η_t. As the order increases, more state variables are required to describe the dynamics of higher-order expectations. In fact, the k-th order higher-order expectation follows an ARMA$(k+1, k-1)$ process.

A rescaling of the single-agent forecast with unmodified signal precision would allow one to match the initial dampening of the aggregate outcome response. However, this modification cannot match the entire impulse response, as higher-order expectations exhibit less amplitude *and* more persistence, with the peak of the hump happening at later periods. In contrast, a single-agent forecast with appropriately discounted private signal precision matches the entire dynamics of the aggregate outcome. To see why this is possible, note that the optimal forecast about the fundamental under Bayesian learning is a weighted average of its prior mean and the new signal. With a less precise signal, agents lower the weight attached to the new signal, and simultaneously increase the weight put on the prior mean. The former reduces the amplitude of the direct effect of a shock to θ_t, and the latter increases its persistence.

In general, with persistent information, solving and analyzing beauty-contest models becomes significantly more difficult. In contrast to the static case, the standard method of undetermined coefficients is not useful because, in principle, an infinite number of coefficients needs to be solved for. This is where our results become most useful since the single-agent solution survives under general information structures.

3 Preliminaries for Beauty-Contest Models

In this section, we introduce a dynamic setting with a much richer information structure. In the following section, we prove the equivalence result.

3.1 Beauty-Contest Model

Best Response Denote agent i's action in period t by y_{it}. For presentation purposes, here we only consider the case in which the action is univariate. In Section 4.3, we extend this model to allow for multiple, and possibly correlated, actions. The best response function is similar to the one used in Section 2,

$$y_{it} = (1-\alpha)\,\mathbb{E}_{it}\left[\theta_t\right] + \alpha\,\mathbb{E}_{it}\left[y_t\right], \qquad \text{with} \qquad y_t = \int y_{it}. \tag{3.1}$$

Analogously to Section 2, agents i's optimal action can be expressed as

$$y_{it} = (1-\alpha)\sum_{k=0}^{\infty}\alpha^k\,\mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k[\theta_t]\right]. \tag{3.2}$$

Note that with dynamic information, the higher-order expectations are also time varying. To guarantee the existence and uniqueness of the equilibrium, we assume that the degree of strategic complementarity is less than 1 in absolute value.

Assumption 1. *The degree of strategic complementarity* $\alpha \in (-1, 1)$.

This best response can arise in various economic environments. In Section 2, we see that it can be the result of a quadratic loss function, as in Morris and Shin (2002), Angeletos and Pavan (2007), and Hellwig and Veldkamp (2009), for instance. The same type of best response can arise in real business cycle models (Angeletos and La'O, 2010, 2013; Venkateswaran, 2014; Huo and Takayama, 2017a), and in monetary economies (Woodford, 2002; Maćkowiak and Wiederholt, 2009; Drenik and Perez, 2015). We include two detailed examples in Online Appendix E for illustration.

Though the best response function we consider here is widely used in the literature, there are two limitations. First, it is implicitly assumed that all agents are symmetric, which excludes the possibility that agents have heterogeneous payoff or information structures, such as in a network game. Second, the individual action can only depend on the current aggregate action. This excludes intertemporal complementarities, in which agents' payoffs today may depend on future or past aggregate actions. In Section 7, we discuss to what extent our results can be extended to these environments.

Information Structure Next, we specify the stochastic process for the economic fundamental and the signals agents receive about it. We first consider the exogenous information case, by which we mean that the signals do not depend on agents' actions. This specification allows us to present our results more transparently. However, our single-agent solution is not limited to the exogenous information case. In Section 6, we show that our main results also hold if information is endogenous, with signals that do depend on agents' actions.

Throughout, we consider the following signal process with discrete time. In period t, agent i observes

$$\mathbf{x}_{it} = \mathbf{M}_t(L)\,\boldsymbol{\varepsilon}_{it}, \tag{3.3}$$

where the signal $\mathbf{x}_{it}$ is an $r \times 1$ stochastic vector, the shock $\boldsymbol{\varepsilon}_{it}$ is an $m \times 1$ stochastic vector, and $\mathbf{M}_t(L)$ is an $r \times m$ polynomial matrix in the lag operator L, which can be time dependent. We make the following assumptions about the signal process.

Assumption 2. *The polynomial matrix* $\mathbf{M}_t(L)$ *satisfies*

$$\mathbf{M}_t(L) = \sum_{\tau=0}^{\infty} \mathbf{M}_{t,\tau} L^{\tau},$$

where each element of $\{\mathbf{M}_{t,\tau}\}_{\tau=0}^{\infty}$ *is square-summable.*

The elements of $\boldsymbol{\varepsilon}_{it}$ are uncorrelated[12] *Gaussian shocks. The covariance matrix of $\boldsymbol{\varepsilon}_{it}$ is denoted by $\boldsymbol{\Sigma}^2$, where $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_m)$. Further, $\boldsymbol{\varepsilon}_{it}$ can be partitioned into two parts,*

$$\boldsymbol{\varepsilon}_{it} = \begin{bmatrix} \boldsymbol{\eta}_t \\ \boldsymbol{\nu}_{it} \end{bmatrix}, \qquad \textit{with} \qquad \int \boldsymbol{\nu}_{it} = \mathbf{0},$$

where the first $u < m$ shocks, $\boldsymbol{\eta}_t$, are common to all agents and the last $m - u$ shocks, $\boldsymbol{\nu}_{it}$, are idiosyncratic.

For now, we assume that the fundamental θ_t is driven only by the common shocks $\boldsymbol{\eta}_t$. A common-value best response is one that satisfies this condition—we relax this assumption in Section 4.3. We allow the stochastic process of the fundamental, as for the signals, to be time dependent.

Assumption 3. *The process of the exogenous fundamental θ_t is given by*

$$\theta_t = \begin{bmatrix} \boldsymbol{\phi}_t'(L) & \mathbf{0} \end{bmatrix} \boldsymbol{\varepsilon}_{it} = \boldsymbol{\phi}_t'(L)\, \boldsymbol{\eta}_t, \tag{3.4}$$

where $\boldsymbol{\phi}_t(L) = \sum_{\tau=0}^{\infty} \boldsymbol{\phi}_{t,\tau} L^{\tau}$ is a $u \times 1$ vector of polynomials in L and each element of $\{\boldsymbol{\phi}_{t,\tau}\}_{\tau=0}^{\infty}$ is square-summable.

The superscript t denotes the history up to t, for example $\mathbf{x}_i^t \equiv \{\mathbf{x}_{it}, \mathbf{x}_{it-1}, \mathbf{x}_{it-2}, \ldots\}$. Agent i's information set in period t, $\mathcal{I}_{it}$, includes the history of observed signals, and also the structure of the stochastic processes,

$$\mathcal{I}_{it} = \{\mathbf{x}_i^t, \mathbf{M}^t(L), \boldsymbol{\phi}^t(L)\}.$$

The expectation operator $\mathbb{E}_{it}[\cdot]$ stands for $\mathbb{E}[\cdot \mid \mathcal{I}_{it}]$, where $\mathbb{E}[\cdot]$ denotes the unconditional expectation. Note that $\boldsymbol{\phi}_t(L)$ and $\mathbf{M}_t(L)$ can change in a deterministic or stochastic way, and that agents only need to know their realizations up to the current period, t.[13]

The information structure we have specified above is quite general. The following examples show that signal processes commonly used in the literature are special cases of it.

1. Static Case: Suppose that $\boldsymbol{\phi}_t(L) = \boldsymbol{\phi}$ and $\mathbf{M}_t(L) = \mathbf{M}$. In this case, the past signals are not informative about the current fundamental. For example, consider

$$\boldsymbol{\phi} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}, \qquad \mathbf{M} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} x_{it}^1 &= \theta_t + \chi_t \\ x_{it}^2 &= \theta_t + \xi_t + u_{it} \end{aligned}$$

[12] The assumption that the shocks are uncorrelated is immaterial since a signal process with correlated shocks can be rewritten in terms of one with uncorrelated shocks.

[13] Also notice that the freedom to choose the primitive $\boldsymbol{\phi}_t$ in Assumption 3 already implies that the best-response function could depend on the fundamental θ_t in any period. This is because for $\theta_{t+k} = \boldsymbol{\phi}_t'(L)L^{-k}\boldsymbol{\eta}_t$, with any k, positive or negative, we can always define a new fundamental $\xi_t \equiv \left[\boldsymbol{\phi}_t'(L)L^{-k}\right]_+ \boldsymbol{\eta}_t$, so that $\mathbb{E}_{it}[\theta_{t+k}] = \mathbb{E}_{it}[\xi_t]$.

where all the shocks, $\boldsymbol{\varepsilon}_{it} = [\theta_t, \chi_t, \xi_t, u_{it}]'$, follow i.i.d. processes. This structure is similar to the specification in Angeletos and Lian (2016b), and it reduces to the case considered in Bergemann and Morris (2013) when setting the variance of ξ_t to zero.

2. Stationary Signal: Suppose that $\boldsymbol{\phi}_t(L) = \boldsymbol{\phi}(L)$ and $\mathbf{M}_t(L) = \mathbf{M}(L)$. In this case, the signal structure is time independent. For example, consider

$$\boldsymbol{\phi}(L) = \begin{bmatrix} \frac{1}{1-\rho_1 L} & 0 & 0 \end{bmatrix}, \qquad \mathbf{M}(L) = \begin{bmatrix} \frac{1}{1-\rho_1 L} & \frac{1}{1-\rho_2 L} & 0 \\ \frac{1}{1-\rho_1 L} & 0 & \frac{1}{1-\rho_3 L} \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} \theta_t &= \tfrac{1}{1-\rho_1 L}\eta_t \\ x^1_{it} &= \theta_t + \tfrac{1}{1-\rho_2 L}\xi_t \\ x^2_{it} &= \theta_t + \tfrac{1}{1-\rho_3 L}u_{it} \end{aligned}$$

with i.i.d. shocks $\boldsymbol{\varepsilon}_{it} = [\eta_t, \xi_t, u_{it}]'$. This structure allows a persistent fundamental and a persistent aggregate noise, which can be viewed as variations of the information structure adopted in Huo and Takayama (2017a) and Angeletos, Collard, and Dellas (2018).

3. Non-Stationary Signal: Since we allow $\boldsymbol{\phi}_t(L)$ and $\mathbf{M}_t(L)$ to be time dependent, the signals can be non-stationary. For example, consider

$$\boldsymbol{\phi}_t(L) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}, \qquad \mathbf{M}_t(L) = \begin{bmatrix} 1 & 0 & 1 \\ 1 & \sigma_t & 0 \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} x^1_{it} &= \theta_t + u_{it} \\ x^2_{it} &= \theta_t + \sigma_t \xi_t \end{aligned}$$

with i.i.d. shocks $\boldsymbol{\varepsilon}_{it} = [\theta_t, \xi_t, u_{it}]'$. The time-varying parameter σ_t allows the precision of x^2_{it} to change over time. This is similar to Fajgelbaum, Schaal, and Taschereau-Dumouchel (2017) where the endogenous learning effectively generates time-varying informativeness. Also, Drenik and Perez (2015) explore how higher uncertainty about inflation influences the effectiveness of monetary policy using a similar device. In general, time-varying variances or other structural parameters can be easily incorporated into $\boldsymbol{\phi}_t(L)$ and $\mathbf{M}_t(L)$.

Another type of non-stationary signal process occurs when one of the random variables has a unit root. For example, suppose that $\{\theta_{-1}, \theta_{-2}, \ldots\}$ is publicly known, and from $t = 0$ on, the fundamental process and signal structure are given by

$$\Delta\theta_t = \rho\, \Delta\theta_{t-1} + \eta_t, \qquad \text{and} \qquad x_{it} = \theta_t + u_{it}.$$

In this case, θ_t itself is not stationary if one looks forward in time, but at any given time, $\boldsymbol{\phi}_t(L)$ and $\mathbf{M}_t(L)$ are still stationary.[14] This specification is similar to the one in Woodford (2002).

[14] With θ_t known for $t < 0$, it is straightforward to see that $\boldsymbol{\phi}_t(L)$ and $\mathbf{M}_t(L)$ both represent finite moving-average processes. If θ_t is unknown for $t < 0$, we would then need to require that the past signals before $t < 0$ follow stationary processes.

Equilibrium As is standard in the dispersed information literature, we employ the Bayesian-Nash equilibrium concept. The following proposition establishes the existence, uniqueness, and linearity of the equilibrium.

Proposition 1. *Under Assumptions 1-3, there exists a unique Bayesian-Nash equilibrium and the equilibrium is linear in signals,*

$$y_{it} = \mathbf{h}_t'(L)\mathbf{x}_{it},$$

where $\mathbf{h}_t'(L)$ is a time-dependent vector of polynomial functions.[15]

Proof. See Appendix A.3. □

Heuristically, with exogenous information and with $\alpha \in (-1, 1)$, the right-hand side of equation (3.2) implies that agent i's equilibrium action is well defined and unique. Since the fundamental and signals are driven by Gaussian shocks, the optimal forecasts are always linear functions of signals.[16]

3.2 Single-Agent Problem

To define the single-agent problem corresponding to a beauty-contest model, we need to first describe the appropriately modified signal process.

Definition 2. *Given the original signal process defined in equation (3.3) with Assumptions 1 and 2 being satisfied, the* ***α-modified signal process*** *is*

$$\widetilde{\mathbf{x}}_{it} = \mathbf{M}_t\left(L\right)\widetilde{\boldsymbol{\varepsilon}}_{it},$$

where

$$\widetilde{\boldsymbol{\varepsilon}}_{it} \equiv \boldsymbol{\Gamma}\boldsymbol{\varepsilon}_{it}, \qquad \textit{and} \qquad \boldsymbol{\Gamma} \equiv \begin{bmatrix} \mathbf{I}_u & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{1-\alpha}}\mathbf{I}_{m-u} \end{bmatrix}.$$

Note that the α-modified signal process simply reduces the precision of all the idiosyncratic shocks $\boldsymbol{\nu}_{it}$ by a factor of $1-\alpha$, and leaves the precision of common shocks $\boldsymbol{\eta}_t$ unchanged. A public signal is, by definition, one that depends only on the common shocks $\boldsymbol{\eta}_t$, and, therefore, is unaffected by the modification. A private signal, on the other hand, must contain some of the idiosyncratic shocks $\boldsymbol{\nu}_{it}$ and has its precision adjusted in the α-modified signal process.

[15]The results in this proposition have all been previously established in the literature, we include them here for completeness.

[16]With endogenous information, the uniqueness is no longer guaranteed. Online Appendix H.2 provides an example with multiple linear equilibria. Furthermore, with endogenous information, we cannot exclude the possibility of nonlinear equilibria, and we restrict our attention to linear equilibria there.

We let $\widetilde{\mathbb{E}}_{it}[\cdot]$ denote the expectation operator conditional on the α-modified signal process. The single-agent problem is to forecast the economic fundamental without paying attention to other agents' actions, that is to solve for $\widetilde{\mathbb{E}}_{it}[\theta_t]$.

4 The Single-Agent Solution

In this section, we connect the beauty-contest model with the single-agent problem. We first lay out the main theoretical result, and then proceed to its several variations.

4.1 Equivalence Result

Theorem 1. *Under Assumptions 1-3, let $\mathbf{h}_t(L)$ denote the equilibrium policy rule, that is*

$$y_{it} = \mathbf{h}'_t(L)\mathbf{x}_{it}.$$

Then, the same $\mathbf{h}_t(L)$ is the forecasting rule of the fundamental θ_t conditional on the α-modified signal process,

$$\widetilde{\mathbb{E}}_{it}[\theta_t] = \mathbf{h}'_t(L)\widetilde{\mathbf{x}}_{it}.$$

Proof. See Appendix A. □

Theorem 1 characterizes the individual's policy rule concisely. The equilibrium in the beauty-contest model involves the following fixed point problem: an agent's action depends on her individual forecast about the aggregate action, but the law of motion of aggregate action is the result of individuals' forecasts. Theorem 1 shows that this fixed point problem can be avoided. The policy rule is exactly the same as the one that solves a simple forecasting problem, which can be viewed as a single agent's first-order expectation. The required transformation of the signal process takes a simple form, and this simple form is preserved across a large class of information structures.

In beauty-contest models, agents' actions depend on the forecast of the fundamental and the forecast of others' actions. The former is pinned down by a first-order expectation, and the latter by higher-order expectations. A higher α raises the relative importance of higher-order expectations and effectively shifts agents' actions towards higher-order expectations. In the corresponding single-agent problem, to induce the same policy rule, the precision of private signals has to be discounted.

The proof of Theorem 1 follows two steps: (1) we show that the equivalence holds when the fundamental and the signal processes follow finite MA(q) processes; and (2) we show that the equivalence remains true when q goes to infinity.[17] In step (1) we effectively use the method of undetermined

[17] This is where the square-summable property in Assumptions 2 and 3 become necessary.

coefficients to solve for the equilibrium policy rule, however, we are not interested on the coefficients themselves but on a property that they satisfy, namely that they are equivalent to a forecasting problem with the appropriately modified signal process. This is because solving for each coefficient becomes intractable as q goes to infinity, which is not the case for the corresponding forecasting problem. In most applications, the signal process permits a state-space representation with a recursive structure. If this is the case, the following proposition gives an explicit procedure for obtaining the equilibrium policy rule using the standard Kalman filter.

Proposition 2. *In addition to Assumptions 1-3, suppose the signal structure allows the following state-space representation*

$$\boldsymbol{\xi}_{it} = \mathbf{F}_t \boldsymbol{\xi}_{it-1} + \mathbf{Q}_t \boldsymbol{\varepsilon}_{it}, \qquad \text{with} \qquad \boldsymbol{\xi}_{i,-1} = \mathbf{0},$$

together with the observation equation and the law of motion of the fundamental,[18]

$$\mathbf{x}_{it} = \mathbf{H}_t \boldsymbol{\xi}_{it}, \qquad \text{and} \qquad \theta_t = \mathbf{G}_t \boldsymbol{\xi}_{it}.$$

Then, the optimal action is given by

$$y_{it} = \mathbf{G}_t \mathbf{z}_{it},$$

where the evolution of the sufficient statistics, $\mathbf{z}_{it}$, follows

$$\begin{aligned} \mathbf{z}_{it} &= (\mathbf{I} - \mathbf{K}_t \mathbf{H}_t) \mathbf{F}_t \mathbf{z}_{it-1} + \mathbf{K}_t \mathbf{x}_{it}, \\ \mathbf{K}_t &= \mathbf{P}_t \mathbf{H}_t' \left(\mathbf{H}_t \mathbf{P}_t \mathbf{H}_t' \right)^{-1}, \qquad (4.1) \\ \mathbf{P}_{t+1} &= \mathbf{F}_{t+1} \left(\mathbf{P}_t - \mathbf{K}_t \mathbf{H}_t \mathbf{P}_t \right) \mathbf{F}_{t+1}' + \mathbf{Q}_{t+1} \boldsymbol{\Gamma}^2 \mathbf{Q}_{t+1}'. \qquad (4.2) \end{aligned}$$

where $\mathbf{z}_{i,-1} = \mathbf{0}$ and $\mathbf{P}_0$ is the unconditional variance of $\boldsymbol{\xi}_{i,0}$ with the α-modified signal process.

Proof. The result follows directly from Theorem 1 and the time-varying Kalman filter construction in Hamilton (1994). □

This proposition allows time-varying parameters and non-stationary processes, and therefore, can be used to study transition dynamics. Readers who are familiar with the Kalman filter should recognize that the law of motion for the vector of sufficient statistics, $\mathbf{z}_{it}$, resembles that of the prior mean of

[18] It is common to write the state space as

$$\begin{aligned} \boldsymbol{\xi}_{it} &= \mathbf{F}_t \boldsymbol{\xi}_{it-1} + \mathbf{Q}_t \boldsymbol{\varepsilon}_{it}, \\ \mathbf{x}_{it} &= \mathbf{H}_t \boldsymbol{\xi}_{it} + \mathbf{R}_t \boldsymbol{\varepsilon}_{it}. \end{aligned}$$

By redefining the state variable $\mathbf{z}_{it}$, one can always include the noise in the observation equation as part of the state variable and set $\mathbf{R}_t = \mathbf{0}$.

$\boldsymbol{\xi}_{it}$. To convert the system into the α-modified signal process, the variance of the shocks is modified by $\boldsymbol{\Gamma}^2$ in equation (4.2) for the covariance matrix $\mathbf{P}_t$. This proposition states the result with a particular starting point. For a stationary process without time-varying parameters, one could set $\mathbf{P}_0$ to be the covariance matrix of the steady-state Kalman filter and obtain a time-invariant policy rule. In quantitative applications, the information structure might be complicated, but even for large scale state-space systems, the Kalman filter problem can be solved in a fast and robust way since it involves simply iterating on equations (4.1) and (4.2). The following example shows how to apply the single-agent solution.

Example 1. We return to the case introduced at the end of Section 2, where the fundamental, θ_t, is assumed to follow an AR(1) process as in (2.8). For simplicity, suppose that agents only receive a private signal about θ_t, so that the original signal and the corresponding α-modified signal are given by

$$\begin{aligned} x_{it} &= \theta_t + \nu_{it}, \qquad \nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1}), \\ \widetilde{x}_{it} &= \theta_t + \widetilde{\nu}_{it}, \qquad \widetilde{\nu}_{it} \sim \mathcal{N}(0, ((1-\alpha)\tau_\nu)^{-1}). \end{aligned}$$

Instead of worrying about higher-order expectations, one can simply turn to our single-agent solution. We start by solving the simple forecasting problem for the α-modified signal process, which can be obtained via Proposition 2,

$$\widetilde{\mathbb{E}}_{it}[\theta_t] = (1-\phi)\rho\, \widetilde{\mathbb{E}}_{it-1}[\theta_{t-1}] + \phi\, \widetilde{x}_{it}, \tag{4.3}$$

Thus, the optimal forecast is a weighted average of the prior $\rho\, \widetilde{\mathbb{E}}_{it-1}[\theta_{t-1}]$ and the new signal $\widetilde{x}_{it}$, where the weight on the signal is the Kalman gain ϕ given by

$$\phi \equiv \frac{(1-\alpha)\tau_\nu}{\kappa + (1-\alpha)\tau_\nu}, \quad \text{with} \quad \kappa \equiv \left[\rho^2 \left(\kappa + (1-\alpha)\tau_\nu\right)^{-1} + \tau_\eta^{-1}\right]^{-1}.$$

These two equations are the versions of equations (4.1) and (4.2) for the example at hand and κ is the precision of the prior. Using Theorem 1, we immediately obtain the individual policy function and the law of motion of the aggregate action in the beauty-contest model from equation (4.3),[19]

$$y_{it} = (1-\phi)\rho\, y_{it-1} + \phi\, x_{it}, \qquad \text{and} \qquad y_t = (1-\phi)\rho\, y_{t-1} + \phi\, \theta_t. \tag{4.4}$$

Recall from Figure 1, higher-order expectations have significantly more complex dynamics than first-order expectations. In contrast, the single-agent solution (4.4) is surprisingly simple. This demonstrates the usefulness our method both for computation and for equilibrium characterization.

[19] Woodford (2002) solves this model numerically by a guess-and-verify approach and Huo and Takayama (2017b) solve it analytically by a more complicated frequency domain method.

We explore the dynamic properties of beauty-contest models further in Section 5.1.

4.2 Implications for Aggregates

Theorem 1 shows that the policy rule of an individual agent in the equilibrium and in the single-agent problem is the same. The following corollary, shows that the aggregate outcome in these two problems are the same since idiosyncratic shocks wash out in the aggregate.

Corollary 1. *Under Assumptions 1-3, the aggregate outcome, y_t, in equilibrium is the same as the average expectation in the single-agent problem, i.e.,*

$$y_t = \int \widetilde{\mathbb{E}}_{it}[\theta_t] \equiv \widetilde{\mathbb{E}}_t[\theta_t].$$

This result has several direct implications. First, without precise knowledge about the strength of the strategic complementarity and about the degree of information frictions, Corollary 1 raises an identification issue. For instance, in Example 1, if the aggregate outcome can be justified by a particular α and τ_ν, there is a continuum of economies with the same product $(1-\alpha)\tau_\nu$ that imply the same aggregate outcomes. In particular, there is always a single-agent problem with no strategic interaction that yields the same allocation.

In practice, there may exist independent micro evidence that helps to pin down the strength of strategic complementarity. For example, in the monetary and the business cycle models, specified in Online Appendix E, the strategic complementarity is determined by the elasticity of substitution across varieties, the Frisch elasticity, and the degree of risk aversion, which have been estimated based on various microdata. In this case, our single-agent result can be used to infer the required degree of information frictions necessary to account for the aggregate allocation.

On the other hand, additional evidence on the degree of information frictions may be available. For example, the Survey of Professional Forecasters and the Michigan Survey of Consumers contain information about the magnitude and persistence of forecast errors, which can be used to discipline information frictions. In this case, with knowledge also on information frictions, our equivalence result can help quantify its role in shaping aggregate fluctuations.[20]

Finally, Corollary 1 does not imply the equivalence between models in terms of cross-sectional allocation. In fact, the individual choices are different,

$$y_{it} = \mathbf{h}_t'(L)\mathbf{x}_{it} \neq \mathbf{h}_t'(L)\widetilde{\mathbf{x}}_{it} = \widetilde{\mathbb{E}}_{it}[\theta_t].$$

[20] See, for instance, Huo and Takayama (2017a), Drenik and Perez (2015), Melosi (2016), and Angeletos and Huo (2018).

In particular, to generate the same aggregate allocation, the signals in the single-agent model are more dispersed, which translates into larger forecast errors and action dispersion. This difference can help achieve identification when cross-sectional moments are available.

4.3 Generalized Best Response

So far, we have considered the classic beauty-contest game where agents' best responses depend only on a common fundamental, and agents only choose a single action each period. In many relevant environments, however, agents' payoffs depend on idiosyncratic factors, and agents choose more than one action at the same time. In this section, we show how the single-agent solution extends to these more general beauty-contest models.

We start by providing a corollary of Theorem 1 that relates higher-order expectations with the first-order expectation in the single-agent model.

Corollary 2. *Under Assumptions 2-3, for any $\alpha \in \mathbb{Z}$ such that $|\alpha| < 1$, the forecasting rule of a geometric sum of infinite higher-order expectations about θ_t, is the same as that of the single-agent problem*[21]

$$(1-\alpha)\sum_{k=0}^{\infty}\alpha^k \, \mathbb{E}_{it}\overline{\mathbb{E}}_t^k[\theta_t] = \mathbf{h}_t'(L)\mathbf{x}_{it},$$

$$\widetilde{\mathbb{E}}_{it}[\theta_t] = \mathbf{h}_t'(L)\widetilde{\mathbf{x}}_{it}.$$

Proof. Given Theorem 1, this follows straightforwardly from equation (3.2). □

Corollary 2 is a result about linear projections with Gaussian signals and can be used independently of any equilibrium concept. It says that a geometric sum of infinite higher-order expectations is equivalent to a particular first-order expectation. As already mentioned, expectations of higher order typically become more complicated in the sense that more state variables are required to forecast those higher-order expectations. If one wants to compute each of the infinite higher-order expectations independently, then an infinite number of state variables are needed. Corollary 2 shows that a first-order expectation with a slightly modified signal process yields the same result. This result turns out to be useful in extending our single-agent solution to more general best responses.

[21] If α is a complex number, the modified precisions are complex numbers as well. In this case, $\widetilde{\mathbb{E}}_{it}[\cdot]$ only stands for the operation required to conduct a forecast.

Independent-Value Best Response Consider a best response function in which the relevant fundamental contains idiosyncratic factors

$$y_{it} = \gamma \mathbb{E}_{it}[\theta_{it}] + \alpha \mathbb{E}_{it}[y_t], \tag{4.5}$$

which is commonly referred as independent-value best response. For this type of best response function, the original single-agent solution does not apply exactly, but a slightly modified version does.

Proposition 3. *Suppose that Assumptions 1-3 are satisfied.*[22] *Denote the forecast about the fundamental by*

$$\mathbb{E}_{it}[\theta_{it}] \equiv \mathbf{g}'_t(L)\mathbf{x}_{it} \equiv \varphi_t + \xi_{it}.$$

where ξ_{it} is only driven by idiosyncratic shocks. There exists a unique equilibrium policy rule $\mathbf{h}_t(L)$ such that

$$y_{it} = \gamma \mathbb{E}_{it}[\theta_{it}] + \alpha\ \mathbb{E}_{it}[y_t] = \gamma(\varphi_t + \xi_{it}) + \frac{\alpha\gamma}{1-\alpha}\mathbf{h}'_t(L)\mathbf{x}_{it},$$

and $\mathbf{h}_t(L)$ is the same as the forecasting rule of φ_t conditional on the α-modified signal process,

$$\widetilde{\mathbb{E}}_{it}[\varphi_t] = \mathbf{h}'_t(L)\widetilde{\mathbf{x}}_{it}.$$

Proof. See Appendix D.1. □

Multiple Actions Next, suppose that agents make multiple choices, each of which may depend on multiple choices made by other agents in its own way. For example, an employer needs to decide not only how many workers to hire, but also the number of working hours and the amount of effort each worker needs to exert. A firm may decide simultaneously the level of production and the amount of resources to spend on advertising. In these situations, the best response becomes a multivariate system for a vector of actions, and a matrix instead of a number summarizes the primitive motive for strategic interaction among agents. We show that a similar, though slightly more complicated, single-agent solution holds in this type of environment.

Consider the environment described in Section 3, but with the following more general best response function which allows for multiple actions depending on multiple fundamentals:

$$\mathbf{y}_{it} = \mathbb{E}_{it}\left[\boldsymbol{\theta}_t\right] + \mathbf{A}\mathbb{E}_{it}[\mathbf{y}_t], \qquad \text{with} \qquad \mathbf{y}_t = \int \mathbf{y}_{it}, \tag{4.6}$$

where $\boldsymbol{\theta}_t = [\theta_{1t}, \theta_{2t}, \ldots, \theta_{nt}]'$, $\mathbf{y}_{it}$, and $\mathbf{y}_t$ are $n \times 1$ vectors, and $\mathbf{A}$ is an $n \times n$ matrix. We denote the

[22] More precisely, the fundamental depends on both the aggregate and idiosyncratic shocks, $\theta_{it} = \sum_{\tau=0}^{\infty} \boldsymbol{\psi}_{t,\tau} L^{\tau} \boldsymbol{\varepsilon}_{it}$ and each element of $\{\boldsymbol{\psi}_{t,\tau}\}_{\tau=0}^{\infty}$ must be square-summable.

eigenvalues of $\mathbf{A}$ by $\boldsymbol{\alpha} = \{\alpha_j\}_{j=1}^n$ and, in what follows, assume that:

Assumption 4. *The matrix* $\mathbf{A}$ *is diagonalizable*

$$\mathbf{A} = \mathbf{Q}\boldsymbol{\Omega}\mathbf{Q}^{-1}, \quad \textit{and} \quad \boldsymbol{\Omega} = \mathrm{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n),$$

and the absolute value of all of its eigenvalues is less than one, that is $|\alpha_j| < 1$ *for all* j.

Since there are multiple degrees of strategic complementarity, the policy function in the equilibrium of the multi-variate beauty-contest model is equivalent, not to that of the simple forecasting with one α-modified signal process, but a particular linear combination of all α_j-modified signal processes. We denote the analogues of $\widetilde{\mathbb{E}}_{it}[\cdot]$ and $\widetilde{\mathbf{x}}_{it}$, with $\alpha = \alpha_j$, by $\widetilde{\mathbb{E}}_{it}[\,\cdot\,;\alpha_j]$ and $\widetilde{\mathbf{x}}_{it}(\alpha_j)$. The coefficients of this linear combination can be obtained directly from the eigenvectors of $\mathbf{A}$. The following proposition is a natural extension of Theorem 1.

Proposition 4. *Suppose Assumptions 2-3 and 4 are satisfied. Let* $\mathbf{g}_{jt}(L)$ *be the policy rule to forecast* $\boldsymbol{\theta}_t$ *conditional on the* α_j*-modified signal process,*

$$\widetilde{\mathbb{E}}_{it}[\boldsymbol{\theta}_t;\alpha_j] \equiv \mathbf{g}'_{jt}(L)\widetilde{\mathbf{x}}_{it}(\alpha_j),$$

Then, the equilibrium policy rule $\mathbf{h}_t(L)$ *is given by*

$$\mathbf{h}'_t(L) = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}'_j\mathbf{Q}^{-1}\left(\mathbf{I}-\mathbf{A}\right)^{-1}\mathbf{g}'_{jt}(L),$$

where $\mathbf{e}_j$ *denotes the* j*-th column of an* $n \times n$ *identity matrix.*

Proof. See Appendix D.2. □

Proposition 4 provides a sharp characterization of the equilibrium. Crucially, the policy rules for all n actions are based on the same set of modified signals. The heterogeneity between these actions is only reflected in the associated weights, which are obtained from the eigenvectors of $\mathbf{A}$. The analysis of the equilibrium can be reduced to the model's primitives: the matrix $\mathbf{A}$ and the exogenous forecasts associated with the modified signals, again without the need to solve a fixed point problem.

The equivalence in Corollary 2 hinges on a special geometric weighting structure of higher-order expectations. An implication of Proposition 4 is that this equivalence can be generalized to much more flexible weighting structures.

Corollary 3. *Under Assumptions 2-3 and 4,*

$$\sum_{k=0}^{\infty} \mathbf{A}^k \boldsymbol{\varphi}\, \overline{\mathbb{E}}_t^{k+1}[\theta_t] = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\mathbf{Q}^{-1}(\mathbf{I}-\mathbf{A})^{-1}\boldsymbol{\varphi}\, \widetilde{\mathbb{E}}_t[\theta_t;\alpha_j], \tag{4.7}$$

where $\boldsymbol{\varphi}$ is an arbitrary row vector.

On the left-hand side of equation (4.7), each row is a weighted sum of higher-order expectations. Different $\mathbf{A}$'s and $\boldsymbol{\varphi}$'s imply different weighting structures, which can lead to complex combinations of higher-order expectations. Despite this complexity, they are always equal to a weighted sum of a finite number of first-order expectations.

5 Applications

In this section, we discuss several applications that utilize our single-agent solution to characterize and understand the equilibrium of beauty-contest models with persistent information. Some of the applications are simple in the interest of clarity, and it should be evident from the exposition above that they, by no means, exhaust the environments in which the results are applicable.[23]

5.1 Inertia and Sentiments

When information is persistent and dispersed, the aggregate outcome can display rich dynamics, which helps match some salient features of the data. For example, one empirical regularity is that aggregate variables (such as output and inflation) respond to underlying shocks in a sluggish way relative to what perfect information models predict. At the same time, higher-order expectations are more anchored by common priors and often display inertia. In beauty contest models, the aggregate outcome partially inherits properties of higher-order expectations, and therefore, can be more persistent than its perfect information counterpart. Another popular theme is that information frictions break the tight connection between the fundamental and the allocation, and rationalize fluctuations driven by *sentiments* or *animal-spirits*.[24] As mentioned earlier, though dispersed information is crucial in understanding the propagation mechanism of these fluctuations, solving and especially characterizing these models is typically difficult. The single-agent solution allows us to characterize the dynamic properties of the equilibrium outcome under a general class of information structures.

Consider a setup with the best response given by equation (3.1) and signals that follow a stationary process and permit a state-space representation as in Proposition 2,

$$\boldsymbol{\xi}_{it} = \mathbf{F}\boldsymbol{\xi}_{it-1} + \mathbf{Q}\boldsymbol{\varepsilon}_{it}, \quad \text{with} \quad \mathbf{x}_{it} = \mathbf{H}\boldsymbol{\xi}_{it} \quad \text{and} \quad \theta_t = \mathbf{G}\boldsymbol{\xi}_{it}.$$

[23] We provide an example how to use the multi-action solution in Online Appendix H.1.

[24] See Lorenzoni (2009), Barsky and Sims (2012), and Angeletos and La'O (2010) among many others.

We assume, further, that the signal process satisfies

$$\int \mathbf{x}_{it} = \mathbf{A}\theta_t + \mathbf{B}(L)\mathbf{v}_t,$$

for some matrices $\mathbf{A}$ and $\mathbf{B}(L)$. The assumption that $\mathbf{A}$ is constant helps isolate non-fundamental driven fluctuations. In the equation above, $\mathbf{v}_t$ represents aggregate shocks that are orthogonal to the fundamental θ_t, but can shift agents' expectations about it. Therefore, they may be interpreted as sentiments or animal spirits. The information structure we have specified above includes most signal processes used in the literature.

Proposition 5. *The aggregate outcome is given by*

$$y_t = \mathbf{C}(L)\big(\mathbf{A}\theta_t + \mathbf{B}(L)\mathbf{v}_t\big),$$

where

$$\mathbf{C}(L) \equiv \frac{\mathbf{G}\ \mathrm{adj}(\mathbf{I} - (\mathbf{F} - \mathbf{KHF})L)\ \mathbf{K}}{\Pi_{k=1}^{\ell}(1 - \lambda_k L)},$$

$\mathbf{K}$ *is the steady-state Kalman gain under the* α*-modified signal process, and* $\{\lambda_k\}_{k=1}^{\ell}$ *are the non-zero eigenvalues of* $\mathbf{F} - \mathbf{KHF}$.

Proof. See Appendix B. □

This result sheds light on how information frictions can shape the dynamics of aggregate outcomes. In the literature, Woodford (2002) and Angeletos and La'O (2010) use a guess-and-verify approach to solve their models—with relatively simple signal processes—numerically. In contrast, a direct application of the single-agent solution provides a formula that characterizes the dynamics for a large class of signal and fundamental processes.

Consider first the response to a shock to the fundamental. Relative to θ_t itself, the aggregate outcome exhibits additional dynamics captured by the term $\mathbf{C}(L)$. Generically, this extra term adds more persistence, since relative to the fundamental, θ_t, the current outcome, y_t, is further anchored by its past realizations. Recall Example 1, where the aggregate outcome in equation (4.4) depends on y_{t-1} besides θ_t, and this connection is intensified with higher complementarity or more information frictions. Proposition 5 formalizes this idea for a general class of information processes. The linkage with past outcomes is precisely determined by the eigenvalues of $\mathbf{F} - \mathbf{KHF}$.

Next, consider the response to sentiment shocks, $\mathbf{v}_t$. For simplicity, focus on the case in which $\mathbf{B}(L) = \mathbf{B}$. The sentiment shocks inherit the same dynamic component as the fundamental, $\mathbf{C}(L)$. Therefore, even when the sentiment shocks are not persistent themselves, they can have prolonged effects on the aggregate outcome. The persistence of sentiments explored in Angeletos and La'O (2010) and Huo and Takayama (2017a) can be viewed as special cases of this proposition.

As a final remark, the additional dynamics of the aggregate outcome relative to the fundamental can arise from the interaction between the persistence of the fundamental with noise, or from the persistence of noise itself that confounds the learning about the fundamental. In an extreme case in which the private noise shocks are correlated over time, aggregate outcomes may display persistent fluctuations even when all the aggregate shocks are i.i.d.

5.2 Comparative Statics, Volatility, and GE Effects

In this sub-section, we show that the single-agent solution also facilitates obtaining predictions about the volatility of the aggregate outcome and its covariance with fundamental. These results are derived without imposing detailed assumptions on the information structure, and are, therefore, robust in the sense used by Bergemann and Morris (2013) and Bergemann, Heumann, and Morris (2015).[25]

We consider two types of best response functions. The first type is the common-value best response (CVBR), as described in equation (3.1). The second is a particular widely used independent-value best response (IVBR), in which agents observe their individual fundamental, θ_{it}, perfectly,

$$y_{it} = (1-\alpha)\,\theta_{it} + \alpha\,\mathbb{E}_{it}[y_t], \tag{5.1}$$

but are unsure about the aggregate fundamental, given by $\theta_t \equiv \int \theta_{it}$. The assumption that θ_{it} is perfectly observable is a natural one if agents optimally allocate their limited cognitive resources.[26] In this economy, the first-order uncertainty about their own payoff relevant fundamental is muted, but higher-order uncertainty about the aggregate fundamental and outcome remains.

We establish the following comparative statics results.

Proposition 6. *Suppose that Assumptions 1-3 are satisfied:*

1. *For CVBR with $\alpha \in (-1,1)$ and IVBR with $\alpha > 0$, the volatility of the aggregate outcome and its covariance with the fundamental are decreasing in the degree of strategic complementarity,*

$$\frac{\partial \mathrm{Var}(y_t)}{\partial \alpha} \leq 0, \quad \textit{and} \quad \frac{\partial \mathrm{Cov}(y_t,\theta_t)}{\partial \alpha} \leq 0.$$

[25] Note that the beauty-contest model considered in this paper satisfies the 'global stability under uncertainty' condition specified in Weinstein and Yildiz (2007b), so that the maximum change in equilibrium actions, due to changes in beliefs at orders higher than k, is exponentially decreasing in k. This property makes the robust predictions possible. We thank Muhamet Yildiz and an anonymous referee for pointing this out.

[26] As in Maćkowiak and Wiederholt (2009), agents rationally choose to pay more attention to their individual fundamental shocks than to the aggregate shocks.

2. *For CVBR, if $\alpha > 0$ ($\alpha < 0$) the volatility of the aggregate outcome is lower (higher) than the volatility of the average forecast of the fundamental,*

$$\mathrm{Var}\,(y_t) \leq \mathrm{Var}\left(\overline{\mathbb{E}}_t(\theta_t)\right), \quad \textit{if } \alpha > 0, \quad \textit{and} \quad \mathrm{Var}\,(y_t) \geq \mathrm{Var}\left(\overline{\mathbb{E}}_t(\theta_t)\right), \quad \textit{if } \alpha < 0.$$

3. *For IVBR if $\alpha > 0$ ($\alpha < 0$) the volatility of the aggregate outcome is lower (higher) than the volatility of the fundamental,*

$$\mathrm{Var}(y_t) \leq \mathrm{Var}(\theta_t), \quad \textit{if } \alpha > 0, \quad \textit{and} \quad \mathrm{Var}(y_t) \geq \mathrm{Var}(\theta_t), \quad \textit{if } \alpha < 0.$$

Proof. See Appendix C. □

The results for the CVBR can be understood using the single-agent solution, which implies that $y_t = \widetilde{\mathbb{E}}_t[\theta_t]$. With the modified signal process as α increases, the precision of private signals decreases, and the forecast is less accurate. Therefore, the volatility of the average forecast decreases since agents become less responsive to signals, and the covariance between the fundamental and its forecast declines as well. Part 2 of the proposition follows directly from Part 1 by noticing that $y_t = \overline{\mathbb{E}}_t[\theta_t]$ when $\alpha = 0$.

Next, consider the IVBR case. From Proposition 3, the aggregate outcome is given by

$$y_t = (1 - \alpha)\,\theta_t + \alpha\,\widetilde{\mathbb{E}}_t[\theta_t]. \tag{5.2}$$

The first term $(1 - \alpha)\,\theta_t$ is associated with the agents' response to their own fundamentals, which can be interpreted as partial equilibrium (PE) effects, and the second term is the agents' response to others' actions which can be interpreted as general equilibrium (GE) effects.[27] A larger strategic complementarity implies noisier private signals in the α-modified signal process, which attenuates the response of $\widetilde{\mathbb{E}}_t[\theta_t]$ to changes in the fundamental. When α is positive, an increase in α also increases the weight of the less responsive GE effects, $\widetilde{\mathbb{E}}_t[\theta_t]$, relative to the fundamental, θ_t, in equation (5.2). This explains why the variance of the aggregate outcome and its covariance with the fundamental decrease with α.

Part 3 for $\alpha > 0$ follows from the fact that $y_t = \theta_t$ when $\alpha = 0$ and from the result in the Part 1. For strategic substitutes, $\alpha < 0$, information frictions amplify the volatility of the aggregate outcome. To see why, rearrange equation (5.2) to obtain

$$y_t = \theta_t - \alpha(\theta_t - \widetilde{\mathbb{E}}_t[\theta_t]).$$

[27] For a more detailed discussion of this interpretation, see Angeletos and Lian (2016a).

Loosely speaking, θ_t is more responsive than the expectation of θ_t. With a negative α, it follows that the aggregate outcome also becomes more responsive than the fundamental itself.[28]

Bergemann and Morris (2013) and Bergemann, Heumann, and Morris (2015) characterize comparative statics about the maximum volatility of the aggregate outcome. Exploiting properties of correlated equilibria, their approach can be used to identify the boundary of the set of equilibria and how it changes with the degree of strategic complementarity. Our results are substantially different from theirs. First, in terms of comparative statics, we fix the underlying information structure when varying α while the information structure that achieves the maximum volatility may vary with α in those papers. Second, we compare the volatility of the aggregate outcome with that of the fundamental and its conditional expectation. Without the single-agent solution, analyzing this type of equilibrium properties would require a case-by-case study for the information structure at hand. Using our result, this task boils down to analyzing statistical properties of a simple forecasting problem, which allows the general statements made above.

5.3 Social Value of Public Information in Transition

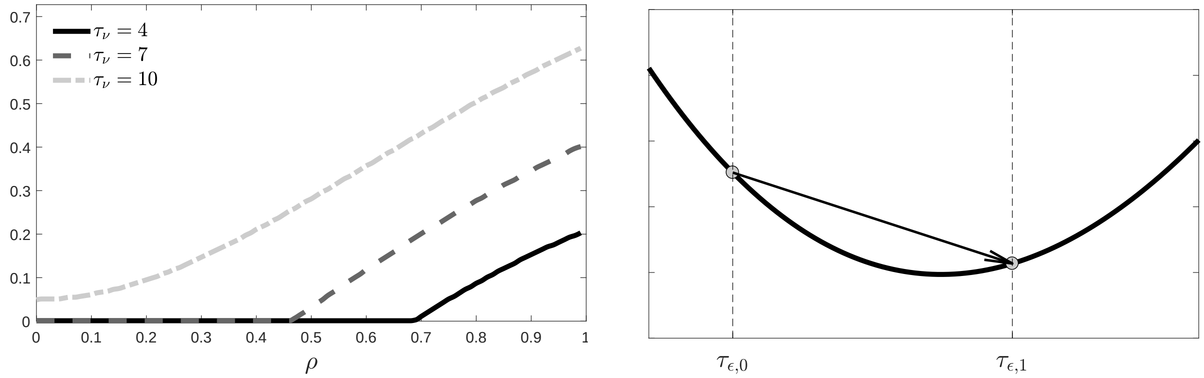
In this section, we explore the transition dynamics of social welfare following a change in the precision of public information. An interesting point made by Morris and Shin (2002) is that increasing the precision of public information can reduce social welfare. Angeletos and Pavan (2007) discuss the social value of information with a more general payoff structure. These studies focus on the static case, while here we consider a dynamic fundamental. We first show that the result from Morris and Shin (2002) is strengthened when the fundamental is more persistent. However, this is based on a comparison between steady states, which can be misleading in dynamic settings. To make this point clear, we provide an example in which, even though increasing the precision of public information is eventually detrimental to welfare, welfare is improved for a long time during transition.

Consider a fundamental, θ_t, that follows an AR(1) process as in (2.8), and suppose that agents observe a public and a private signal about it. Purposely, we allow the precision of the public signal to change over time,

$$z_t = \theta_t + \varepsilon_t, \quad \text{and} \quad x_{it} = \theta_t + \nu_{it},$$

where $\varepsilon_t \sim \mathcal{N}(0, \tau_{\varepsilon,t}^{-1})$, and $\nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1})$. Before moving to the transition dynamics, we first discuss the properties of the social value of information in the stationary case, i.e., with $\tau_{\varepsilon,t} = \tau_\varepsilon$. We assume that the agents' best response is given by equation (3.1), and that the (period) social welfare function is $W_t \equiv -\int (y_{it} - \theta_t)^2$, so that it is optimal to keep all agents' actions as close as

[28] For the CVBR, Angeletos and Lian (2016b) prove Part 1 of Proposition 6 under a particular static signal structure. We generalize the signal process to allow for a larger class of dynamic information structures. With respect to Part 2, Angeletos and Lian (2016b) show that $\text{Var}(y_t) \leq \text{Var}(\theta_t)$, while we obtain a tighter bound by comparing $\text{Var}(y_t)$ with $\text{Var}\left(\overline{\mathbb{E}}_t(\theta_t)\right)$. In addition, we also explore properties of the IVBR, whereas that paper does not.



(A) $\widehat{\tau}_\varepsilon$ for Different ρ (B) Steady State Welfare for Different τ_ε

FIGURE 2: Welfare Change in Steady State and Transition

Parameters (A & B): $\tau_\eta = 1$, and $\alpha = 0.85$; (B): $\tau_\nu = 0.3$, and $\rho = 0.95$.

possible to the fundamental, θ_t. Using the equivalence result, it follows from Proposition 2 that we can solve for the steady-state welfare analytically.[29]

To investigate under which conditions increasing the level of public information precision can reduce steady state welfare, let $\widehat{\tau}_\varepsilon$ be the precision that minimizes welfare,

$$\widehat{\tau}_\varepsilon \equiv \min_{\tau_\varepsilon} W.$$

It is straightforward to verify that $\partial W / \partial \tau_\varepsilon < 0$ if and only if $\tau_\varepsilon \in [0, \widehat{\tau}_\varepsilon)$. Figure 2A shows how $\widehat{\tau}_\varepsilon$ changes with the persistence ρ. Interestingly, as ρ increases, the welfare-reducing region becomes larger.

However, this result can be misleading. When signals are persistent, following a change to the precision of public information, the distribution of agents' priors does not change immediately, and the economy can experience a relatively long transition before it converges to the new steady state. During this process, the dynamics of social welfare can reverse the steady state comparative statics. Figure 2B shows welfare in the steady state for various values of τ_ε. The U-shaped curve indicates the existence of a welfare-reducing region. In our experiment, there is a permanent increase of the precision of public signal at time $t = 0$. As depicted in the figure, an increase in τ_ε induces a reduction in steady-state welfare. However, if we look at the transition path towards the new steady state, the

[29]The formula for the steady-state welfare is given by

$$W = \frac{\tau_\eta((1-\alpha)^2 \tau_\nu + \tau_\varepsilon) + \kappa^2}{(\tau_\varepsilon + (1-\alpha)\tau_\nu)\left((\tau_\varepsilon + (1-\alpha)\tau_\nu) + 2\kappa\right)\tau_\eta + (1-\rho^2)\tau_\eta \kappa^2}, \quad \text{with} \quad \kappa = \frac{\tau_\eta\left(\kappa + (\tau_\varepsilon + (1-\alpha)\tau_\nu)\right)}{\kappa + (\rho^2 \tau_\eta + \tau_\varepsilon + (1-\alpha)\tau_\nu)}.$$

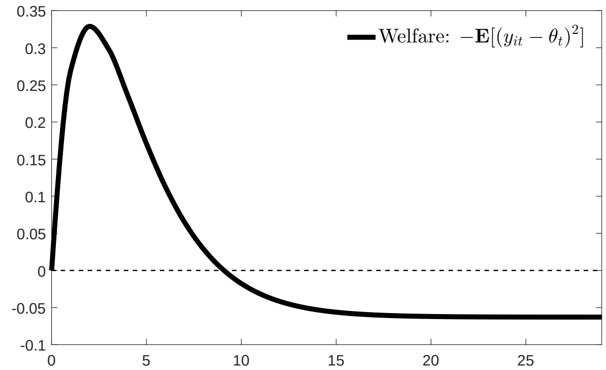


FIGURE 3: Welfare in Transition

Parameters: $\tau_\eta = 1$, $\alpha = 0.85$, $\tau_\nu = 0.3$, $\rho = 0.95$, $\tau_{\varepsilon,0} = 0.001$, and $\tau_{\varepsilon,t} = 0.004$, for $t \geq 1$.

message is quite different. Figure 3 shows how welfare changes over time. For many periods welfare is actually improved, and the magnitude of the improvement is larger than the eventual reduction.

Social welfare can be decomposed into

$$W_t = -\Bigg(\underbrace{v_t}_{\text{Var}(y_{it})} - \underbrace{2c_t}_{\text{Cov}(y_{it},\theta_t)} + \underbrace{\omega}_{\text{Var}(\theta_t)} \Bigg).$$

The variance of θ_t is invariant, $\omega \equiv 1/(1-\rho^2)\tau_\eta$, and, using Proposition 2, the dynamics of the other components of welfare can be written recursively as

$$v_t = \frac{\kappa_t^2}{(\tau_t+\kappa_t)^2}\rho^2 v_{t-1} + \frac{2\tau_t\kappa_t}{(\tau_t+\kappa_t)^2}\rho^2 c_{t-1} + \frac{\tau_t^2}{(\tau_t+\kappa_t)^2}\omega + \frac{\tau_t - \alpha(1-\alpha)\tau_\nu}{(\tau_t+\kappa_t)^2}, \tag{5.3}$$

$$c_t = \frac{\kappa_t}{\tau_t+\kappa_t}\rho^2 c_{t-1} + \frac{\tau_t}{\tau_t+\kappa_t}\omega, \tag{5.4}$$

$$\kappa_{t+1} = \frac{\tau_\eta(\kappa_t+\tau_{t+1})}{\kappa_t+\tau_{t+1}+\tau_\eta\rho^2}, \tag{5.5}$$

with the trigger of the transition coming from the time-varying precision $\tau_t \equiv (1-\alpha)\tau_\nu + \tau_{\varepsilon,t}$. The system (5.3-5.5) captures the slow-moving nature of the social welfare. Agents' forecasts rely on their prior precision about the aggregate state (5.5), which is updated gradually due to Bayesian learning. The variance of y_{it} and its covariance with the fundamental depend on the precision of the prior, and thus exhibit persistence as well. The different updating speeds of v_t and c_t allows welfare to increase initially and decrease later on. The takeaway from this example is that the results derived from a static setting may not readily extend to a dynamic setting since the slow-moving nature of agents' behavior can induce non-trivial dynamics.

6 Endogenous Information

So far, we have assumed that the signals follow exogenous stochastic processes, that is, that they are independent of agents' actions in equilibrium. In this section, we show that the single-agent solution remains valid when the source of information is endogenous. To better simplify the exposition, we focus here on stationary signal processes.

To include the possibility that agents can learn from endogenous aggregate variables, we consider signal structures with the following form

$$\mathbf{x}_{it} = \mathbf{M}(L)\boldsymbol{\varepsilon}_{it} + \mathbf{p}(L)y_t. \tag{6.1}$$

The first part $\mathbf{M}(L)\boldsymbol{\varepsilon}_{it}$ captures the information that is exogenously specified, and the second part $\mathbf{p}(L)y_t$ allows the signal to be a function of the aggregate action y_t. The informativeness of the signal depends on the process of y_t, which is determined endogenously in equilibrium.[30] A natural requirement is that agents can only learn from current and past aggregate actions, that is, $\mathbf{p}(L)$ is a one-sided square-summable polynomial vector.

With endogenous signals, the equilibrium imposes an additional consistency requirement that the law of motion for the aggregate action in the signals is consistent with that implied by agents' actions. We define the equilibrium in a way that highlights this requirement.

Definition 3. *A linear Bayesian-Nash* ***equilibrium*** *with endogenous information is a policy rule* $\mathbf{h}(L)$ *for agents and a law of motion* $\boldsymbol{\mathcal{H}}(L)$ *for the aggregate action, such that*

1. *The individual action* $y_{it} = \mathbf{h}'(L)\mathbf{x}_{it}$ *satisfies the best response (3.1), and the signal is given by the following exogenous process:*

$$\mathbf{x}_{it} = \mathbf{M}(L)\boldsymbol{\varepsilon}_{it} + \mathbf{p}(L)\ \boldsymbol{\mathcal{H}}'(L)\boldsymbol{\eta}_t. \tag{6.2}$$

2. *The aggregate action is consistent with individual actions:* $y_t = \int y_{it}$.

3. *The aggregate action is consistent with agents' signals:* $y_t = \boldsymbol{\mathcal{H}}'(L)\boldsymbol{\eta}_t$.

Condition 3 distinguishes an endogenous from an exogenous information equilibrium. It implies,

[30] Note that, in our setup, only the aggregate action enters the agents' information sets and all agents behave competitively. As a result, agents do not take into account the effect that their action might have on others' information sets.

together with Condition 1, the cross-equation restriction

$$\boldsymbol{\mathcal{H}}'(L) = \mathbf{h}'(L)\left(\mathbf{M}(L)\begin{bmatrix}\mathbf{I}_u \\ \mathbf{0}\end{bmatrix} + \mathbf{p}(L)\boldsymbol{\mathcal{H}}'(L)\right), \tag{6.3}$$

which is clearly a fixed point problem for $\boldsymbol{\mathcal{H}}(L)$. We want to emphasize that given any $\boldsymbol{\mathcal{H}}(L)$, the signal structure of an individual agent is well defined by equation (6.2), and, for Conditions 1 and 2 to be satisfied, it must be an exogenous information equilibrium from Definition 1. Therefore, all of our results from Section 4 still apply. If there exists a particular $\boldsymbol{\mathcal{H}}(L)$ that also satisfies Condition 3, then it is an endogenous information equilibrium. We formalize this argument in the following proposition:

Proposition 7. *Under Assumptions 1-3, if $\mathbf{h}(L)$ and $\boldsymbol{\mathcal{H}}(L)$ are an equilibrium from Definition 3, then the policy rule $\mathbf{h}(L)$ satisfies*

$$\widetilde{\mathbb{E}}_{it}[\theta_t] = \mathbf{h}'(L)\widetilde{\mathbf{x}}_{it}.$$

where $\widetilde{\mathbf{x}}_{it}$ the α-modified signal process of $\mathbf{x}_{it}$, and the law of motion of the aggregate action $\boldsymbol{\mathcal{H}}(L)$ satisfies

$$y_t = \boldsymbol{\mathcal{H}}'(L)\boldsymbol{\eta}_t = \widetilde{\mathbb{E}}_t[\theta_t].$$

The proof of this proposition follows from the observation that, if an endogenous information equilibrium exists, each individual agent still regards their information process as exogenous.[31] This is due to the fact that all agents behave competitively and do not take into account the effects of their action on the aggregate outcome. Hence, the policy rule is still the forecasting rule of the fundamental with modified signals.

When computing these types of models, obtaining the signal process requires a recursive algorithm that solves an exogenous information equilibrium in each iteration. This step can be efficiently implemented via the single-agent solution. By the same logic, our single-agent solution can also be extended to environments with an information acquisition choice (Hellwig and Veldkamp, 2009; Maćkowiak and Wiederholt, 2009). Once all agents have chosen the precision of their signals in equilibrium, they effectively receive signals as if the process for the signal is exogenously specified.[32]

7 Extensions

As mentioned in Section 3, the type of best response functions we have considered so far excludes two possibilities: (1) heterogeneity among agents' payoff and information structures; and (2) intertemporal strategic complementarity. Though our single-agent solution does not apply exactly to

[31] Proposition 7 does not establish the existence of an equilibrium, but a property that an equilibrium must satisfy if it exists.

[32] We provide an example to how to use Proposition 7 in Online Appendix H.2.

these types of best response functions, we show that the basic insight still works for some relevant applications.

7.1 Network: Heterogeneous Payoff and Information Structure

An assumption that has been maintained throughout the paper, up to this point, is that of *ex-ante* symmetry between the agents. Though agents may experience different private shocks over time, *ex ante* they have the same information structure and best response functions. In this section, we extend our result to asymmetric settings where there are agents of different types. These asymmetries lead to a network setup where agents depend on each other in different ways, and the information structure can also differ from agent to agent.

Suppose that there are n different agents in the economy.[33] Each agent has its own fundamental and the agents' actions can depend on each of the other agents actions. Abusing notation, we denote $\mathbb{E}_t = [\mathbb{E}_{1t}, \cdots, \mathbb{E}_{nt}]'$. Then, the best response function can be written as[34]

$$\mathbf{y}_t = \mathbb{E}_t[\boldsymbol{\theta}_t] + \mathbb{E}_t[\mathbf{W}\mathbf{y}_t], \tag{7.1}$$

where $\boldsymbol{\theta}_t$, and $\mathbf{y}_t$ are $n \times 1$ vectors, and $\mathbf{W}$ is an $n \times n$ matrix that summarizes the cross dependence of agents' payoffs on others' actions, we call this the *payoff matrix*. The heterogeneity in the payoff structure could be interpreted as the network structure in a network game. Moreover, one can also relate this framework to macroeconomic models in which multiple groups of agents (such as firms and households) with heterogeneous information interact with each other, as in Maćkowiak and Wiederholt (2015) and Angeletos and Lian (2018).

With perfect information, the equilibrium outcome would be given by

$$\mathbf{y}_t = (\mathbf{I} - \mathbf{W})^{-1} \boldsymbol{\theta}_t. \tag{7.2}$$

The matrix $(\mathbf{I} - \mathbf{W})^{-1}$ is related to Katz-Bonacich centrality of the network, which appear in the characterizations of equilibrium in De Martí and Zenou (2015) and Bergemann, Heumann, and Morris (2017). We are interested in how this formula changes when information is incomplete.

[33]It can also be interpreted in a way that the economy has n group of agents, and there are a large number of identical agents in each group.

[34]In more detail,

$$\begin{bmatrix} y_{1t} \\ \vdots \\ y_{nt} \end{bmatrix} = \begin{bmatrix} \mathbb{E}_{1t}[\theta_{1t}] \\ \vdots \\ \mathbb{E}_{nt}[\theta_{nt}] \end{bmatrix} + \begin{bmatrix} \sum_{j\neq 1} \mathbb{E}_{1t}[W_{1j} y_{jt}] \\ \vdots \\ \sum_{j\neq n} \mathbb{E}_{nt}[W_{nj} y_{jt}]] \end{bmatrix}.$$

Information In period t, agent i observes a vector of signals[35]

$$\mathbf{x}_{it} = \mathbf{M}(L)\boldsymbol{\varepsilon}_{it}, \tag{7.3}$$

where $\mathbf{M}(L)$ is common to all agents. The heterogeneous information structure is captured in the covariance matrix of $\boldsymbol{\varepsilon}_{it}$.

Assumption 5. *The elements of $\boldsymbol{\varepsilon}_{it}$ are uncorrelated normal shocks. The covariance matrix of $\boldsymbol{\varepsilon}_{it}$ is denoted by $\boldsymbol{\Sigma}_i^2$, where $\boldsymbol{\Sigma}_i^2 = \operatorname{diag}\left(\tau_1^{-1}, \ldots, \tau_{m-1}^{-1},\ \gamma_i^{-1}\tau_m^{-1}\right)$, with the first $m-1$ shocks being common shocks and only the last one being private.*[36] *Otherwise, $\mathbf{M}(L)$ satisfies Assumption 2 and each*

$$\theta_{it} = \begin{bmatrix} \boldsymbol{\phi}_i'(L) & 0 \end{bmatrix} \boldsymbol{\varepsilon}_{it}$$

satisfies Assumption 3.

The first $m-1$ shocks are common shocks, which could drive fundamentals or serve as common noise. The last shock is a private noise, and its variance is agent specific. This *ex-ante* heterogeneity in agents' information structures is absent in previous sections. We use the diagonal matrix $\boldsymbol{\Upsilon} = \operatorname{diag}\left(\gamma_1^{-1}, \ldots, \gamma_n^{-1}\right)$ to summarize this heterogeneity.

In the multi-action setup from Section 4.1, the relevant modification of the signal process only depends on the degree of strategic complementarity. With heterogeneous payoff and information structures, the required modification is more involved and hinges on both $\mathbf{W}$ and $\boldsymbol{\Upsilon}$.

Assumption 6. *The matrix $(\mathbf{I}-\mathbf{W})^{-1}\boldsymbol{\Upsilon}$ is diagonalizable with*

$$(\mathbf{I}-\mathbf{W})^{-1}\boldsymbol{\Upsilon} = \mathbf{Q}\operatorname{diag}(\widetilde{\gamma}_1^{-1}, \ldots, \widetilde{\gamma}_n^{-1})\,\mathbf{Q}^{-1}$$

where $\widetilde{\gamma}_j^{-1}$ is its j-th eigenvalue and the absolute values of all eigenvalues are less than one.

In a similar way to the multi-action analysis, we define the j-th **network-modified signal process** to be

$$\widetilde{\mathbf{x}}_{jt} = \mathbf{M}(L)\,\widetilde{\boldsymbol{\varepsilon}}_{jt},$$

where the covariance matrix of $\widetilde{\boldsymbol{\varepsilon}}_{jt}$ is $\widetilde{\boldsymbol{\Sigma}}_j^2 = \operatorname{diag}\left(\tau_1^{-1}, \ldots, \tau_{m-1}^{-1},\ \widetilde{\gamma}_j^{-1}\tau_m^{-1}\right)$. The modified precision takes into account both the relative importance of the network linkages and the relative precision of agents' signals. These modified signals turn out to be sufficient to establish our single-agent solution.

[35]In Online Appendix F, we allow $\mathbf{M}(L)$ to depend on time t.

[36]The fact that agents are assumed to have only one private shock is made mostly for presentation purposes, the assumption is relaxed in the proof presented in Online Appendix F.

Theorem 2. *The forecast rule conditional on the j-th network modified signal process is defined as*

$$\widetilde{\mathbb{E}}\left[\boldsymbol{\theta}_t | \, \widetilde{\mathbf{x}}_j^t\right] = \mathbf{g}_j'(L)\widetilde{\mathbf{x}}_{jt}.$$

Under Assumptions 5 and 6, the equilibrium policy rule $\mathbf{h}(L)$ is given by[37]

$$\mathbf{h}'(L) = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\mathbf{Q}^{-1}(\mathbf{I}-\mathbf{W})^{-1}\mathbf{g}_j'(L).$$

Proof. See Online Appendix F. □

The solution in Theorem 2 modifies its perfect information counterpart (7.2) in two ways: first, the fundamental $\boldsymbol{\theta}_t$ is replaced by forecasts of the fundamental; second, the policy rule is a weighted average of n different Katz-Bonacich centrality measures. This result is, unsurprisingly, more complicated than the case in which all agents are symmetric. Instead of forming expectations about the average expectations, agents need to keep track of the average of their neighbors' expectations of the average of their neighbors' expectations of the fundamental, and so on.[38] Incidentally, this is why the proof does not follow directly from Corollary 2. Regardless, the main properties of the single-agent solution are maintained. With this network structure, solving the equilibrium via the fixed point approach is a very demanding task, as one needs to solve for n sets of policy rules. In contrast, our single-agent solution remains tractable. Besides, it shows how the payoff and information structures matter for the equilibrium in a transparent way.

In a similar, though static, setting, De Martí and Zenou (2015) and Bergemann, Heumann, and Morris (2017) characterize the equilibrium by Katz-Bonacich centrality measures weighted by shocks and covariances between fundamentals and signals, respectively. Our characterization extends to dynamic information structures. Importantly, when information is persistent, agents naturally want to keep track of the entire history of signals, which is an infinite-dimension object. By expressing the policy rule in terms of finite forecasts of the fundamental, one can utilize the Kalman filter and derive the policy rule recursively. We think this tractability will be useful for applied work.

We conclude this section by providing an example to show how the network structure interacts with incomplete information.

Example 2. Consider a ring network. The length of the network or the number of agents is n. Assume that in the network, only agent 1 cares about the fundamental θ_t directly and all other

[37] As a corollary it follows that the equilibrium exists and is unique.

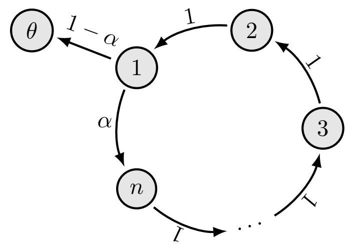
[38] By iteration on equation (7.1), we obtain $\mathbf{y}_t = \sum_{k=0}^{\infty} \overline{\mathbb{E}}_t^k[\mathbf{W}\overline{\mathbb{E}}_t[\boldsymbol{\theta}_t]]$ where the higher-order expectations $\overline{\mathbb{E}}_t^k[\mathbf{W}\overline{\mathbb{E}}_t[\boldsymbol{\theta}_t]]$ are defined recursively as $\overline{\mathbb{E}}_t^0[\mathbf{W}\overline{\mathbb{E}}_t[\boldsymbol{\theta}_t]] = \overline{\mathbb{E}}_t[\boldsymbol{\theta}_t]$, and $\overline{\mathbb{E}}_t^k[\mathbf{W}\overline{\mathbb{E}}_t[\boldsymbol{\theta}_t]] = \overline{\mathbb{E}}_t[\mathbf{W}\overline{\mathbb{E}}_t^{k-1}[\boldsymbol{\theta}_t]]$.

agents only care about the previous agent's action. The best response functions are given by

$$y_{it} = \begin{cases} (1-\alpha)\mathbb{E}_{1t}[\theta_t] + \alpha\mathbb{E}_{1t}[y_{nt}], & i = 1 \\ \mathbb{E}_{it}[y_{i-1,t}], & i > 1 \end{cases}$$

The corresponding $\mathbf{W}$ matrix, which we present next to a diagram of the network structure, is given by

$$\mathbf{W} = \begin{bmatrix} 0 & 0 & \cdots & \alpha \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}$$



Suppose, further, that the fundamental, θ_t, follows a random walk, and that, every period, each agent observes a private signal about the fundamental,

$$x_{it} = \theta_t + \nu_{it}, \quad \nu_{it} \sim \mathcal{N}(0, (\gamma_i \tau_\nu)^{-1}).$$

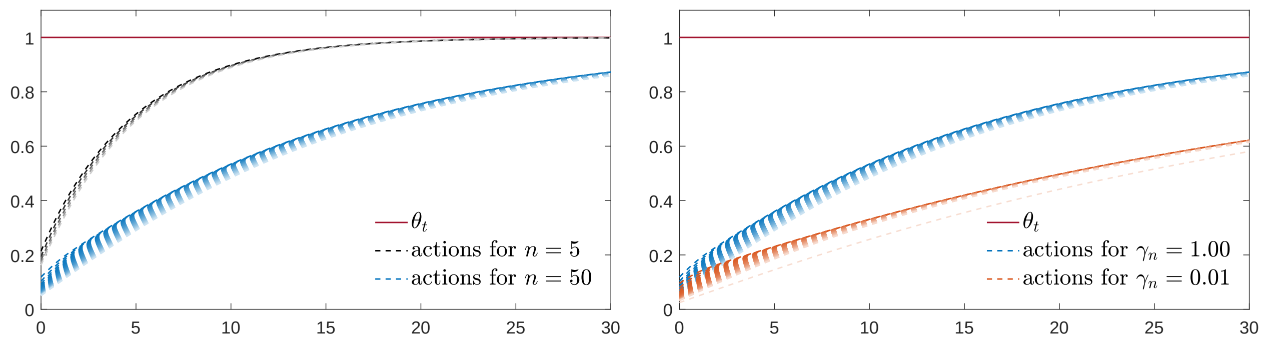
Note that the precision is agent specific.

With perfect information, it is easy to verify that agents' actions are independent of their position in the network, and all the responses are simply equal to the fundamental itself, $y_{it} = \theta_t$.

In contrast, the network structure matters a lot when information is incomplete. First, consider the case in which $\gamma_i = 1$ for all i. If $\alpha = 0$, agent 1's action is equal to the first-order expectation about the fundamental, agent 2's action is equal to the second-order expectation, and, by induction, agent n's action is equal to the n-th order expectation. Clearly, an agent's position in the network shapes their reaction. The further away an agent is from agent 1, the deeper the depth of reasoning required. When $\alpha \neq 0$, agent 1 also cares about agent n's expectations, and all the higher-order expectations start to play a role. However, the extent to which a particular higher-order expectation matters depends on the length of the network. In particular,

$$y_{it} = (1-\alpha)\mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^{i-1}[\theta_t]\right] + (1-\alpha)\sum_{k=1}^{\infty}\alpha^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^{nk+i-1}[\theta_t]\right], \quad \text{for } i \in 1, \ldots, n.$$

Note that, with a larger n, the weight on higher-order expectations increases as agents have to think through more layers of others' expectations. Figure 4A shows the impulse responses to a fundamental shock. Clearly, as the length of the network increases from $n = 5$ to $n = 50$, the outcome becomes more dampened and more persistent. A higher n brings the outcome closer to expectations of higher order.



(A) Varying the length of the ring

(B) Varying the precision of the n-th agent

FIGURE 4: Impuse Responses for a Ring Network with Incomplete Information

Parameters (A & B): $\tau_\eta = \tau_\nu = \gamma_i = 1$, and $\alpha = 0.9$; (B): $n = 50$.

In Figure 4B we consider another experiment. We fix the length of the network to be $n = 50$. Instead of setting $\gamma_i = 1$ for all agents, we lower the precision of the n-th agent by setting $\gamma_n = 0.01$. Even though all other agents still have relatively accurate forecasts about the fundamental, a strong coordination motive makes them reluctant to respond in order to behave similarly to the n-th agent, which has the most dampened response. Given the large number of agents in the network, it is interesting to notice how much the change in precision of only one agent affects the action of all other agents. This pattern is similar to the tyranny of the least informed discussed in Golub and Morris (2017)—here we see its dynamic effects.

7.2 Beyond Static Beauty-Contest Model

So far, we have focused on beauty-contest models with static strategic complementarity, where an agent's action only depends on others' actions within the same period. This is the most extensively studied case in the literature on dispersed information, but there is a growing interest in understanding properties of beauty-contest models with intertemporal strategic complementarities. In this subsection, we consider the following best response function in which agents also care about the future aggregate action,

$$y_{it} = \gamma \mathbb{E}_{it}[\theta_t] + \alpha \mathbb{E}_{it}[y_t] + \beta \mathbb{E}_{it}[y_{t+1}]. \tag{7.4}$$

This specification is similar to Allen, Morris, and Shin (2006), Nimark (2017), Rondina and Walker (2018), Angeletos and Lian (2018). It arises when agents need to make intertemporal decisions, such as the consumers' Euler equation, the New Keynesian Phillips curve, or asset pricing equations. Angeletos and Huo (2018) obtain the solution to a similar forward-looking model. Here, we focus on its equivalence to a single-agent problem and provide the appropriate transformation to the signal process.

We assume that agents receive a public and a private signal

$$z_t = \theta_t + \varepsilon_t, \quad \text{and} \quad x_{it} = \theta_t + \nu_{it},$$

where $\varepsilon_t \sim \mathcal{N}(0, \tau_\varepsilon^{-1})$, and $\nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1})$. Suppose that the fundamental θ_t follows an AR(1) process as in (2.8) with $\eta_t \sim \mathcal{N}(0,1)$. The information set of agent i at time t is $\mathcal{I}_t = \{z^t, x_i^t\}$. Also, normalize $\gamma \equiv 1 - \alpha - \rho\beta$, such that the perfect information solution is simply $y_t = \theta_t$.

In the simple forecasting problem, we have that

$$\mathbb{E}_{it}[\theta_t] = \frac{\lambda}{\rho(1-\rho\lambda)} \frac{\tau_\varepsilon z_t + \tau_\nu x_{it}}{1 - \lambda L}, \tag{7.5}$$

where the weights on the signals are proportional to their precision and λ is given by a function Λ of the precision of the signals, and of the persistence of the fundamental,

$$\lambda \equiv \Lambda(\tau_\varepsilon, \tau_\nu, \rho) \equiv \frac{1}{2}\left(\frac{\tau_\varepsilon + \tau_\nu}{\rho} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_\varepsilon + \tau_\nu}{\rho} + \frac{1}{\rho} + \rho\right)^2 - 4}\right).$$

Proposition 8. *Suppose* $\left|\frac{\beta}{1-\alpha}\right| < 1$. *Then, the policy rule in equilibrium is given by*

$$y_{it} = \frac{\widetilde{\lambda}}{\rho(1-\rho\widetilde{\lambda})} \frac{\tau_\varepsilon z_t + \widetilde{\tau}_\nu x_{it}}{1 - \widetilde{\lambda} L}. \tag{7.6}$$

The modified precision of the private signal, $\widetilde{\tau}_\nu$, *and the persistence,* $\widetilde{\lambda}$, *are jointly determined by the system*

$$\widetilde{\tau}_\nu \equiv (1 - \alpha - \widetilde{\lambda}\beta)\tau_\nu, \quad \textit{and} \quad \widetilde{\lambda} \equiv \Lambda(\tau_\varepsilon, \widetilde{\tau}_\nu, \rho).$$

Proof. See Online Appendix G.[39] □

Comparing equations (7.5) and (7.6), one can see that, in this forward-looking beauty-contest model, the policy rule is still equivalent to a pure forecasting problem with a modified signal process that can be interpreted as a single-agent solution. The key difference is in how the signal is modified. When $\beta = 0$, the solution collapses to the one in our baseline case without forward-looking behavior,

[39] See the remark at the end of Appendix A for a discussion of why the proof strategy applied in Theorem 1 cannot be applied when there are dynamic complementarities. This is why the proof of Proposition 8 is different and more involved.

and the precision is simply discounted by the degree of static complementarity, that is $\widetilde{\tau}_\nu = (1-\alpha)\tau_\nu$. When $\beta \neq 0$, the modification is more involved as it requires solving a nonlinear equation,

$$\widetilde{\tau}_\nu \equiv (1 - \alpha - \Lambda(\tau_\varepsilon, \widetilde{\tau}_\nu, \rho)\beta)\, \tau_\nu.$$

The higher-order expectations, in this case, involve forecasts that take place across different horizons since agents need to forecast how others think about the fundamental tomorrow, how others think about how others think about the fundamental the day after tomorrow, and so on. The required discounting is, therefore, not just the intertemporal complementarity, because an adjustment is necessary to capture the effects of the anticipation of learning by other agents in the future as well. In any case, the basic insight that the equilibrium policy rule is akin to a pure forecasting of the fundamental with a discounted precision of private signals remains true.[40]

More generally, the best response function could be written as

$$y_{it} = \gamma \mathbb{E}_{it}[\theta_t] + \mathbb{E}_{it}[g(L)y_t],$$

where $g(L)$ is a two-sided polynomial in L. The best response (3.1) corresponds to $g(L) = \alpha$, and the best response (7.4) corresponds to $g(L) = \alpha + \beta L^{-1}$, which are two common cases discussed in the literature. Despite the simplicity of our single-agent solution, it is subject to the limitation that it may not work for general $g(L)$. Notice that Proposition 8 holds when there are forward-looking complementarities and the fundamental follows an AR(1) process. Alternatively, if there are backward-looking complementarities or if the fundamental follows a more complicated process, the result may not hold. We provide an example of this kind in Online Appendix G.6.[41] This helps draw a limit to our results, though the single-agent solution could still serve as a good approximation for the exact equilibrium dynamics in those cases.

8 Conclusion

This paper establishes the equivalence between the equilibrium of beauty-contest models and the solution to a single-agent forecasting problem with a modified signal process. We have shown that the policy rule in the latter also solves the former. What makes our single-agent solution powerful is the fact that the required modification of the signal structure takes a simple form and that it works for a general class of information structures. This allows us to explore general properties of beauty-contest models and makes it suitable for quantitative applications. We extend this result to models with multiple actions and to network games.

[40] Differently from the static beauty-contest model in which the single-agent solution works for essentially all linear signal processes, the type of single-agent solution in Proposition 8 may not generalize to more arbitrary signal processes.

[41] This example includes backward-looking complementarity, which can result, for instance, from an environment with endogenous capital accumulation.

We believe the methods developed in this paper can potentially be applied to explore many interesting questions. For example, with incomplete information, how do the GE attenuation effects change when market concentration increases? How does the persistence of inflation change when monetary policy becomes more or less aggressive? How does information dispersion affect the propagation mechanism for sectoral shocks in an economy with input-output linkages? The guess-and-verify approach or the numerical approximation methods commonly used in the literature, to some extent, limit the understanding and applicability of dispersed information models. We hope that our results can help overcome these limitations.

References

Allen, F., S. Morris, and H. S. Shin (2006): "Beauty contests and iterated expectations in asset markets," *The Review of Financial Studies*, 19(3), 719–752.

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

Angeletos, G.-M., F. Collard, and H. Dellas (2018): "Quantifying confidence," *Econometrica*, 86(5), 1689–1726.

Angeletos, G.-M., and Z. Huo (2018): "Myopia and Anchoring," Discussion paper, National Bureau of Economic Research.

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

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

Angeletos, G.-M., and C. Lian (2016a): "Dampening General Equilibrium: From Micro Elasticities to Macro Effects," Discussion paper, Working paper.

——— (2016b): "Incomplete Information in Macroeconomics: Accommodating Frictions in Coordination," *Handbook of Macroeconomics*, 2, 1065–1240.

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

Angeletos, G.-M., and A. Pavan (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–1142.

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

Benhabib, J., P. Wang, and Y. Wen (2015): "Sentiments and Aggregate Demand Fluctuations," *Econometrica*, 83, 549–585.

Bergemann, D., T. Heumann, and S. Morris (2015): "Information and volatility," *Journal of Economic Theory*, 158, Part B, 427 – 465, Symposium on Information, Coordination, and Market Frictions.

——— (2017): "Information and Interaction," Discussion paper, Yale University, Cowles Foundation Discussion Paper No. 2088.

Bergemann, D., and S. Morris (2013): "Robust Predictions in Games With Incomplete Information," *Econometrica*, 81(4), 1251–1308.

Carlsson, H., and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.

De Martí, J., and Y. Zenou (2015): "Network games with incomplete information," *Journal of Mathematical Economics*, 61, 221–240.

Drenik, A., and D. Perez (2015): "Price Setting Under Uncertainty About Inflation," 2015 Meeting Papers 1429, Society for Economic Dynamics.

Ely, J. C., and M. Pęski (2011): "Critical Types," *Review of Economic Studies*, 78(3), 907–937.

Fajgelbaum, P. D., E. Schaal, and M. Taschereau-Dumouchel (2017): "Uncertainty traps," *The Quarterly Journal of Economics*, 132(4), 1641–1692.

Golub, B., and S. Morris (2017): "Expectations, Networks, and Conventions," Discussion paper.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press, Princeton, 2 edn.

Hellwig, C. (2004): "Heterogeneous Information and the Benefits of Public Information Disclosures (October 2005)," UCLA Economics Online Papers 283, UCLA Department of Economics.

Hellwig, C., and L. Veldkamp (2009): "Knowing what others Know: Coordination motives in information acquisition," *Review of Economic Studies*, 76, 223–251.

Huo, Z., and N. Takayama (2017a): "Higher Order Beliefs, Confidence and Business Cycles," Manuscript, Yale University.

——— (2017b): "Rational Expectations Models with Higher Order Beliefs," Manuscript, Yale University.

Kajii, A., and S. Morris (1997): "The Robustness of Equilibria to Incomplete Information," *Econometrica*, 65(6), 1283–1310.

Keynes, J. M. (1936): *The General Theory of Employment, Interest and Money*. Cambridge University Press.

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

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

——— (2015): "Business Cycle Dynamics under Rational Inattention," *The Review of Economic Studies*, 82(4), 1502–1532.

Melosi, L. (2016): "Signalling effects of monetary policy," *The Review of Economic Studies*, 84(2), 853–884.

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

Nimark, K. (2017): "Dynamic higher order expectations," Working paper.

Rondina, G., and T. B. Walker (2018): "Dispersed Information and Confounding Dynamics," Working paper.

Rubinstein, A. (1989): "The Electronic Mail Game: Strategic Behavior under "Almost Common Knowledge."," *American Economic Review*, 79(3), 385–391.

Sargent, T. J. (1991): "Equilibrium with signal extraction from endogenous variables," *Journal of Economic Dynamics and Control*, 15(2), 245–273.

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

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

Weinstein, J., and M. Yildiz (2007a): "A Structure Theorem for Rationalizability with Application to Robust Predictions of Refinements," *Econometrica*, 75(2), 365–400.

——— (2007b): "Impact of Higher-Order Uncertainty," *Games and Economic Behavior*, 60(1), 200–212.

Woodford, M. (2002): "Imperfect Common Knowledge and the Effects of Monetary Policy," in *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honour of Edmund S. Phelps*, ed. by P. Aghion, R. Frydman, J. Stiglitz, and M.Woodford. Princeton University Press.

Appendix

A Proof of Theorem 1

The proof is presented as a series of lemmas and propositions. In Section A.1 we show that the forecasting problem set up in Section 3 is equivalent to a limit of a truncated version of it. Section A.2 sets up and solves a static forecasting problem equivalent to the truncated version. Section A.3 presents and solves the associated fixed point problem that gives the equilibrium of a beauty-contest problem; in it, we also establish the existence, uniqueness, and linearity of the equilibrium. Section A.4 describes the α-modified signal process and Section A.5 proves the equivalence between the policy function of the solution to the static version of the forecasting problem with the α-modified signal process and the solution to the fixed point problem.

A.1 Limit of Truncated Forecasting Problem

Fix t. Section 3 sets up the problem of forecasting θ_t given $\mathbf{x}_i^t \equiv \{\mathbf{x}_{it}, \mathbf{x}_{it-1}, \ldots\}$. For ease of notation, from now on we suppress the dependence on t and i. Notice that each element of θ and $\mathbf{x}$ can be represented as an MA(∞) process and that there is an infinite history of signals.

Consider a truncated version of this problem.[42] Let θ_q be the MA(q) truncation of θ, that is,

$$\theta = \sum_{k=0}^{\infty} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} \Rightarrow \theta_q = \sum_{k=0}^{q} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k}.$$

Let $\mathbf{x}^{(N)} \equiv \{\mathbf{x}_{it}, \ldots, \mathbf{x}_{it-N}\}$ and let each element of $\mathbf{x}_p^{(N)}$ be the MA(p) truncation of the corresponding element of $\mathbf{x}^{(N)}$. The next proposition shows that the limit as q, p, and N go to infinity of the forecast of θ_q given $\mathbf{x}_p^{(N)}$ is equivalent to the forecast of θ given $\mathbf{x}$. Throughout, the concept of convergence between random variables is mean square. For example, we can say that $\lim_{q\to\infty} \theta_q = \theta$, since

$$\lim_{q\to\infty} \mathbb{E}\left[(\theta - \theta_q)^2\right] = \lim_{q\to\infty} \mathbb{E}\left[\left(\theta - \sum_{k=0}^{q} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k}\right)^2\right] = \lim_{q\to\infty} \sum_{k=q+1}^{\infty} \boldsymbol{\phi}_k \mathbb{E}[\boldsymbol{\eta}_{t-k}^2] \boldsymbol{\phi}_k' = 0,$$

where the last equality is due to the assumption that $\boldsymbol{\phi}(L)$ is square summable and that $\mathbb{E}[\boldsymbol{\eta}_{t-k}^2]$ is finite.

Proposition 9. $\mathbb{E}[\theta|\mathbf{x}] = \lim_{p,q,N\to\infty} \mathbb{E}\left[\theta_q|\mathbf{x}_p^{(N)}\right]$.

Proof. The strategy is to establish the following equalities

$$\mathbb{E}[\theta|\mathbf{x}] \overset{[3]}{=} \lim_{p\to\infty} \mathbb{E}[\theta|\mathbf{x}_p] \overset{[2]}{=} \lim_{p\to\infty} \lim_{N\to\infty} \mathbb{E}\left[\theta|\mathbf{x}_p^{(N)}\right] \overset{[1]}{=} \lim_{p\to\infty} \lim_{N\to\infty} \lim_{q\to\infty} \mathbb{E}\left[\theta_q|\mathbf{x}_p^{(N)}\right].$$

We start from the last and move to the first.

[42] Note that in this truncation, we do not assume shocks become public after a certain number of periods, differently from the common assumption made in the literature (e.g. Townsend (1983)).

[1]: To show that

$$\mathbb{E}\left[\theta \mid \mathbf{x}_p^{(N)}\right] = \lim_{q\to\infty} \mathbb{E}\left[\theta_q | \mathbf{x}_p^{(N)}\right],$$

note that there exists K large enough such that, for any $k > K$,

$$\mathbb{E}\left[\boldsymbol{\eta}_{t-k} \mid \mathbf{x}_p^{(N)}\right] = 0.$$

It follows that

$$\mathbb{E}\left[\theta \mid \mathbf{x}_p^{(N)}\right] = \mathbb{E}\left[\sum_{k=0}^{K} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} + \sum_{k=K+1}^{\infty} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} \mid \mathbf{x}_p^{(N)}\right] = \mathbb{E}\left[\sum_{k=0}^{K} \boldsymbol{\phi}_k \boldsymbol{\eta}_{t-k} \mid \mathbf{x}_p^{(N)}\right] = \mathbb{E}\left[\theta_K | \mathbf{x}_p^{(N)}\right],$$

and

$$\lim_{q\to\infty} \mathbb{E}\left[\theta_q | \mathbf{x}_p^{(N)}\right] = \lim_{q\to\infty} \mathbb{E}\left[\sum_{k=0}^{K} \boldsymbol{\phi}_k \boldsymbol{\varepsilon}_{-k} + \sum_{k=K+1}^{q} \boldsymbol{\phi}_k \boldsymbol{\varepsilon}_{-k} \mid \mathbf{x}_p^{(N)}\right] = \lim_{q\to\infty} \mathbb{E}\left[\theta_K | \mathbf{x}_p^{(N)}\right] = \mathbb{E}\left[\theta_K | \mathbf{x}_p^{(N)}\right].$$

[2]: Next, to show that

$$\mathbb{E}[\theta | \mathbf{x}_p] = \lim_{N\to\infty} \mathbb{E}\left[\theta | \mathbf{x}_p^{(N)}\right],$$

we simply need to establish that the limit on the right hand side exists. First notice that forecast errors are decreasing in the number of signals and that the stationarity of θ guarantees that the mean squared error is well defined, which implies that

$$0 \leq \mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N+1)}\right]\right)^2\right] \leq \mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N)}\right]\right)^2\right].$$

Therefore, there exists σ^2 such that

$$\lim_{N\to\infty} \mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N)}\right]\right)^2\right] = \sigma^2.$$

Moreover, for any N_1, N_2,

$$\mathbb{E}\left[\left(\theta - \frac{\mathbb{E}\left[\theta | \mathbf{x}_p^{(N_1)}\right] + \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_2)}\right]}{2}\right)^2\right] \geq \sigma^2.$$

It follows that

$$\begin{aligned}
&\mathbb{E}\left[\left(\mathbb{E}\left[\theta | \mathbf{x}_p^{(N_1)}\right] - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_2)}\right]\right)^2\right] \\
&\quad = 2\mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_1)}\right]\right)^2\right] + 2\mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_2)}\right]\right)^2\right] - 4\mathbb{E}\left[\left(\theta - \frac{\mathbb{E}\left[\theta | \mathbf{x}_p^{(N_1)}\right] + \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_2)}\right]}{2}\right)^2\right] \\
&\quad \leq 2\mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_1)}\right]\right)^2\right] + 2\mathbb{E}\left[\left(\theta - \mathbb{E}\left[\theta | \mathbf{x}_p^{(N_2)}\right]\right)^2\right] - 4\sigma^2.
\end{aligned}$$

In the limit, the right-hand side converges to zero, and therefore

$$\lim_{N\to\infty} \mathbb{E}\left[\theta|\mathbf{x}_p^{(N)}\right]$$

is indeed well defined.

[3]: By the Winner-Hopf prediction formula, it follows that

$$\begin{aligned}
\mathbb{E}[\theta|\mathbf{x}] &= [\boldsymbol{\phi}(L)\mathbf{M}'(L^{-1})\mathbf{B}'(L^{-1})^{-1}]_+\mathbf{B}(L)^{-1}\mathbf{x} \equiv \mathbf{D}(L)\boldsymbol{\varepsilon}, \\
\mathbb{E}[\theta|\mathbf{x}_p] &= [\boldsymbol{\phi}(L)\mathbf{M}_p'(L^{-1})\mathbf{B}_p'(L^{-1})^{-1}]_+\mathbf{B}_p(L)^{-1}\mathbf{x}_p \equiv \mathbf{D}_p(L)\boldsymbol{\varepsilon}.
\end{aligned}$$

where $\mathbf{B}(z)$ and $\mathbf{B}_p(z)$ are the corresponding fundamental representations of $\mathbf{M}(z)$ and $\mathbf{M}_p(z)$, respectively. The mean-squared difference between these forecasts is

$$\mathbb{E}\left[\left(\mathbb{E}[\theta|\mathbf{x}] - \mathbb{E}[\theta|\mathbf{x}_p]\right)^2\right] = \sum_{k=0}^{\infty}(\mathbf{D}_{t-k} - \mathbf{D}_{p,t-k})\boldsymbol{\Sigma}^2(\mathbf{D}_{t-k} - \mathbf{D}_{p,t-k})'.$$

If $\lim_{p\to\infty} \mathbf{D}_p(z) = \mathbf{D}(z)$, then $\lim_{p\to\infty} \mathbb{E}\left[\left(\mathbb{E}[\theta|\mathbf{x}] - \mathbb{E}[\theta|\mathbf{x}_p]\right)^2\right] = 0$. First, by construction,

$$\mathbf{M}(z) = \lim_{p\to\infty}\sum_{k=0}^{p}\mathbf{M}_k z^k = \lim_{p\to\infty}\mathbf{M}_p(z).$$

Second, given a signal process $\mathbf{M}_p(z)$, its corresponding fundamental representation of $\mathbf{B}_p(z)$ is uniquely determined when the covariance matrix of the fundamental innovation is normalized, and satisfies

$$\mathbf{B}_p(z)\mathbf{B}_p'(z^{-1}) = \mathbf{M}_p(z)\mathbf{M}_p'(z^{-1}).$$

As a result,

$$\mathbf{B}(z) = \lim_{p\to\infty}\mathbf{B}_p(z).$$

Then, by continuity of the annihilation operator, it follows that

$$\lim_{p\to\infty}\mathbf{D}_p(z) = \mathbf{D}(z).$$

□

A.2 Static Forecasting Problem

For the sake of simplicity, from this point on, we suppress the dependence on the indexes used in the limits above and rewrite the truncated forecasting problem as a static problem. We also continue to suppress the dependence on t, unless any ambiguity may arise, but bring back the dependence on i.

Define

$$\boldsymbol{\varepsilon}_i \equiv \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\nu}_i \end{bmatrix}, \quad \boldsymbol{\eta} \equiv \begin{bmatrix} \boldsymbol{\eta}_t \\ \vdots \\ \boldsymbol{\eta}_{t-\max\{q,p\}-N} \end{bmatrix}, \quad \boldsymbol{\nu}_i \equiv \begin{bmatrix} \boldsymbol{\nu}_{it} \\ \vdots \\ \boldsymbol{\nu}_{it-\max\{q,p\}-N} \end{bmatrix},$$

and notice that, there exists a $u(q+1)\times 1$ vector $\mathbf{a}$ and a $r(N+1)\times M$ matrix $\mathbf{B}$, with $M \equiv m(\max\{q,p\}+N+2)$, such that the forecasting problem at time t becomes that of forecasting

$$\theta = \begin{bmatrix} \mathbf{a}' & \mathbf{0}' \end{bmatrix} \boldsymbol{\varepsilon}_i = \mathbf{a}'\boldsymbol{\eta}, \quad \text{given} \quad \mathbf{x}_i = \mathbf{B}\boldsymbol{\varepsilon}_i = \mathbf{B} \begin{bmatrix} \boldsymbol{\eta} \\ \boldsymbol{\nu}_i \end{bmatrix}.$$

Let $\boldsymbol{\Omega}^2$ denote the covariance matrix of $\boldsymbol{\varepsilon}_i$, let $\mathbf{A} \equiv \begin{bmatrix} \mathbf{a}' & \mathbf{0}' \end{bmatrix}$, and let $\boldsymbol{\Lambda}$ be the $M \times M$ matrix given by

$$\boldsymbol{\Lambda} \equiv \begin{bmatrix} \mathbf{I}_U & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

where $U \equiv u(\max\{q,p\} + N + 2)$. It follows that

$$\mathbb{E}[\theta \mid \mathbf{x}_i] = \mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{x}_i.$$

It is convenient for what follows to write the forecast in this way. To obtain this formula we use, in particular, the fact that $\mathbf{A}' = \mathbf{A}'\boldsymbol{\Lambda}$ and that $\boldsymbol{\Lambda}\boldsymbol{\Omega} = \boldsymbol{\Omega}\boldsymbol{\Lambda}$.

A.3 Fixed Point Problem

Suppose we also want to forecast y. We do not know the stochastic process for y, so let $\mathbf{h}$ denote the agent's equilibrium policy function, i.e. $y_i = \mathbf{h}'\mathbf{x}_i$, then

$$y = \int \mathbf{h}'\mathbf{x}_i = \mathbf{h}'\mathbf{B}\boldsymbol{\Lambda}\boldsymbol{\varepsilon}_i.$$

Then, the forecast of y is given by

$$\mathbb{E}[y \mid \mathbf{x}_i] = \mathbf{h}'\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{x}_i.$$

In equilibrium, we have that

$$y_i = (1-\alpha)\mathbb{E}[\theta \mid \mathbf{x}_i] + \alpha\mathbb{E}[y \mid \mathbf{x}_i],$$

and, therefore

$$\mathbf{h}'\mathbf{x}_i = \left[(1-\alpha)\mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} + \alpha\mathbf{h}'\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\right] \mathbf{x}_i.$$

It follows from the fact that equation above holds for any $\mathbf{x}_i$ that

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{d}, \tag{A.1}$$

where

$$\mathbf{C} \equiv \mathbf{I} - \alpha \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}', \tag{A.2}$$

$$\mathbf{d} \equiv (1-\alpha) \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A}. \tag{A.3}$$

Lemma A.1. $\mathbf{C}$ *is invertible.*

Proof. We start by showing that $\mathbf{M} \equiv \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}'$ has real eigenvalues in $[0, 1]$. First notice that $\mathbf{M}$ has real, non-negative eigenvalues since it is similar to

$$\begin{aligned}(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{1/2}\mathbf{M}(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2} &= (\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{1/2}(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1}(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}')(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2} \\ &= (\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2}(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}')(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}')^{-1/2}\end{aligned}$$

which is positive semidefinite. On the other hand,

$$\mathbf{I} - \mathbf{M} = \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\boldsymbol{\Omega}(\mathbf{I} - \boldsymbol{\Lambda})\boldsymbol{\Omega}\mathbf{B}',$$

which, analogously to $\mathbf{M}$, is also similar to a positive semidefinite matrix. If λ is an eigenvalue of $\mathbf{M}$, then $1 - \lambda$ is an eigenvalue of $\mathbf{I} - \mathbf{M}$. Therefore, the fact that the eigenvalues of $\mathbf{I} - \mathbf{M}$ are positive implies that the eigenvalues of $\mathbf{M}$ must be less than or equal to 1, as desired. It follows that $\mathbf{S} \equiv \sum_{j=0}^{\infty}(\alpha\mathbf{M})^j$ converges, and $\mathbf{S}(\mathbf{I} - \alpha\mathbf{M}) = \mathbf{I}$. ☐

In particular, it follows from this lemma that there exists a unique equilibrium to the beauty-contest model. It also follows that the equilibrium actions of an agent are linear functions of their signals.[43]

A.4 α-Modified Signal Process and Prediction Formula

Define $\boldsymbol{\Gamma}$ to be the $M \times M$ matrix given by

$$\boldsymbol{\Gamma} \equiv \begin{bmatrix} \mathbf{I}_U & \mathbf{0} \\ \mathbf{0} & \frac{1}{\sqrt{1-\alpha}}\mathbf{I}_{M-U} \end{bmatrix},$$

and suppose that the signals observed by agent i are a α-modified version of $\mathbf{x}_i$ given by

$$\widetilde{\mathbf{x}}_i = \mathbf{B}\widetilde{\boldsymbol{\varepsilon}}_i, \quad \text{with} \quad \widetilde{\boldsymbol{\varepsilon}}_i \equiv \boldsymbol{\Gamma}\boldsymbol{\varepsilon}_i.$$

It follows that

$$\widetilde{\mathbb{E}}[\theta \mid \widetilde{\mathbf{x}}_i] = \mathbf{A}'\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' \left(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}'\right)^{-1} \widetilde{\mathbf{x}}_i. \tag{A.4}$$

[43] In the arguments made above we have implicitly used the well known result that the optimal forecast for Gaussian processes is linear, see Hamilton (1994) Section 4.6 for a formal proof.

A.5 Equivalence Result

Proposition 10 establishes that the right hand side of equation terms in equation (A.1) is to the right hand side of equation (A.4), which completes the proof of Theorem 1.

Proposition 10. *Using the definitions above, it follows that*

$$\mathbf{C}^{-1}\mathbf{d} = \left(\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A}.$$

Proof. From the definition of $\mathbf{C}$, equation (A.2), we obtain

$$\begin{aligned}\mathbf{C} &= \mathbf{I} - \alpha\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{B}' \\ &= \left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}(\mathbf{I}-\alpha\boldsymbol{\Lambda})\boldsymbol{\Omega}\mathbf{B}'.\end{aligned}$$

Thus, since

$$\mathbf{I} - \alpha\boldsymbol{\Lambda} = (1-\alpha)\boldsymbol{\Gamma}^2,$$

it follows that

$$\mathbf{C} = (1-\alpha)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}'.$$

Finally, using Lemma A.1 and equation (A.3),

$$\begin{aligned}\mathbf{C}^{-1}\mathbf{d} &= \left[(1-\alpha)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}'\right]^{-1}(1-\alpha)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A} \\ &= (\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}')^{-1}\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)\left(\mathbf{B}\boldsymbol{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A} \\ &= (\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Gamma}^2\boldsymbol{\Omega}\mathbf{B}')^{-1}\mathbf{B}\boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}\mathbf{A}.\end{aligned}$$

□

Remark It is crucial for the proof to work that the matrix $\mathbf{I}-\alpha\boldsymbol{\Lambda}$ is real, symmetric and positive semidefinite so that it can be interpreted as a covariance matrix. When there are forward-looking or backward-looking strategic complementarities this is in general not the case. This does not mean that a transformation to the information structure that yields an equivalence result cannot exist with dynamic complementarities, but simply that this particular proof strategy is not suitable in those cases. See, for instance, Proposition 8.

B Proof of Proposition 5

Proof. By Proposition 2, the individual action can be written as

$$y_{it} = \mathbf{G}\mathbf{z}_{it},$$

where

$$\mathbf{z}_{it} = (\mathbf{I}-\mathbf{K}\mathbf{H})\mathbf{F}\mathbf{z}_{it-1} + \mathbf{K}\mathbf{x}_{it} = (\mathbf{I}-(\mathbf{F}-\mathbf{K}\mathbf{H}\mathbf{F})L)^{-1}\mathbf{K}\mathbf{x}_{it},$$

and $\mathbf{K}$ is the steady state Kalman gain matrix with α-modified signals. Note that by Cramer's rule,

$$(\mathbf{I}-(\mathbf{F}-\mathbf{KHF})L)^{-1}=\frac{\text{adj}(\mathbf{I}-(\mathbf{F}-\mathbf{KHF})L)}{\det(\mathbf{I}-(\mathbf{F}-\mathbf{KHF})L)}=\frac{\text{adj}(\mathbf{I}-(\mathbf{F}-\mathbf{KHF})L)}{\Pi_{k=1}^{\ell}(1-\lambda_k L)},$$

where $\{\lambda_k\}_{k=1}^{\ell}$ are non-zero eigenvalues of $(\mathbf{F}-\mathbf{KHF})$. Let r denote the dimension of $\mathbf{F}$, the last equality follows from the fact that

$$\det\left(\mathbf{I}-(\mathbf{F}-\mathbf{KHF})\,L\right)=L^r\det\left(\mathbf{I}L^{-1}-(\mathbf{F}-\mathbf{KHF})\right)=L^r L^{-(r-\ell)}\prod_{k=1}^{\ell}\left(L^{-1}-\lambda_k\right)=\prod_{k=1}^{\ell}\left(1-\lambda_k L\right).$$

The aggregate outcome y_t then follows

$$y_t=\int y_{it}=\mathbf{C}(L)\int \mathbf{x}_{it}=\mathbf{C}(L)\Big(\mathbf{A}\theta_t+\mathbf{B}(L)\mathbf{v}_t\Big).$$

□

C Variance and Covariance of Average Forecast

Let $\mathbf{\Omega}=\text{diag}(\sigma_1,\ldots,\sigma_M)$, and $\mathcal{I}\equiv\{U+1,\ldots,M\}$ be the σ-indexes associated with the idiosyncratic shocks $\boldsymbol{\nu}_i$. Also, let $\mathbf{e}_j$ be the j-th column of the $M\times M$ identity matrix and let $\overline{\mathbb{E}}[\theta]\equiv\int\mathbb{E}[\theta\mid\mathbf{x}_i]$.

Increasing the variance of any element of $\boldsymbol{\nu}_i$ does not affect θ, but makes the signals $\mathbf{x}_i$ noisier which, in turn, makes the forecast less accurate. This implies that the forecast reacts less to signals and, as a result, it is less volatile. It also implies that it is less correlated with the actual θ. This motivates the following lemmas.

Lemma C.1. *The variance* $\text{Var}(\overline{\mathbb{E}}[\theta])$ *is decreasing in how noisy the signals are,*

$$\frac{\partial\text{Var}(\overline{\mathbb{E}}[\theta])}{\partial\sigma_j^2}\leq 0,\qquad \textit{for } j\in\mathcal{I}.$$

Proof. First notice that

$$\text{Var}(\overline{\mathbb{E}}[\theta])=\mathbf{A}'\mathbf{\Omega}^2\mathbf{B}'\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}'\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\mathbf{\Omega}^2\mathbf{A}.$$

For any $j\in\mathcal{I}$, $\mathbf{A}'\mathbf{\Omega}^2$ and $\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}$ do not depend on σ_j^2, and therefore, taking derivatives yields

$$\frac{\partial\text{Var}(\overline{\mathbb{E}}[\theta])}{\partial\sigma_j^2}=-\,\sigma_j^2\mathbf{A}'\mathbf{\Omega}^2\mathbf{B}'\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\Big[\left(\mathbf{B}\mathbf{e}_j\mathbf{e}_j'\mathbf{B}'\right)\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}'+ \\ +\mathbf{B}\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}'\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\left(\mathbf{B}\mathbf{e}_j\mathbf{e}_j'\mathbf{B}'\right)\Big]\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{B}\mathbf{\Omega}^2\mathbf{A}.$$

The matrix in the inner bracket is symmetric, so let $\mathbf{L}\mathbf{L}'$ denote its Cholesky decomposition. Then, letting $\mathbf{z}\equiv\mathbf{A}'\mathbf{\Omega}^2\mathbf{B}'\left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1}\mathbf{L}$, the right hand side is equal to $-\sigma_j^2\mathbf{z}\mathbf{z}'$, which is less than or equal to 0. □

Lemma C.2. *The covariance* $\text{Cov}(\theta, \overline{\mathbb{E}}[\theta])$ *is decreasing in how noisy the signals are,*

$$\frac{\partial \text{Cov}(\theta, \overline{\mathbb{E}}[\theta])}{\partial \sigma_j^2} \leq 0, \quad \text{for } j \in \mathcal{I}.$$

Proof. First notice that

$$\text{Cov}(\theta, \overline{\mathbb{E}}[\theta]) = \mathbf{A}'\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega}\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\mathbf{\Omega}^2\mathbf{A}.$$

For any $j \in \mathcal{I}$, $\mathbf{A}'\mathbf{\Omega}^2$ does not depend on σ_j^2, and $\mathbf{A}'\mathbf{\Omega}\mathbf{\Lambda}\mathbf{\Omega} = \mathbf{A}'\mathbf{\Omega}^2$, therefore

$$\frac{\partial \text{Cov}(\theta, \overline{\mathbb{E}}[\theta])}{\partial \sigma_j^2} = -\mathbf{A}'\mathbf{\Omega}^2\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} \left(\mathbf{B}\mathbf{e}_j\mathbf{e}_j'\mathbf{B}'\right) \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\mathbf{\Omega}^2\mathbf{A}.$$

Finally, letting $\mathbf{z} \equiv \mathbf{A}'\mathbf{\Omega}^2\mathbf{B}' \left(\mathbf{B}\mathbf{\Omega}^2\mathbf{B}'\right)^{-1} \mathbf{B}\mathbf{e}_j$, the right hand side of this equation can be written as $-\mathbf{z}\mathbf{z}'$ which is less than or equal to 0. □

For the following proofs we denote $\widetilde{\mathbb{E}}[\theta] \equiv \int \mathbb{E}[\theta \mid \widetilde{\mathbf{x}}_i]$, where $\widetilde{\mathbf{x}}_i$ is the modified signal defined in Section A.4.

C.1 Proof of Proposition 6 for the Common-Value Best Response

Proof. In the static version of the problem, under Assumptions 1-3 and with the best response given by

$$y_i = (1-\alpha)\mathbb{E}_i[\theta] + \alpha\mathbb{E}_i[y],$$

it follows from Corollary 1 that

$$y = \widetilde{\mathbb{E}}[\theta],$$

which is equivalent to $\overline{\mathbb{E}}[\theta]$ with the variance of the idiosyncratic shocks discounted by α, that is with $\widetilde{\sigma}_j^2 = \sigma_j^2/(1-\alpha)$ for all $j \in \mathcal{I}$. Hence, an increase in α is equivalent to an increase in the variance of all the idiosyncratic shocks. Thus, Part 1 in Proposition 6 for the CVBR follows from Lemmas C.1 and C.2. Part 2 follows directly from Lemma C.1 and the fact that $y = \overline{\mathbb{E}}[\theta]$ when $\alpha = 0$. Finally, it follows from Proposition 9 that these results generalize to the setting in Proposition 6 by a continuity argument. □

C.2 Proof of Proposition 6 for the Independent-Value Best Response

Proof. With the best response given by

$$y_i = (1-\alpha)\theta_i + \alpha\mathbb{E}_i[y],$$

it follows from Proposition 3 that

$$y = (1-\alpha)\theta + \alpha\widetilde{\mathbb{E}}[\theta].$$

Notice that

$$\text{Var}(y) = (1-\alpha)^2\text{Var}(\theta) + \alpha^2\text{Var}(\widetilde{\mathbb{E}}[\theta]) + 2\alpha(1-\alpha)\text{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]).$$

Then, using the law of total variance and Lemmas C.1 and C.2, it follows that, if $\alpha > 0$,

$$\begin{aligned}
\frac{\partial \text{Var}(y)}{\partial \alpha} =& -2(1-\alpha)\text{Var}(\theta) + 2\alpha\text{Var}(\widetilde{\mathbb{E}}[\theta]) + 2(1-2\alpha)\text{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]) \\
&+ \alpha^2 \frac{\partial \text{Var}(\widetilde{\mathbb{E}}[\theta])}{\partial \alpha} + 2\alpha(1-\alpha)\frac{\partial \text{Cov}(\widetilde{\mathbb{E}}[y], \theta)}{\partial \alpha} \\
\leq& -2(1-\alpha)\text{Var}(\theta) + 2\alpha\text{Var}(\theta) + 2(1-2\alpha)\text{Var}(\theta) = 0.
\end{aligned}$$

Similarly,

$$\text{Cov}(y, \theta) = (1-\alpha)\text{Var}(\theta) + \alpha\text{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]),$$

and if $\alpha > 0$,

$$\begin{aligned}
\frac{\partial \text{Cov}(y, \theta)}{\partial \alpha} &= -\text{Var}(\theta) + \text{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]) + \alpha \frac{\partial \text{Cov}(\widetilde{\mathbb{E}}[y], \theta)}{\partial \alpha} \\
&\leq -\text{Var}(\theta) + \text{Var}(\theta) = 0.
\end{aligned}$$

This establishes Part 1 for IVBR. Part 3 for $\alpha > 0$ follows from the fact that $y_t = \theta_t$ when $\alpha = 0$ and the result in Part 1. Finally, we can also write the aggregate action y as

$$y = \theta + \alpha(\widetilde{\mathbb{E}}[\theta] - \theta).$$

and it follows that

$$\text{Var}(y) = \text{Var}(\theta) + \alpha^2\text{Var}(\widetilde{\mathbb{E}}[\theta] - \theta) + 2\alpha\,\text{Cov}(\theta,\, \widetilde{\mathbb{E}}[\theta] - \theta).$$

To show that, for $\alpha < 0$, $\text{Var}(y) \geq \text{Var}(\theta)$, it is sufficient to show that $\text{Cov}(\theta,\, \widetilde{\mathbb{E}}[\theta] - \theta) \leq 0$. Note that

$$\text{Cov}(\theta,\, \widetilde{\mathbb{E}}[\theta] - \theta) = \text{Cov}(\theta, \widetilde{\mathbb{E}}[\theta]) - \text{Var}(\theta) \leq 0.$$

Proposition 9 implies that these results generalize to the setting in Proposition 6 by continuity. □

D Proofs for Generalized Best Response and Multiple Actions

This appendix contains the proofs for the extensions to generalized best responses and multiple actions from Section 4.3.

D.1 Proof of Proposition 3

Proof. Iterating on the best response function in equation (4.5) we obtain

$$y_{it} = \gamma(\varphi_t + \xi_{it}) + \alpha\gamma \sum_{k=0}^{\infty} \alpha^k\, \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\varphi_t\right]\right].$$

By Corollary 2, the infinite sum of higher-order expectations can be rewritten as a first-order expectation

$$y_{it} = \gamma(\varphi_t + \xi_{it}) + \frac{\alpha\gamma}{1-\alpha}\mathbf{h}_t'(L)\mathbf{x}_{it}.$$

□

D.2 Proof of Proposition 4

Proof. Iterating on equation (4.6) leads to

$$\mathbf{y}_{it} = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\boldsymbol{\theta}_t\right]\right].$$

Then, notice that for any $k \in \{0, 1, 2, \ldots\}$,

$$\mathbf{A}^k = \mathbf{Q}\boldsymbol{\Omega}^k\mathbf{Q}^{-1} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\mathbf{Q}^{-1}\alpha_j^k.$$

Therefore, $\mathbf{y}_{it}$ can be written as

$$\mathbf{y}_{it} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\mathbf{Q}^{-1}\sum_{k=0}^{\infty} \alpha_j^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\boldsymbol{\theta}_t\right]\right].$$

From Corollary 2 we have that each row of $\sum_{k=0}^{\infty} \alpha_j^k \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k\left[\boldsymbol{\theta}_t\right]\right]$ is equal to the corresponding row of $(1-\alpha_j)^{-1}\mathbf{g}_{jt}'(L)\mathbf{x}_{it}$ and it follows that

$$\mathbf{y}_{it} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'(1-\alpha_j)^{-1}\mathbf{Q}^{-1}\mathbf{g}_{jt}'(L)\mathbf{x}_{it} = \sum_{j=1}^{n} \mathbf{Q}\mathbf{e}_j\mathbf{e}_j'\left(\mathbf{I}-\boldsymbol{\Omega}\right)^{-1}\mathbf{Q}^{-1}\mathbf{g}_{jt}'(L)\mathbf{x}_{it},$$

and the fact that $\left(\mathbf{I}-\boldsymbol{\Omega}\right)^{-1}\mathbf{Q}^{-1} = \mathbf{Q}^{-1}\left(\mathbf{I}-\mathbf{A}\right)^{-1}$ concludes the proof. □

ONLINE APPENDIX: NOT FOR PUBLICATION

E Example Economies

In this appendix, we describe two different economic environments that have equilibria that can be summarized by a system of equations with a best response equation and an aggregation equation of the kind we work with in the paper.

E.1 Monetary Model with Dispersed Information

In this section, we describe a simple monetary model with information frictions; it is a version of the model in Woodford (2002). A representative household has period utility given by

$$U(C_t, N_t) = \frac{C_t^{1-\sigma}}{1-\sigma} - \chi \frac{N_t^{1+\kappa}}{1+\kappa}$$

where N_t denotes their labor supply and C_t is their consumption of a composite good defined to be the CES aggregator of a continuum of differentiated goods,

$$C_t = \left(\int C_{it}^{\frac{\eta-1}{\eta}} \right)^{\frac{\eta}{\eta-1}}.$$

The demand for good i and supply of labor in period t are given by

$$C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\eta} C_t, \quad \text{and} \quad N_t = \left(\frac{W_t}{\chi P_t C_t^{\sigma}} \right)^{\frac{1}{\kappa}}, \qquad \text{with} \quad P_t \equiv \left(\int P_{it}^{1-\eta} \right)^{\frac{1}{1-\eta}},$$

and where W_t denotes the wage. There is continuum of firms, each producing one of the differentiated goods with the following production function,

$$C_{it} = AN_{it}^{\varepsilon}.$$

Firms have private information about the state of the world and, in period t, solve

$$\max_{P_{it}, C_{it}, N_{it}} \mathbb{E}_{it}[P_{it}C_{it} - W_t N_{it}], \qquad \text{subject to} \quad C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\eta} C_t, \quad \text{and} \quad C_{it} = A_t N_{it}^{\varepsilon}.$$

It follows that

$$P_{it} = \frac{\eta}{\eta - 1} \frac{\mathbb{E}_{it}\left[W_t C_{it}^{\frac{1-\varepsilon}{\varepsilon}} \right]}{\varepsilon A^{\frac{1}{\varepsilon}}}.$$

Nominal GDP is determined exogenously by a monetary shock, Θ_t, so that

$$P_t C_t = \Theta_t.$$

Using this equation and the household's optimality conditions we obtain

$$P_{it} = \left(\frac{\chi\eta}{A^{\frac{1}{\varepsilon}}\varepsilon\left(\eta-1\right)}\right)^{\frac{\varepsilon}{\eta+\varepsilon(1-\eta)}} \mathbb{E}_{it}\left[\Theta_t^{\frac{(\sigma-1)\varepsilon+1}{\eta+\varepsilon(1-\eta)}} P_t^{1-\frac{(\sigma-1)\varepsilon+1}{\eta+\varepsilon(1-\eta)}} N_t^{\frac{\kappa\varepsilon}{\eta+\varepsilon(1-\eta)}}\right].$$

Let lower-case variables denote log-deviations from steady state. Integrating the production function we obtain as first order approximation (which is exact if all shocks are log-normal) that $c_t = \varepsilon n_t$, and it follows that

$$p_{it} = (1-\alpha)\,\mathbb{E}_{it}[\theta_t] + \alpha\mathbb{E}_{it}[p_t], \qquad \text{with} \quad p_t = \int p_{it},$$

and the degree of strategic complementarity given by

$$\alpha \equiv \frac{(1-\varepsilon)\left(\eta-1\right)+\varepsilon\left(1-\sigma\right)-\kappa}{\eta+\varepsilon(1-\eta)}.$$

E.2 Business Cycles Model

In this section, we describe a stylized real business cycle model with information frictions, a simplified version of the model in Angeletos and La'O (2010).

Environment There is a continuum of islands indexed by i in the economy. In each island lives a representative agent who specializes in producing differentiated good i. Each period, agent i consumes C_{ijt} of the good produced on island j, and C_{it} is the CES aggregator of agent i's consumption of all goods,

$$C_{it} = \left(\int_j C_{ijt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}.$$

The production technology is

$$Y_{it} = \Theta_{it} N_{it}^{\varepsilon},$$

where η_{it} is the productivity level on island i in period t, and N_{it} denotes the labor input by agent i. The period utility of agent i is given by

$$U(C_{it}, N_{it}) = \frac{C_{it}^{1-\sigma}}{1-\sigma} - \frac{N_{it}^{1+\kappa}}{1+\kappa}.$$

Each period has two stages: In the first stage, agents decide how much to produce, that is, choose N_{it} which determines Y_{it}, conditional on their information about the island-specific productivity and aggregate output. In the second stage, taking prices and Y_{it} as given, they choose how much to consume of each good, i.e., $(C_{ijt})_j$, subject to the budget constraint

$$\int_j P_{jt} C_{ijt} = P_{it} Y_{it},$$

where P_{jt} is the price of the good produced on island j. Trading, in this second stage, occurs in a centralized market. In the first stage, agents have to decide how much to produce before the goods market opens, and therefore they have to forecast aggregate output to infer the price of their own goods. We allow agents to have different information sets which include their own productivity Θ_{it}, but otherwise, we remain agnostic about the information structure.

Equilibrium Characterization Beginning with second stage, the optimal demand of the representative agent i for the good from island j whose price is P_{jt} is given by

$$C_{ijt} = C_{it}\left(\frac{P_{jt}}{P_t}\right)^{-\eta}, \qquad \text{where} \quad P_t \equiv \left(\int P_{jt}^{1-\eta}\right)^{\frac{1}{1-\eta}}.$$

Together with the budget constraint and market clearing condition $\int_i C_{ijt} = Y_{jt}$, it follows that

$$C_{it} = Y_t^{\frac{1}{\eta}} Y_{it}^{\frac{\eta-1}{\eta}} \qquad \text{where} \quad Y_t \equiv \left(\int Y_{jt}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}},$$

that is, the consumption of island i is a weighted geometric mean of the aggregate output and the output produced in the island. Using this equation and the production function the first-stage problem becomes

$$\max_{Y_{it}} \mathbb{E}_{it}\left[\frac{1}{1-\sigma}\left(Y_t^{\frac{1}{\eta}} Y_{it}^{\frac{\eta-1}{\eta}}\right)^{1-\sigma} - \frac{1}{1+\kappa}\left(\frac{Y_{it}}{\Theta_{it}}\right)^{\frac{1+\kappa}{\varepsilon}}\right].$$

which implies

$$Y_{it} = \left(\varepsilon\frac{\eta-1}{\eta}\Theta_{it}^{\frac{1+\kappa}{\varepsilon}}\mathbb{E}_{it}\left[Y_t^{\frac{1-\sigma}{\eta}}\right]\right)^{\frac{\varepsilon\eta}{\eta(1+\kappa)+\varepsilon(1-\eta)(1-\sigma)}}.$$

Letting lower-case letters denote log-deviations from steady state, it follows that

$$y_{it} = \gamma\theta_{it} + \alpha\mathbb{E}_{it}[y_t], \qquad \text{with} \quad y_t = \int y_{it},$$

and the following definitions

$$\alpha \equiv \frac{\varepsilon\,(1-\sigma)}{\eta\,(1+\kappa) + \varepsilon\,(1-\eta)\,(1-\sigma)}, \qquad \text{and} \qquad \gamma \equiv \alpha\frac{\eta(1+\kappa)}{\varepsilon(1-\sigma)}.$$

This type of best response function is analyzed in Section 4.3.

F Proof of Theorem 2

The proof is presented as a series of lemmas and propositions. An analogous argument to the one made in Section A.1 holds so that, if we prove the result for an arbitrary static information structure, the result follows. Section F.1 sets up and solves the aforementioned arbitrary static forecasting problem. Section F.2 presents and solves the associated fixed point problem that gives the equilibrium of a beauty-contest problem. Section F.3 describes the modified signal process and Section F.4 proves the equivalence between the policy function of the solution to the static version of the forecasting problem with the modified signal process and the solution to the fixed point problem.

F.1 Setup with Static Information Structure

Best response function The best response function is

$$\mathbf{y} = \mathbb{E}[\boldsymbol{\theta}] + \mathbb{E}[\mathbf{W}\mathbf{y}],$$

where $\mathbf{y}$ is a vector of individual actions

$$\mathbf{y} \equiv \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}',$$

$\boldsymbol{\theta}$ is a vector of exogenous variables

$$\boldsymbol{\theta} \equiv \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_n \end{bmatrix}',$$

and the expectation operator $\mathbb{E}$ is given by

$$\mathbb{E} \equiv \begin{bmatrix} \mathbb{E}_1 & \mathbb{E}_2 & \cdots & \mathbb{E}_n \end{bmatrix}'.$$

Agents can have heterogeneous information sets, and $\mathbb{E}_i$ may be different from $\mathbb{E}_j$. The matrix $\mathbf{W}$ represents the network structure.

Information structure Let $\boldsymbol{\varepsilon}_i$ be the vector of, normally distributed, shocks the first u being common shocks, and the last r private. The covariance matrix of $\boldsymbol{\varepsilon}_i$ is the identity matrix. Suppose that the process for fundamentals is given by

$$\theta_i = \begin{bmatrix} \boldsymbol{\phi}_i' & \mathbf{0}_{1\times r} \end{bmatrix} \boldsymbol{\varepsilon}_i$$

and that agent i's signal is given by

$$\mathbf{x}_i = \mathbf{M}_i \boldsymbol{\varepsilon}_i$$

where

$$\mathbf{M}_i \equiv \mathbf{M}\boldsymbol{\Sigma}_i,$$

and

$$\boldsymbol{\Sigma}_i \equiv \text{diag}\left(\tau_1^{-1/2}, \dots, \tau_u^{-1/2}, \tau_{1i}^{-1/2}, \dots, \tau_{ri}^{-1/2}\right),$$

with $u + r \equiv m$. Moreover, let

$$\mathbf{\Delta} \equiv \text{diag}(\tau_1^{-1/2}, \dots, \tau_u^{-1/2}, 0, \dots, 0),$$

and define $\mathbf{\Lambda}$ and $\mathbf{\Gamma}$ to be

$$\mathbf{\Lambda} \equiv \begin{bmatrix} \mathbf{I}_u \\ \mathbf{0}_{r\times u} \end{bmatrix}_{m\times u}, \quad \text{and} \quad \mathbf{\Gamma} \equiv \mathbf{I}_m - \mathbf{\Lambda}\mathbf{\Lambda}',$$

so that, in particular, we have

$$\mathbf{\Sigma}_i \mathbf{\Lambda} = \mathbf{\Delta}\mathbf{\Lambda}, \quad \text{and} \quad \mathbf{\Sigma}_i \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{\Sigma}_j = \mathbf{\Delta}^2.$$

Let $\mathbf{E}_i^n$ be the $n \times n$ matrix with zeros everywhere and 1 at the position (i, i), and define

$$\overline{\mathbf{M}} = \sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{M}_i, \quad \text{and} \quad \mathbf{\Sigma} = \sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{\Sigma}_i.$$

Finally, for each private shock, indexed by p, collect the associate variance for each agent i in the diagonal of the following matrix:

$$\mathbf{\Upsilon}_p \equiv \text{diag}\left(\tau_{p1}^{-1}, \tau_{p2}^{-1}, \dots, \tau_{pn}^{-1}\right).$$

Forecast The forecast of the fundamental of agent j by agent i is given by

$$\mathbb{E}[\theta_j \mid \mathbf{x}_i] = \boldsymbol{\phi}_j' \mathbf{\Lambda}' \mathbf{M}_i' \left(\mathbf{M}_i \mathbf{M}_i'\right)^{-1} \mathbf{x}_i \equiv \mathbf{g}_{ji}' \mathbf{x}_i. \tag{F.1}$$

F.2 Fixed Point Problem

Let $\mathbf{h}_i$ be the equilibrium policy function for agent i, that is

$$y_i = \mathbf{h}_i' \mathbf{x}_i = \mathbf{h}_i' \mathbf{M}_i \boldsymbol{\varepsilon}_i.$$

Then, since the agents do not have information about others' private shocks, the forecast of y_j by agent i is given by

$$\mathbb{E}[y_j \mid \mathbf{x}_i] = \mathbf{h}_j' \mathbf{M}_j \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_i' \left(\mathbf{M}_i \mathbf{M}_i'\right)^{-1} \mathbf{x}_i.$$

In equilibrium, we have that

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \mathbb{E}[\theta_1 \mid \mathbf{x}_1] \\ \mathbb{E}[\theta_2 \mid \mathbf{x}_2] \\ \vdots \\ \mathbb{E}[\theta_n \mid \mathbf{x}_n] \end{bmatrix} + \begin{bmatrix} 0\ \mathbb{E}[y_1 \mid \mathbf{x}_1] + w_{12}\mathbb{E}[y_2 \mid \mathbf{x}_1] + \cdots + w_{1n}\mathbb{E}[y_n \mid \mathbf{x}_1] \\ w_{21}\mathbb{E}[y_1 \mid \mathbf{x}_2] + \ 0\ \mathbb{E}[y_2 \mid \mathbf{x}_2] + \cdots + w_{2n}\mathbb{E}[y_n \mid \mathbf{x}_2] \\ \vdots \\ w_{n1}\mathbb{E}[y_1 \mid \mathbf{x}_n] + w_{n2}\mathbb{E}[y_2 \mid \mathbf{x}_n] + \cdots + \ 0\ \mathbb{E}[y_n \mid \mathbf{x}_n] \end{bmatrix},$$

so that

$$\begin{bmatrix} \mathbf{h}_1' \mathbf{x}_1 \\ \mathbf{h}_2' \mathbf{x}_2 \\ \vdots \\ \mathbf{h}_n' \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_1' \mathbf{\Lambda}' \mathbf{M}_1' \left(\mathbf{M}_1 \mathbf{M}_1'\right)^{-1} \mathbf{x}_1 \\ \boldsymbol{\phi}_2' \mathbf{\Lambda}' \mathbf{M}_2' \left(\mathbf{M}_2 \mathbf{M}_2'\right)^{-1} \mathbf{x}_2 \\ \vdots \\ \boldsymbol{\phi}_n' \mathbf{\Lambda}' \mathbf{M}_n' \left(\mathbf{M}_n \mathbf{M}_n'\right)^{-1} \mathbf{x}_n \end{bmatrix} + \begin{bmatrix} \sum_{k\neq 1} w_{1k} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_1' \left(\mathbf{M}_1 \mathbf{M}_1'\right)^{-1} \mathbf{x}_1 \\ \sum_{k\neq 2} w_{2k} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_2' \left(\mathbf{M}_2 \mathbf{M}_2'\right)^{-1} \mathbf{x}_2 \\ \vdots \\ \sum_{k\neq n} w_{nk} \mathbf{h}_k' \mathbf{M}_k \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{M}_n' \left(\mathbf{M}_n \mathbf{M}_n'\right)^{-1} \mathbf{x}_n \end{bmatrix},$$

and, therefore, using the fact that this equation holds for any $(\mathbf{x}_1, \dots, \mathbf{x}_n)$ we can write

$$\begin{bmatrix} \mathbf{h}_1'\mathbf{M}_1\mathbf{M}_1' \\ \mathbf{h}_2'\mathbf{M}_2\mathbf{M}_2' \\ \vdots \\ \mathbf{h}_n'\mathbf{M}_n\mathbf{M}_n' \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}_1'\boldsymbol{\Lambda}'\mathbf{M}_1' \\ \boldsymbol{\phi}_2'\boldsymbol{\Lambda}'\mathbf{M}_2' \\ \vdots \\ \boldsymbol{\phi}_n'\boldsymbol{\Lambda}'\mathbf{M}_n' \end{bmatrix} + \begin{bmatrix} \sum_{k\neq 1} w_{1k}\mathbf{h}_k'\mathbf{M}_k\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\mathbf{M}_1' \\ \sum_{k\neq 2} w_{2k}\mathbf{h}_k'\mathbf{M}_k\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\mathbf{M}_2' \\ \vdots \\ \sum_{k\neq n} w_{nk}\mathbf{h}_k'\mathbf{M}_k\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\mathbf{M}_n' \end{bmatrix}.$$

Transposing each row we get

$$\begin{bmatrix} \mathbf{M}_1\mathbf{M}_1'\mathbf{h}_1 \\ \mathbf{M}_2\mathbf{M}_2'\mathbf{h}_2 \\ \vdots \\ \mathbf{M}_n\mathbf{M}_n'\mathbf{h}_n \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1\boldsymbol{\Lambda}\boldsymbol{\phi}_1 \\ \mathbf{M}_2\boldsymbol{\Lambda}\boldsymbol{\phi}_2 \\ \vdots \\ \mathbf{M}_n\boldsymbol{\Lambda}\boldsymbol{\phi}_n \end{bmatrix} + \begin{bmatrix} \sum_{k\neq 1} w_{1k}\mathbf{M}_1\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\mathbf{M}_k'\mathbf{h}_k \\ \sum_{k\neq 2} w_{2k}\mathbf{M}_2\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\mathbf{M}_k'\mathbf{h}_k \\ \vdots \\ \sum_{k\neq n} w_{nk}\mathbf{M}_n\boldsymbol{\Lambda}\boldsymbol{\Lambda}'\mathbf{M}_k'\mathbf{h}_k \end{bmatrix},$$

which can be rewritten as

$$\overline{\mathbf{M}}\,\overline{\mathbf{M}}'\mathbf{h} = \overline{\mathbf{M}}(\mathbf{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{\phi} + \overline{\mathbf{M}}(\mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\overline{\mathbf{M}}'\mathbf{h},$$

where $\mathbf{h}$ and $\boldsymbol{\phi}$ are defined to be

$$\boldsymbol{\phi} = \begin{bmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \\ \vdots \\ \boldsymbol{\phi}_n \end{bmatrix}_{nu\times 1}, \quad \text{and} \quad \mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \vdots \\ \mathbf{h}_n \end{bmatrix}.$$

Solving for $\mathbf{h}$ we obtain

$$\mathbf{h} = \mathbf{C}^{-1}\mathbf{d},$$

where

$$\begin{aligned} \mathbf{C} &\equiv \overline{\mathbf{M}}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\overline{\mathbf{M}}', \\ \mathbf{d} &\equiv \overline{\mathbf{M}}(\mathbf{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{\phi}. \end{aligned}$$

The fact that $\mathbf{C}$ is invertible is established below in Section F.4. It follows that the equilibrium to the beauty-contest model exists and is unique.

F.3 Modified Signal Process and Prediction Formula

Define $\mathbf{D}$ to be

$$\mathbf{D} \equiv \sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{D}_i,$$

with each $\mathbf{D}_i$ given by

$$\mathbf{D}_i \equiv \mathbf{I}_n \otimes \boldsymbol{\Delta} + \sum_{p=1}^{r}[\boldsymbol{\Omega}_p \otimes \mathbf{E}^m_{u+p}] = \operatorname{diag}\left(\tau_1^{-1/2}, \ldots, \tau_u^{-1/2}, \widetilde{\tau}_{1i}^{-1/2}, \ldots, \widetilde{\tau}_{ri}^{-1/2}\right),$$

and where $\widetilde{\tau}_{pi}^{-1}$ is the i-th eigenvalue of $(\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p$. Suppose that the signals observed by agent i are a modified version of $\mathbf{x}_i$ given by

$$\widetilde{\mathbf{x}}_i = \mathbf{M}\mathbf{D}_i\boldsymbol{\varepsilon}_i.$$

Notice that, relative to $\boldsymbol{\Sigma}_i$, $\mathbf{D}_i$ simply replaces the precision of the private signals τ_{pi} by the transformed $\widetilde{\tau}_{pi}$. It follows that

$$\mathbb{E}[\theta_j \mid \widetilde{\mathbf{x}}_i] = \boldsymbol{\phi}_j'\boldsymbol{\Lambda}'\boldsymbol{\Delta}\mathbf{M}'\left(\mathbf{M}\mathbf{D}_i^2\mathbf{M}'\right)^{-1}\widetilde{\mathbf{x}}_i \equiv \mathbf{g}_{ji}'\mathbf{x}_i. \tag{F.2}$$

where we used the fact that $\boldsymbol{\Lambda}'\mathbf{D}_i = \boldsymbol{\Lambda}'\boldsymbol{\Delta}$. Moreover, we need the following assumption.

Assumption 7. *The matrix $(\mathbf{I}_n - \mathbf{W})$ is invertible, $(\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p$ is diagonalizable, all of its eigenvalues have absolute value less than 1, and all of its eigenvectors are independent of p.*[44]

Under this assumption, letting $\mathbf{Q}$ denote the matrix composed of the eigenvectors of $(\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p$, its eigendecomposition allows us to write

$$(\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p = \mathbf{Q}\boldsymbol{\Omega}_p\mathbf{Q}^{-1} \tag{F.3}$$

where

$$\boldsymbol{\Omega}_p \equiv \begin{bmatrix} \widetilde{\tau}_{p1}^{-1} & 0 & \cdots & 0 \\ 0 & \widetilde{\tau}_{p2}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \widetilde{\tau}_{pn}^{-1} \end{bmatrix}_{n \times n}.$$

F.4 Equivalence Result

We start with a useful lemma, then, using this lemma we prove a proposition that establishes the invertibility of the matrix $\mathbf{C}$ and a final proposition that establishes the equivalence result.

Lemma F.1. *Under Assumption 7 and using the definitions above, it follows that*

$$\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} = [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m]\mathbf{D}^2[\mathbf{Q}^{-1} \otimes \mathbf{I}_m]$$

[44] A trivial case where this holds is when $\boldsymbol{\Upsilon}_p = \gamma_p\boldsymbol{\Upsilon}$ or when $r = 1$.

Proof. First notice that

$$\begin{aligned}
\mathbf{\Sigma}[(\mathbf{I}_n - \mathbf{W}) \otimes \mathbf{\Lambda}\mathbf{\Lambda}']\mathbf{\Sigma} &= \left(\sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{\Sigma}_i\right) [(\mathbf{I}_n - \mathbf{W}) \otimes \mathbf{\Lambda}\mathbf{\Lambda}'] \left(\sum_{j=1}^{n} \mathbf{E}_j^n \otimes \mathbf{\Sigma}_j\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n (\mathbf{I}_n - \mathbf{W}) \mathbf{E}_j^n \otimes \mathbf{\Sigma}_i \mathbf{\Lambda}\mathbf{\Lambda}' \mathbf{\Sigma}_j \\
&= (\mathbf{I}_n - \mathbf{W}) \otimes \mathbf{\Delta}^2,
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{\Sigma}[\mathbf{I}_n \otimes \mathbf{\Gamma}]\mathbf{\Sigma} &= \left(\sum_{i=1}^{n} \mathbf{E}_i^n \otimes \mathbf{\Sigma}_i\right) [\mathbf{I}_n \otimes \mathbf{\Gamma}] \left(\sum_{j=1}^{n} \mathbf{E}_j^n \otimes \mathbf{\Sigma}_j\right) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n \mathbf{E}_j^n \otimes \mathbf{\Sigma}_i \mathbf{\Gamma} \mathbf{\Sigma}_j \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n \mathbf{E}_j^n \otimes \sum_{p=1}^{r} \sum_{q=1}^{r} \tau_{ip}^{-1/2} \tau_{jq}^{-1/2} \mathbf{E}_{u+p}^m \mathbf{E}_{qq}^m \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n \mathbf{E}_j^n \otimes \sum_{p=1}^{r} \tau_{ip}^{-1/2} \tau_{jp}^{-1/2} \mathbf{E}_{u+p}^m \\
&= \sum_{p=1}^{r} \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n \mathbf{E}_j^n \otimes \tau_{ip}^{-1/2} \tau_{jp}^{-1/2} \mathbf{E}_{u+p}^m \right] \\
&= \sum_{p=1}^{r} \left[\sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{E}_i^n \tau_{ip}^{-1/2} \tau_{jp}^{-1/2} \mathbf{E}_j^n \otimes \mathbf{E}_{u+p}^m \right] \\
&= \sum_{p=1}^{r} \left[\sum_{i=1}^{n} \mathbf{E}_i^n \tau_{ip}^{-1} \otimes \mathbf{E}_{u+p}^m \right] \\
&= \sum_{p=1}^{r} \left[\mathbf{\Upsilon}_p \otimes \mathbf{E}_{u+p}^m \right].
\end{aligned}$$

Hence,

$$
\begin{aligned}
\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} &= \boldsymbol{\Sigma}(\mathbf{I}_n \otimes (\boldsymbol{\Gamma} + \boldsymbol{\Lambda}\boldsymbol{\Lambda}') - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} \\
&= \boldsymbol{\Sigma}[\mathbf{I}_n \otimes \boldsymbol{\Gamma}]\boldsymbol{\Sigma} + \boldsymbol{\Sigma}[(\mathbf{I}_n - \mathbf{W}) \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}']\boldsymbol{\Sigma} \\
&= \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}^m_{u+p}\right] + (\mathbf{I}_n - \mathbf{W}) \otimes \boldsymbol{\Delta}^2 \\
&= \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}^m_{u+p}\right] - (\mathbf{I}_n - \mathbf{W}) \otimes \boldsymbol{\Gamma} + (\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma}) \\
&= \sum_{p=1}^{r} \left[\boldsymbol{\Upsilon}_p \otimes \mathbf{E}^m_{u+p}\right] - \sum_{p=1}^{r} \left[(\mathbf{I}_n - \mathbf{W}) \otimes \mathbf{E}^m_{u+p}\right] + (\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma}) \\
&= \sum_{p=1}^{r} \left[(\boldsymbol{\Upsilon}_p - \mathbf{I}_n + \mathbf{W}) \otimes \mathbf{E}^m_{u+p}\right] + (\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma}) \\
&= [(\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} \left[(\mathbf{I}_n - \mathbf{W})^{-1}(\boldsymbol{\Upsilon}_p - \mathbf{I}_n + \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1}\mathbf{E}^m_{u+p}\right] + \mathbf{I}_{nm} \right\} \\
&= [(\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} \left[((\mathbf{I}_n - \mathbf{W})^{-1}\boldsymbol{\Upsilon}_p - \mathbf{I}_n) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1}\mathbf{E}^m_{u+p}\right] + \mathbf{I}_{nm} \right\}.
\end{aligned}
$$

Next, using equation (F.3),

$$
\begin{aligned}
\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma} &= [(\mathbf{I}_n - \mathbf{W}) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} [(\mathbf{Q}\boldsymbol{\Omega}_p\mathbf{Q}^{-1} - \mathbf{I}_n) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1}\mathbf{E}^m_{u+p}] + \mathbf{I}_{nm} \right\} \\
&= [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \left\{ \sum_{p=1}^{r} [(\boldsymbol{\Omega}_p - \mathbf{I}_n) \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})^{-1}\mathbf{E}^m_{u+p}] + \mathbf{I}_{nm} \right\} [\mathbf{Q}^{-1} \otimes \mathbf{I}_m] \\
&= [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m] \left\{ \sum_{p=1}^{r} [(\boldsymbol{\Omega}_p - \mathbf{I}_n) \otimes \mathbf{E}^m_{u+p}] + [\mathbf{I}_n \otimes (\boldsymbol{\Delta}^2 + \boldsymbol{\Gamma})] \right\} [\mathbf{Q}^{-1} \otimes \mathbf{I}_m] \\
&= [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m] \left\{ \sum_{p=1}^{r} [\boldsymbol{\Omega}_p \otimes \mathbf{E}^m_{u+p}] + (\mathbf{I}_n \otimes \boldsymbol{\Delta}^2) \right\} [\mathbf{Q}^{-1} \otimes \mathbf{I}_m] \\
&= [(\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m]\mathbf{D}^2[\mathbf{Q}^{-1} \otimes \mathbf{I}_m].
\end{aligned}
$$

□

Proposition 11. *Under Assumption 7 and using the definitions above, it follows that*

$$\begin{bmatrix} \mathbf{h}_1' \\ \mathbf{h}_2' \\ \vdots \\ \mathbf{h}_n' \end{bmatrix} = \sum_{k=1}^{n} \mathbf{Q}\mathbf{E}_{kk}^{n}\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \begin{bmatrix} \mathbf{g}_{1k}' \\ \mathbf{g}_{2k}' \\ \vdots \\ \mathbf{g}_{nk}' \end{bmatrix}.$$

Proof. First notice that

$$\begin{aligned} (\mathbf{I}_n \otimes \mathbf{M})\boldsymbol{\Sigma}(\mathbf{I}_n \otimes \boldsymbol{\Lambda}) &= (\mathbf{I}_n \otimes \mathbf{M}) \left(\sum_{i=1}^{n} \mathbf{E}_i^n \otimes \boldsymbol{\Sigma}_i \right) (\mathbf{I}_n \otimes \boldsymbol{\Lambda}) \\ &= (\mathbf{I}_n \otimes \mathbf{M}) \left(\sum_{i=1}^{n} \mathbf{E}_i^n \otimes \boldsymbol{\Sigma}_i \boldsymbol{\Lambda} \right) \\ &= (\mathbf{I}_n \otimes \mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda}) \end{aligned}$$

So that, using this fact and Lemma F.1,

$$\begin{aligned} \mathbf{h} &= (\overline{\mathbf{M}}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\overline{\mathbf{M}}')^{-1}\overline{\mathbf{M}}(\mathbf{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{\phi} \\ &= \left[(\mathbf{I}_n \otimes \mathbf{M})\boldsymbol{\Sigma}(\mathbf{I}_{nm} - \mathbf{W} \otimes \boldsymbol{\Lambda}\boldsymbol{\Lambda}')\boldsymbol{\Sigma}(\mathbf{I}_n \otimes \mathbf{M}')\right]^{-1} (\mathbf{I}_n \otimes \mathbf{M})\boldsymbol{\Sigma}(\mathbf{I}_n \otimes \boldsymbol{\Lambda})\boldsymbol{\phi} \\ &= \left[(\mathbf{I}_n \otimes \mathbf{M}) \left((\mathbf{I}_n - \mathbf{W})\mathbf{Q} \otimes \mathbf{I}_m\right) \mathbf{D}^2 \left(\mathbf{Q}^{-1} \otimes \mathbf{I}_m\right) (\mathbf{I}_n \otimes \mathbf{M}')\right]^{-1} (\mathbf{I}_n \otimes \mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda})\boldsymbol{\phi} \\ &= (\mathbf{Q} \otimes \mathbf{I}_r) \left[(\mathbf{I}_n \otimes \mathbf{M})\mathbf{D}^2(\mathbf{I}_n \otimes \mathbf{M}')\right]^{-1} (\mathbf{I}_n \otimes \mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda}) \left(\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \otimes \mathbf{I}_u\right) \boldsymbol{\phi} \\ &= (\mathbf{Q} \otimes \mathbf{I}_r) \left(\sum_{k=1}^{n} \mathbf{E}_{kk}^n \otimes (\mathbf{M}\mathbf{D}_k^2\mathbf{M}')^{-1}\mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda} \right) \left(\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \otimes \mathbf{I}_u\right) \boldsymbol{\phi} \\ &= \sum_{k=1}^{n} \left[\mathbf{Q}\mathbf{E}_{kk}^n\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1} \otimes (\mathbf{M}\mathbf{D}_k^2\mathbf{M}')^{-1}\mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda}\right] \boldsymbol{\phi}. \end{aligned}$$

Finally, notice that, using the vectorization formula, we can rewrite this equation as

$$\mathbf{h} = \operatorname{vec}\left(\sum_{k=1}^{n} (\mathbf{M}\mathbf{D}_k^2\mathbf{M}')^{-1}\mathbf{M}\boldsymbol{\Delta}\boldsymbol{\Lambda} \begin{bmatrix} \boldsymbol{\phi}_1 & \boldsymbol{\phi}_2 & \cdots & \boldsymbol{\phi}_n \end{bmatrix} \left(\mathbf{Q}\mathbf{E}_k^n\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1}\right)' \right),$$

and, it follows from equation (F.2) that

$$\mathbf{h} = \operatorname{vec}\left(\sum_{k=1}^{n} \begin{bmatrix} \mathbf{g}_{1k} & \mathbf{g}_{2k} & \cdots & \mathbf{g}_{nk} \end{bmatrix} \left(\mathbf{Q}\mathbf{E}_k^n\mathbf{Q}^{-1}(\mathbf{I}_n - \mathbf{W})^{-1}\right)' \right).$$

The result follows by rearranging this equation. ☐

Notice that if all agents forecast the same fundamental, i.e. $\boldsymbol{\phi}_j = \boldsymbol{\phi}_i$ then

$$\mathbf{h}_i = \sum_{k=1}^{n} \omega_{ik} \mathbf{g}_k,$$

where

$$\omega_{ik} \equiv \sum_{j=1}^{n} \mathbf{e}_i' \mathbf{Q} \mathbf{E}_k^n \mathbf{Q}^{-1} (\mathbf{I}_n - \mathbf{W})^{-1} \mathbf{e}_j,$$

and and $\mathbf{e}_i$ is the i-th column of $\mathbf{I}_n$. Finally, notice that in the proof of Proposition 11 we also obtain the following corollary.

Corollary 4. *Under Assumption 7,* $\mathbf{C}$ *is invertible which implies that the equilibrium exists and is unique.*

G Proof of Proposition 8

In Section G.1 we obtain the canonical factorization of the auto-covariance generating function for the signal process which is necessary to apply the Wiener-Hopf prediction formula in Section G.2. Section G.3 presents and solves the fixed point problem that allows us to solve for the equilibrium explicitly. Section G.4 describes the modified signal process and Section G.5 shows the equivalence between the equilibrium policy function and the forecasting problem with the modified signal process.

G.1 Canonical Factorization

This information structure is tractable enough that we can solve for the equilibrium analitically using the Wiener-Hopf prediction formula to solve the necessary forecasting problems explicitly. The observation equation is

$$\begin{bmatrix} z_t \\ x_{it} \end{bmatrix} = \underbrace{\begin{bmatrix} \tau_\varepsilon^{-1/2} & 0 & \frac{1}{1-\rho L} \\ 0 & \tau_\nu^{-1/2} & \frac{1}{1-\rho L} \end{bmatrix}}_{\equiv \mathbf{M}(L)} \underbrace{\begin{bmatrix} \hat{\varepsilon}_t \\ \hat{\nu}_{it} \\ \hat{\eta}_t \end{bmatrix}}_{\equiv \hat{\mathbf{s}}_{it}}.$$

where $\hat{\mathbf{s}}_{it}$ is a vector of standardized normal random variables. Let $\mathbf{A}(L)$ be the auto-covariance generating function for the signal process, then

$$\mathbf{A}(L) \equiv \mathbf{M}(L)\mathbf{M}'\left(L^{-1}\right) = \frac{1}{(L-\rho)(1-\rho L)} \begin{bmatrix} L + \frac{(L-\rho)(1-\rho L)}{\tau_\varepsilon} & L \\ L & L + \frac{(L-\rho)(1-\rho L)}{\tau_\nu} \end{bmatrix}.$$

In order to apply the Wiener-Hopf prediction formula we need to obtain the canonical factorization of $\mathbf{A}(L)$. Accordingly, let λ be the inside root of the determinant of $\mathbf{A}(L)$, that is

$$\lambda = \frac{1}{2}\left(\frac{\tau_\varepsilon + \tau_\nu}{\rho} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_\varepsilon + \tau_\nu}{\rho} + \frac{1}{\rho} + \rho \right)^2 - 4} \right).$$

Then notice that

$$\mathbf{V} \equiv \frac{1}{\lambda(\tau_\nu + \tau_\varepsilon)} \begin{bmatrix} \frac{\lambda\tau_\nu + \rho\tau_\varepsilon}{\tau_\varepsilon} & \rho - \lambda \\ \rho - \lambda & \frac{\rho\tau_\nu + \lambda\tau_\varepsilon}{\tau_\nu} \end{bmatrix},$$

and

$$\mathbf{B}(L) \equiv \frac{1}{(\tau_\nu + \tau_\varepsilon)(1-\rho L)} \begin{bmatrix} \tau_\nu + \tau_\varepsilon - (\rho\tau_\nu + \lambda\tau_\varepsilon) L & (\rho - \lambda)\tau_\nu L \\ (\rho - \lambda)\tau_\varepsilon L & \tau_\nu + \tau_\varepsilon - (\lambda\tau_\nu + \rho\tau_\varepsilon) L \end{bmatrix},$$

are such that

$$\mathbf{B}(L)\mathbf{V}\mathbf{B}'\left(L^{-1}\right) = \mathbf{M}(L)\mathbf{M}'\left(L^{-1}\right).$$

G.2 Wiener-Hopf Prediction Formula

Applying the prediction formula, the forecast of $\theta_t = \begin{bmatrix} 0 & 0 & \frac{1}{1-\rho L} \end{bmatrix} \hat{\mathbf{s}}_{it}$ is given by

$$\mathbb{E}_{it}\left[\theta_t\right] = \left[\begin{bmatrix} 0 & 0 & \frac{1}{1-\rho L} \end{bmatrix} \mathbf{M}'\left(L^{-1}\right) \mathbf{B}'\left(L^{-1}\right)^{-1}\right]_+ \mathbf{V}^{-1}\mathbf{B}\left(L\right)^{-1} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix} = \frac{\lambda \begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}}{\rho\left(1-\lambda L\right)\left(1-\rho\lambda\right)} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix}. \tag{G.1}$$

Let $g\left(L\right) \equiv h_1(L) + h_2(L)$, then, the forecast about $y_t = \begin{bmatrix} \tau_\varepsilon^{-1/2} h_1\left(L\right) & 0 & \frac{g(L)}{1-\rho L} \end{bmatrix} \hat{\mathbf{s}}_{it}$ is given by

$$\begin{aligned}
\mathbb{E}_{it}\left[y_t\right] =& \left[\begin{bmatrix} \tau_\varepsilon^{-1/2} h_1\left(L\right) & 0 & \frac{g(L)}{1-\rho L} \end{bmatrix} \mathbf{M}'\left(L^{-1}\right) \mathbf{B}'\left(L^{-1}\right)^{-1}\right]_+ \mathbf{V}^{-1}\mathbf{B}\left(L\right)^{-1} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix} \\
=& \left\{ \frac{\begin{bmatrix} \left(\left(\rho\tau_\nu + \lambda\tau_\varepsilon + \lambda\rho\left(\lambda\tau_\nu + \rho\tau_\varepsilon\right)\right)L - \lambda\rho\left(\tau_\nu + \tau_\varepsilon\right)\left(1+L^2\right)\right)h_1\left(L\right) & \tau_\nu\left(\lambda-\rho\right)\left(1-\rho\lambda\right)Lh_1\left(L\right) \end{bmatrix}}{\rho\left(\tau_\nu+\tau_\varepsilon\right)\left(L-\lambda\right)\left(1-\lambda L\right)} \right. \\
&+ \frac{\begin{bmatrix} \tau_\varepsilon\left(\rho-\lambda\right)\left(1-\rho L\right)\lambda h_1\left(\lambda\right) & -\tau_\nu\left(\lambda-\rho\right)\lambda\left(1-\rho L\right)h_1\left(\lambda\right) \end{bmatrix}}{\rho\left(\tau_\nu+\tau_\varepsilon\right)\left(L-\lambda\right)\left(1-\lambda L\right)} \\
&+ \left. \frac{\lambda\left(L\left(1-\rho\lambda\right)g\left(L\right) - \lambda\left(1-\rho L\right)g\left(\lambda\right)\right)\begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}}{\rho\left(1-\rho\lambda\right)\left(L-\lambda\right)\left(1-\lambda L\right)} \right\} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix},
\end{aligned}$$

and the forecast about $y_{t+1} = \begin{bmatrix} \tau_\varepsilon^{-1/2} L^{-1} h_1\left(L\right) & 0 & \frac{L^{-1}g(L)}{1-\rho L} \end{bmatrix} \hat{\mathbf{s}}_{it}$ is given by

$$\begin{aligned}
\mathbb{E}_{it}\left[y_{t+1}\right] =& \left[\begin{bmatrix} \tau_\varepsilon^{-1/2} L^{-1} h_1\left(L\right) & 0 & \frac{L^{-1}g(L)}{1-\rho L} \end{bmatrix} \mathbf{M}'\left(L^{-1}\right) \mathbf{B}'\left(L^{-1}\right)^{-1}\right]_+ \mathbf{V}^{-1}\mathbf{B}\left(L\right)^{-1} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix} \\
=& \left\{ \frac{\begin{bmatrix} \left(\left(\rho\tau_\nu + \lambda\tau_\varepsilon + \lambda\rho\left(\lambda\tau_\nu + \rho\tau_\varepsilon\right)\right)L - \lambda\rho\left(\tau_\nu + \tau_\varepsilon\right)\left(1+L^2\right)\right)\lambda h_1\left(L\right) & \lambda\tau_\nu\left(\lambda-\rho\right)\left(1-\rho\lambda\right)Lh_1\left(L\right) \end{bmatrix}}{\rho\left(\tau_\nu+\tau_\varepsilon\right)\lambda L\left(L-\lambda\right)\left(1-\lambda L\right)} \right. \\
&+ \frac{\begin{bmatrix} \tau_\varepsilon\left(\rho-\lambda\right)\left(1-\rho L\right)\lambda L h_1\left(\lambda\right) & -\lambda\tau_\nu\left(\lambda-\rho\right)\left(1-\rho L\right)Lh_1\left(\lambda\right) \end{bmatrix}}{\rho\left(\tau_\nu+\tau_\varepsilon\right)\lambda L\left(L-\lambda\right)\left(1-\lambda L\right)} \\
&+ \frac{\begin{bmatrix} -\lambda\rho\left(L-\lambda\right)\left(\tau_\nu+\tau_\varepsilon-\left(\lambda\tau_\nu+\rho\tau_\varepsilon\right)L\right)h_1\left(0\right) & \lambda\tau_\nu\left(\lambda-\rho\right)L\rho\left(L-\lambda\right)h_1\left(0\right) \end{bmatrix}}{\rho\left(\tau_\nu+\tau_\varepsilon\right)\lambda L\left(L-\lambda\right)\left(1-\lambda L\right)} \\
&+ \left. \frac{\lambda\left(\left(1-\rho\lambda\right)g\left(L\right) - \left(1-\rho L\right)g\left(\lambda\right)\right)\begin{bmatrix} \tau_\varepsilon & \tau_\nu \end{bmatrix}}{\rho\left(1-\rho\lambda\right)\left(L-\lambda\right)\left(1-\lambda L\right)} \right\} \begin{bmatrix} z_t \\ x_{it} \end{bmatrix}.
\end{aligned}$$

G.3 Fixed Point

Substituting the forecast formulas into the best response function we obtain the following system

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

where

$$\mathbf{C}(L) \equiv \begin{bmatrix} 1 - \frac{\left(\alpha+\beta L^{-1}\right)}{\rho(L-\lambda)(1-\lambda L)} \frac{\left(-\lambda\rho(\tau_\nu+\tau_\varepsilon)L^2+(\rho\tau_\nu+\lambda\tau_\varepsilon+\lambda\rho(\lambda\tau_\nu+\rho\tau_\varepsilon))L-\lambda\rho(\tau_\nu+\tau_\varepsilon)\right)+\lambda L\tau_\varepsilon(\tau_\nu+\tau_\varepsilon)}{(\tau_\nu+\tau_\varepsilon)} & -\frac{\left(\alpha+\beta L^{-1}\right)\lambda L\tau_\varepsilon}{\rho(L-\lambda)(1-\lambda L)} \\ -\frac{\left(\alpha+\beta L^{-1}\right)}{\rho(L-\lambda)(1-\lambda L)} \frac{-\tau_\nu(\lambda-\rho)(L(\lambda\rho-1))+\lambda L\tau_\nu(\tau_\nu+\tau_\varepsilon)}{(\tau_\nu+\tau_\varepsilon)} & 1 - \frac{\left(\alpha+\beta L^{-1}\right)\lambda L\tau_\nu}{\rho(L-\lambda)(1-\lambda L)} \end{bmatrix},$$

$$\mathbf{D}(L) \equiv \frac{\gamma\lambda\begin{bmatrix}\tau_\varepsilon & \tau_\nu\end{bmatrix}'}{\rho(1-\lambda L)(1-\rho\lambda)} - \varphi_1\frac{(1-\rho L)\begin{bmatrix}\tau_\varepsilon & \tau_\nu\end{bmatrix}'}{(L-\lambda)(1-\lambda L)} - \varphi_2\frac{\begin{bmatrix}\tau_\varepsilon+\tau_\nu-(\lambda\tau_\nu+\rho\tau_\varepsilon)L & \tau_\nu(\lambda-\rho)L\end{bmatrix}'}{L(1-\lambda L)},$$

with

$$\varphi_1 \equiv \frac{\alpha\lambda+\beta}{\lambda}\left(\frac{\lambda(\lambda-\rho)h_1(\lambda)}{\rho(\tau_\nu+\tau_\varepsilon)} + \frac{\lambda^2 g(\lambda)}{\rho(1-\rho\lambda)}\right), \quad \text{and} \quad \varphi_2 \equiv \frac{\beta}{\tau_\nu+\tau_\varepsilon}h_1(0).$$

Using the fact that $\lambda + \frac{1}{\lambda} = \rho + \frac{1}{\rho} + \frac{\tau_\varepsilon+\tau_\nu}{\rho}$ to substitute out for τ_ε, $\mathbf{C}(L)$ simplifies to

$$\mathbf{C}(L) = \begin{bmatrix} 1-\alpha-\beta L^{-1} & -\frac{\alpha+\beta L^{-1}}{\rho(L-\lambda)(1-\lambda L)}\lambda L\tau_\varepsilon \\ 0 & 1-\frac{\alpha+\beta L^{-1}}{\rho(L-\lambda)(1-\lambda L)}\lambda L\tau_\nu \end{bmatrix}.$$

Next notice that the determinant of $\mathbf{C}(L)$,

$$\det(\mathbf{C}(L)) = \frac{-\lambda\left(L^2 - \left(\lambda+\frac{1}{\lambda}\right)L + 1 + \frac{\alpha L+\beta}{\rho}\tau_\nu\right)((1-\alpha)L-\beta)}{L(1-\lambda L)(L-\lambda)}$$

has two inside roots,

$$\omega_1 \equiv \frac{\rho\left(\lambda+\frac{1}{\lambda}\right) - \alpha\tau_\nu - \sqrt{\left(\rho\left(\lambda+\frac{1}{\lambda}\right)-\alpha\tau_\nu\right)^2 - 4(\rho+\beta\tau_\nu)\rho}}{2\rho}, \quad \text{and} \quad \omega_2 \equiv \frac{\beta}{1-\alpha},$$

and one outside root,

$$\omega_3 \equiv \frac{\rho+\beta\tau_\nu}{\rho}\frac{1}{\omega_1}.$$

Solving the system for $h_1(L)$ and $h_2(L)$ and choosing φ_1 and φ_2 to remove the inside poles of $h_1(L)$ at ω_1 and ω_2,[45] and rearranging we obtain the policy functions explicitly in terms of parameters,

$$h_1(L) = \frac{\psi}{\rho(\omega_3 - \rho)} \frac{\tau_\varepsilon}{\left(1 - \frac{1}{\omega_3} L\right)}, \quad \text{and} \quad h_2(L) = \frac{\psi}{\rho(\omega_3 - \rho)} \frac{\left((1-\alpha) - \frac{\beta}{\omega_3}\right)\tau_\nu}{\left(1 - \frac{1}{\omega_3} L\right)}$$

where

$$\psi \equiv \frac{\gamma}{1 - \alpha - \rho\beta}.$$

G.4 Modified Signal Process

Suppose, now, that agents receive the same public signals as before but the private signals are given by

$$\widetilde{x}_{it} = \theta_t + \widetilde{\nu}_{it}$$

where $\widetilde{\nu}_{it} \sim \mathcal{N}(0, \widetilde{\tau}_\nu^{-1})$ and

$$\widetilde{\tau}_\nu \equiv \left((1-\alpha) - \frac{\beta}{\omega_3}\right)\tau_\nu.$$

Notice that, applying the Wiener-Hopf prediction formula, analogously to above, we obtain

$$\widetilde{\mathbb{E}}_{it}[\theta_t] = \frac{\widetilde{\lambda}\begin{bmatrix}\tau_\varepsilon & \widetilde{\tau}_\nu\end{bmatrix}}{\rho\left(1 - \widetilde{\lambda}L\right)\left(1 - \rho\widetilde{\lambda}\right)} \begin{bmatrix} z_t \\ \widetilde{x}_{it} \end{bmatrix},$$

where

$$\widetilde{\lambda} \equiv \frac{1}{2}\left(\frac{\tau_\varepsilon + \widetilde{\tau}_\nu}{\rho} + \frac{1}{\rho} + \rho - \sqrt{\left(\frac{\tau_\varepsilon + \widetilde{\tau}_\nu}{\rho} + \frac{1}{\rho} + \rho\right)^2 - 4}\right).$$

G.5 Equivalence Result

Finally, notice that

$$\widetilde{\lambda} = \omega_3^{-1},$$

and, therefore,

$$\psi\widetilde{\mathbb{E}}_{it}[\theta_t] = \frac{\psi}{\rho(\omega_3 - \rho)} \frac{\begin{bmatrix}\tau_\varepsilon & \left((1-\alpha) - \frac{\beta}{\omega_3}\right)\tau_\nu\end{bmatrix}}{\left(1 - \frac{1}{\omega_3}L\right)} \begin{bmatrix} z_t \\ \widetilde{x}_{it} \end{bmatrix} = \begin{bmatrix} h_1(L) & h_2(L) \end{bmatrix} \begin{bmatrix} z_t \\ \widetilde{x}_{it} \end{bmatrix}.$$

[45] That is,

$$\varphi_1 = \frac{\gamma\rho\left((1-\alpha)\lambda - \kappa\right)(1 - \lambda\omega_3) + \gamma\kappa\lambda\tau_\nu}{\rho^2(1-\rho\lambda)(\omega_3 - \rho)(1 - \alpha - \rho\kappa)}, \quad \text{and} \quad \varphi_2 = \frac{\gamma\kappa\left((\lambda - \rho)(1 - \rho\lambda) + \lambda\tau_\nu\right)}{\rho(1-\rho\lambda)(\omega_3 - \rho)(1 - \alpha - \rho\kappa)(\lambda - \rho)}.$$

G.6 Forward and Backward Looking Best Response

Consider the following best response function where agents also care about future and past aggregate actions,

$$y_{it} = \gamma \mathbb{E}_{it}\left[\theta_t\right] + \alpha \mathbb{E}_{it}[y_t] + \beta \mathbb{E}_{it}\left[y_{t+1}\right] + \kappa \mathbb{E}_{it}\left[y_{t-1}\right].$$

Following a similar procedure to the one above we can solve for the equilibrium best response functions,

$$h_1\left(L\right) = \frac{\psi \tau_\varepsilon}{\left(1 - \frac{1}{\omega_3} L\right)\left(1 - \frac{1}{\omega_4} L\right)}, \quad \text{and} \quad h_2\left(L\right) = \frac{\psi\left(\kappa \omega_3 \omega_4 - \beta\right)\tau_\nu}{\omega_3\left(1 - \frac{1}{\omega_3} L\right)}$$

where

$$\psi \equiv \frac{\gamma\left(\rho + \kappa \tau_\nu\right)}{\left(\kappa \omega_4 - \rho \beta\right)\left(\rho + \kappa \tau_\nu\right)\rho \omega_3 - \rho^2 \kappa\left(\rho + \beta \tau_\nu\right)\omega_4 + \rho \beta\left(\rho^3 + \left(\rho\left(1 - \alpha\right) - \kappa\right)\tau_\nu\right)},$$

$$\omega_3 \equiv \frac{\left(1 - \alpha\right)\tau_\nu + \tau_\varepsilon + 1 + \rho^2 + \sqrt{\left(\left(1 - \alpha\right)\tau_u + \tau_\varepsilon + 1 + \rho^2\right)^2 - 4\left(\rho + \beta \tau_u\right)\left(\rho + \kappa \tau_\nu\right)}}{2\left(\rho + \kappa \tau_\nu\right)},$$

$$\omega_4 \equiv \frac{1 - \alpha + \sqrt{\left(1 - \alpha\right)^2 - 4\beta\kappa}}{2\kappa}.$$

Relative to a forecast of θ_t with the signal structure above (see equation (G.1)) there is an extra lag operator in the denominator of the response to the public signal, $h_1(L)$, here. Hence, the equilibrium actions cannot be represented by a forecast of θ_t by a modification of the precision of the shocks using the same information structure.

H Additional Examples

H.1 Multi-Action Example

In this section, we explore the effects of information frictions on the comovement between aggregate outcomes. To our knowledge, this paper is the first to explore this issue. The joint dynamics of multiple aggregate outcomes results from their intrinsic cross-dependence and from the degree of information frictions. Using the single-agent solution from Proposition 4, we obtain a clear characterization of these two forces. In particular, we show that increasing the degree of information frictions can flip the sign of the correlation between aggregate variables.

Consider an economy in which agents choose two actions simultaneously. For simplicity, assume that the two actions depend on the same aggregate fundamental, θ_t. Then, the agents' best response functions can be written as

$$y_{it}^1 = \mathbb{E}_{it}[\theta_t] + a_{11}\mathbb{E}_{it}[y_t^1] + a_{12}\mathbb{E}_{it}[y_t^2]$$
$$y_{it}^2 = \mathbb{E}_{it}[\theta_t] + a_{21}\mathbb{E}_{it}[y_t^1] + a_{22}\mathbb{E}_{it}[y_t^2]$$

where the matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \mathbf{Q} \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \mathbf{Q}^{-1},$$

summarizes the dependence on aggregate actions. The second equality represents the eigendecomposition of matrix $\mathbf{A}$, where α_1 and α_2 denote its eigenvalues and $\mathbf{Q}$ is a matrix composed of its eigenvectors. Without loss of generality, let ω_1 and ω_2, be such that

$$\mathbf{Q} = \begin{bmatrix} 1 & 1 \\ \frac{\omega_2}{\omega_1} & \frac{1-(1-\alpha_1)\omega_2}{1-(1-\alpha_1)\omega_2} \end{bmatrix}.$$

Notice that choosing α_1, α_2, ω_1 and ω_2 we can generate any matrix $\mathbf{A}$ that satisfies Assumption 4. The policy rule permits a simpler format when represented in terms of these parameters rather than the ones in the original matrix $\mathbf{A}$. Applying Proposition 4, it follows that

$$y_t^1 = \omega_1 \widetilde{\mathbb{E}}_t[\theta_t; \alpha_1] + \phi_1 \widetilde{\mathbb{E}}_t[\theta_t; \alpha_2],$$
$$y_t^2 = \omega_2 \widetilde{\mathbb{E}}_t[\theta_t; \alpha_1] + \phi_2 \widetilde{\mathbb{E}}_t[\theta_t; \alpha_2].$$

where ϕ_1 and ϕ_2 are functions of the primitives

$$\phi_1 \equiv \frac{1}{1-\alpha_2} - \frac{1-\alpha_1}{1-\alpha_2}\omega_1, \quad \text{and} \quad \phi_2 \equiv \frac{1}{1-\alpha_2} - \frac{1-\alpha_1}{1-\alpha_2}\omega_2.$$

This representation makes clear how the degrees of strategic complementarity affect each action. In particular, if $\alpha_1 = \alpha_2$, then $\widetilde{\mathbb{E}}_t[\theta_t; \alpha_1] = \widetilde{\mathbb{E}}_t[\theta_t; \alpha_2]$, and the actions are the same irrespective of how these forecasts are weighted. If $\alpha_1 \neq \alpha_2$, the behavior of the two actions does depend on the weights.

Next, we show in a numerical example how dispersed information can affect the relationship between the two actions. Suppose that the fundamental follows an AR(1) process

$$\theta_t = \rho\theta_{t-1} + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \tau_\eta^{-1}),$$

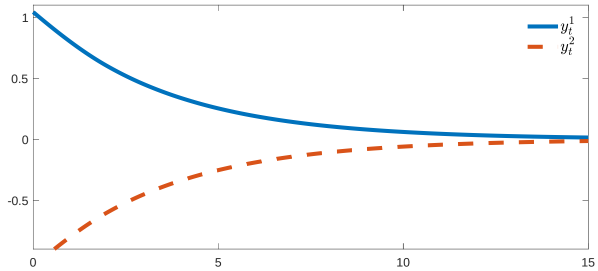
and that agents only receive a private signal about θ_t,

$$x_{it} = \theta_t + \nu_{it}, \qquad \nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1}).$$

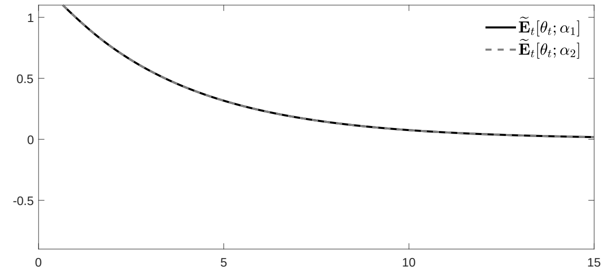
We set the eigenvalues to $\alpha_1 = -0.5$, and $\alpha_2 = 0.5$, so the precision of the α_1-modified signals are higher. We also set the eigenvector parameters to $\omega_1 = 0.6$, and $\omega_2 = 1.4$, which implies that $\phi_1 = 0.2$, and $\phi_2 = -2.2$. In the perfect information benchmark, both modified forecasts collapse to the fundamental itself, that is $\widetilde{\mathbb{E}}_t[\theta_t; \alpha_1] = \widetilde{\mathbb{E}}_t[\theta_t; \alpha_2] = \theta_t$. The aggregate outcomes are, then, uniquely pinned down by the fundamental and the parameters controlling strategic interactions. Explicitly,

$$y_t^1 = (\omega_1 + \phi_1)\theta_t, \quad \text{and} \quad y_t^2 = (\omega_2 + \phi_2)\theta_t.$$

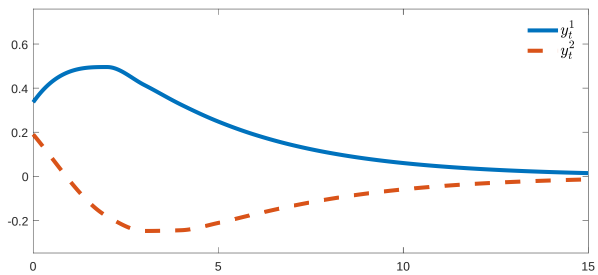
Under our parameterization, $(\omega_1 + \phi_1)$ and $(\omega_2 + \phi_2)$ have opposite signs, and therefore, the actions are



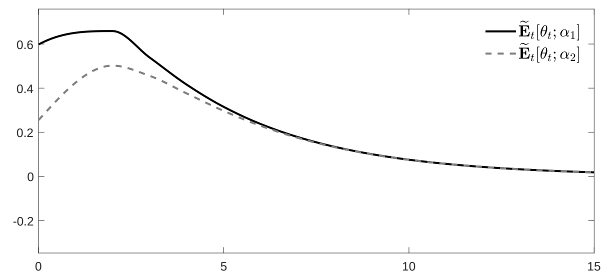
(A) Actions with $\tau_\nu \to \infty$



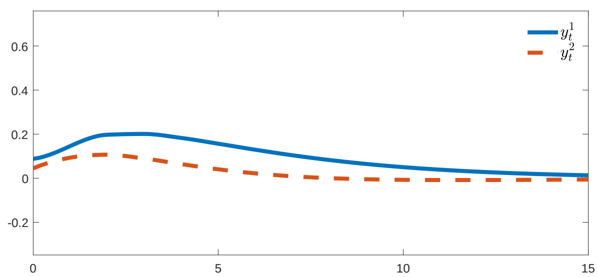
(B) α_j-Modified Forecasts with $\tau_\nu \to \infty$



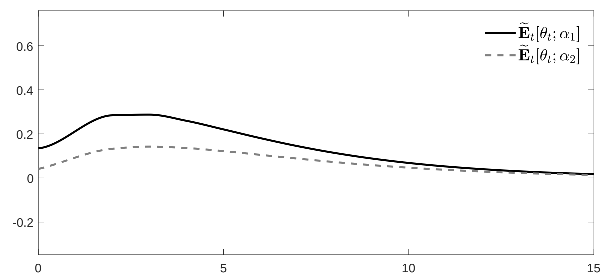
(C) Actions with $\tau_\nu = 1.0$



(D) α_j-Modified Forecasts with $\tau_\nu = 1.0$



(E) Actions with $\tau_\nu = 0.1$



(F) α_j-Modified Forecasts with $\tau_\nu = 0.1$

FIGURE 5: Multi-action Example

Parameters: $\tau_\eta = 1$, $\rho = 0.75$, $\omega_1 = 0.6$ and $\omega_2 = 1.4$, $\alpha_1 = -0.5$ and $\alpha_2 = 0.5$.

perfectly negatively correlated, as shown in Figure 5A. Figure 5C shows that, when $\tau_\nu = 1$, the two actions still move in different directions. The reason is that both $(1-\alpha_1)\tau_\nu$ and $(1-\alpha_2)\tau_\nu$ are large enough, so that $\widetilde{\mathbb{E}}_t[\theta_t; \alpha_1]$ and $\widetilde{\mathbb{E}}_t[\theta_t; \alpha_2]$ are still quite responsive and close to one another, as shown in Figure 5D.

When $\tau_\nu = 0.1$, the modified precisions are low enough yielding a $\widetilde{\mathbb{E}}_t[\theta_t; \alpha_2]$ close to zero; recall that $\alpha_2 = 0.5 > -0.5 = \alpha_1$. As a result, even though ϕ_1 and ϕ_2 have the opposite signs, the terms with $\widetilde{\widetilde{\mathbb{E}}}_t[\theta_t; \alpha_1]$ dominate and the aggregate actions comove. In fact, Figure 6 shows that the correlation between the two actions is decreasing in τ_ν monotonically, and switch from positive to negative.

This result—that the correlation between actions is increasing in the intensity of information frictions—depends on the particular specification of parameters in this example. However, it highlights the fact that dispersed information may significantly affect the joint behavior of interactive actions.

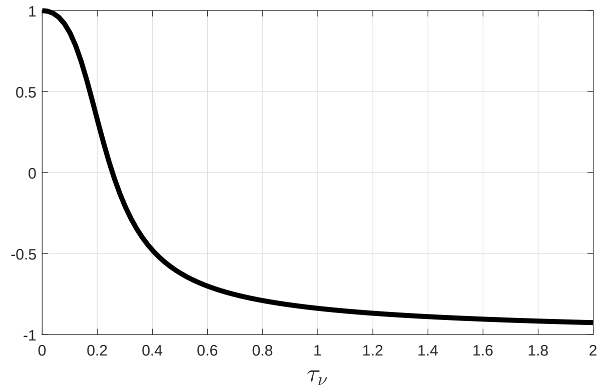


FIGURE 6: Correlation between y_t^1 and y_t^2 for different τ_ν

H.2 Endogenous Learning Example

Relative to Example 1, suppose agents also receive a signal about the aggregate action y_t, which is an endogenous object. Letting $\boldsymbol{\varepsilon}_{it} = [\theta_t,\, \nu_{it},\, \varepsilon_{it}]'$, the information structure can be represented within the framework of equation (6.1):

$$\mathbf{M}(L) = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{p}(L) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} x_{it} &= \theta_t + \nu_{it} \\ z_{it} &= y_t + \varepsilon_{it}. \end{aligned}$$

For simplicity, we also assume that all the shocks are i.i.d.: $\theta_t \sim \mathcal{N}(0, \tau_\theta^{-1})$, $\nu_{it} \sim \mathcal{N}(0, \tau_\nu^{-1})$, and $\varepsilon_{it} \sim \mathcal{N}(0, \tau_\varepsilon^{-1})$. This type of information structure is widely used in the literature on endogenous learning.

The aggregate action, y_t, must depend on the fundamental, θ_t—the only aggregate shock in this economy—so we conjecture that

$$y_t = \mathcal{H}\, \theta_t,$$

for some constant $\mathcal{H}$. Since agents do not internalize the effects of their own actions on others' signals, they take $\mathcal{H}$ as given. Hence, their forecasting problem, given $\mathcal{H}$, is equivalent to one in which their signals are generated by the following exogenous process:

$$\widehat{\mathbf{M}}(L) = \begin{bmatrix} 1 & 1 & 0 \\ \mathcal{H} & 0 & 1 \end{bmatrix} \qquad \Rightarrow \qquad \begin{aligned} x_{it} &= \theta_t + \nu_{it} \\ \widehat{z}_{it} &= \mathcal{H}\, \theta_t + \varepsilon_{it} \end{aligned}$$

With endogenous information, even in the absence of a primitive coordinating motive ($\alpha = 0$), agents implicitly coordinate via their signal processes. Using the single-agent solution, one can focus on the alternative simple forecasting problem in which all the coordination occurs via information. In the example at hand, the aggregate action can be written as

$$y_t = \int \widetilde{\mathbb{E}}_{it}[\theta_t] = \frac{(1-\alpha)\tau_\nu}{\tau_\theta + (1-\alpha)\tau_\nu + \mathcal{H}^2\,(1-\alpha)\tau_\varepsilon} \int x_{it} + \frac{\mathcal{H}\,(1-\alpha)\tau_\varepsilon}{\tau_\theta + (1-\alpha)\tau_\nu + \mathcal{H}^2\,(1-\alpha)\tau_\varepsilon} \int \widehat{z}_{it}.$$

At this stage, the equivalence result spares us the trouble of making an inference about y_t, and this policy rule already satisfies the first two equilibrium conditions in Definition 3. To make sure that the perceived law of motion for y_t is consistent with agents' signal processes, condition (6.3) must also be satisfied, which reduces to a cubic polynomial equation in terms of $\mathcal{H}$,

$$(1-\alpha)\tau_\varepsilon\,\mathcal{H}^3 - (1-\alpha)\tau_\varepsilon\,\mathcal{H}^2 + (\tau_\theta + (1-\alpha)\tau_\nu)\,\mathcal{H} - (1-\alpha)\tau_\nu = 0. \tag{H.1}$$

It is clear that there may exist multiple real solutions to equation (H.1), which correspond to multiple equilibria. The origin of this multiplicity lies in the self-fulfilling property of the signals' informativeness. For example, if all agents respond to the fundamental aggressively, if $\mathcal{H}$ is high, then the signal z_{it} is very informative. As a result, agents can learn more from the endogenous signal and indeed become more responsive.