Dynamic Information Aggregation: Learning from the Past

Huo, Z.; Pedroni, M.

DOI
10.2139/ssrn.3687529

Publication date
2020

Document Version
Final published version

Dynamic Information Aggregation: Learning from the Past*

Zhen Huo
Yale University

Marcelo Pedroni
University of Amsterdam

August, 2020

Abstract

In an environment with dispersed information, how much can agents learn from past endogenous aggregate outcomes such as prices or output? We show that, in a rational expectations equilibrium, two possible regimes can arise endogenously: a perfect revealing regime and a confounding regime. The economic fundamental can be perfectly inferred in the former but is only partially revealed in the latter. The confounding regime arises only when general equilibrium feedback effects are strong enough. Then, even when the past aggregate outcomes are perfectly observed, the effects of informational frictions are persistent over time. Furthermore, in the confounding regime, the aggregate outcomes do not permit a finite-state representation, and they display an initial underreaction relative to their perfect information counterparts followed by a delayed overreaction. In a standard New Keynesian model, we show that the confounding regime is more likely to arise under a dovish monetary policy rule.

Keywords: Information Aggregation, Dispersed Information, Monetary Policy
JEL classifications: D8, E3

*Huo: Yale University, 28 Hillhouse Ave, New Haven, CT, 06510, US, zhen.huo@yale.edu. Pedroni: University of Amsterdam, PO Box 15867, 1001 NJ Amsterdam, Netherlands, m.pedroni@uva.nl. Acknowledgements: We are thankful to Yan Xiong for discussing our paper in the 2019 PHBS Workshop in Macroeconomics and Finance. The paper benefits from comments from Marios Angeletos, Ryan Chahrour, Simas Kucinskas, Keith Kuester, Jeniffer La'O, Christian Stoltenberg, and Pengfei Wang, and an early discussion with Giacomo Rondina and Todd Walker. We would also like to thank the seminar participants at PHBS, University of Bonn, TI Macro Conference in Rotterdam. We thank Flint O'Neil for providing excellent research assistance. Lastly, we thank Zhao Han, Fei Tan, and Jieran Wu for generously sharing their computational results with us.

1. Introduction

In this paper, we revisit a number of classical questions in the information literature: how much can agents learn from endogenous aggregate outcomes about underlying fundamentals? Can monetary policy generate persistent real effects when the history of inflation is observed? How does the informativeness of aggregate statistics depend on the strength of the underlying general equilibrium (GE) effects? We explore these questions in economies with persistent information and strategic interactions between agents.

We consider a beauty-contest model in which agents care about some persistent economic fundamental and also an endogenous aggregate outcome. We assume agents receive two types of signals: private noisy signals about current fundamentals and the perfect observation of the history of past aggregate outcomes. The lack of observation of the current aggregate outcome captures the idea that firms and households often make their production and consumption decisions before contemporaneous aggregate statistics are available. We show that in a rational expectations equilibrium, two possible regimes can arise endogenously: a revealing regime and a confounding regime.[1] In the revealing regime, the history of outcomes perfectly aggregate information, and agents can infer the underlying fundamental without error. This is in line with the conventional wisdom that, even with dispersed information, prices or output help eliminate uncertainty about the aggregate fundamental. However, there also exists a confounding regime in which agents make persistent forecast errors, and consensus about the fundamental is never achieved.

Whether the confounding regime arises depends on the degree of strategic complementarity between agents which captures the strength of GE feedback effects, and on the precision of private signals about the fundamental. Endogenous outcomes aggregate information only if agents respond to their private signals which are direct sources of information. The more agents rely on past aggregate outcomes in their inference, the less information contained in private signals can be encoded into individuals' actions, resulting in potential information dilution. The extent to which agents rely on the history of aggregate outcomes is decreasing in the precision of private signals and increasing in the need to coordinate with other agents.[2] In an environment with persistent information, a sufficient amount of information dilution can make the process of the aggregate outcome non-invertible.

The dynamics of aggregate outcomes also differ in a significant way under these two regimes. In the revealing regime, in response to an innovation to the fundamental, the aggregate outcome underreacts on impact, since information about the current fundamental is dispersed. Afterwards, the aggregate outcome exactly follows the fundamental since, by observing past aggregate outcomes, agents reach common knowledge about the innovation. In contrast, in the confounding regime, past outcomes are not sufficient for agents to infer the underlying innovations, and agents' forecast

[1] The notion of confounding dynamics that results from a non-invertible process was first introduced in an early version of Rondina and Walker (2020).

[2] See a more detailed discussion on this point in Morris and Shin (2002), Angeletos and Pavan (2007), and Huo and Pedroni (2020).

errors are persistent over time. Still, the observation of past aggregate outcomes does not allow for the systematic under- or overestimation of the fundamental. As a result, the aggregate outcome fluctuates around the fundamental, alternating between overreaction and underreaction. Furthermore, we prove that the process for the aggregate outcome is complex in the sense that it does not permit a finite-state representation. In a monetary model, this implies that a money supply shock can have long-lasting real effects even when firms observe the history of inflation.

An implication of the two informational regimes is that there exists a "kink" on how the informativeness of aggregate outcomes is related to the structural parameters. Within the revealing regime, a local variation of parameters only affects the impact response of the aggregate outcome but does not change its response afterwards nor its invertibility. In response to a large enough shift in the parameters, the economy enters the confounding regime, in which case either a larger strategic complementarity or a lower signal precision decrease the informativeness of the aggregate outcome.

To highlight the mechanism, we consider a classical monetary model à la Woodford (2003) with the crucial departure that firms are allowed to observe past aggregate prices to infer monetary shocks. When pricing complementarities are sufficiently strong, prices fail to perfectly aggregate information, and monetary shocks can generate long-lasting real effects. Moreover, the inflation dynamics in this confounding regime can be empirically relevant. The impulse response function of inflation *forecasts* displays an initial underreaction and a delayed overreaction, which is consistent with the evidence on expectations documented in Angeletos, Huo, and Sastry (2020) and Kucinskas and Peters (2018).[3] The model also predicts waves of under- and overshooting of the inflation itself, similar to the identified responses from Christiano, Eichenbaum, and Evans (2005) and Altig, Christiano, Eichenbaum, and Linde (2011).

The aforementioned insights extend to various environments with multiple signals, forward-looking beauty-contest games, and multivariate systems, among which we highlight the following applications. First, we explore the effect of adding additional public signals, and we show it does not necessarily make the economy more likely to be in the revealing regime. On one hand, more signals tend to alleviate informational frictions, which makes the revealing regime more likely to arise. On the other hand, public signals crowd out reliance on private signals, which reduces the information aggregated by aggregate outcomes. These two effects may exactly cancel out.

Second, we show that with dispersed information, the confounding regime is more likely to arise with a lower nominal rigidity in the New Keynesian Phillips curve and a higher marginal propensity to consume (MPC) in the dynamic IS curve. With more flexible prices, those firms that can adjust their prices need to pay more attention to others' prices; with a higher MPC, consumers are more sensitive to the aggregate demand. Both of these map to stronger GE feedback effects, and hence lead to non-invertible aggregate outcomes.

When the central bank employs a Taylor-type monetary policy, a more aggressive response of

[3]In Angeletos, Huo, and Sastry (2020), this pattern is rationalized by imposing a behavioral bias—over-extrapolation—and we show it can be a natural result of imperfect information aggregation.

nominal interest rate to inflation reduces the positive feedback effects between the supply and demand block, resulting in a lower degree of strategic complementarity. With incomplete information, a more aggressive policy not only modifies the magnitude of the economy's response to shocks, but it can also alter the entire dynamic pattern by shifting the economy from the confounding to the revealing regime. It follows that a dovish monetary policy rule tends to induce hump-shaped and oscillatory responses of inflation and output.

Related Literature

The study of how well endogenous variables like prices summarize relevant dispersed information dates back to Hayek (1945). What makes information aggregation nontrivial in our model is the fact that fundamentals and signals are correlated over time. Our work therefore complements the line of research that focuses on endogenous information aggregation that relies on the history of signals, including, for instance, Townsend (1983), Sargent (1991), Kasa (2000), Nimark (2014), and Bacchetta and Wincoop (2006). This paper is closely related to Kasa, Walker, and Whiteman (2014) and Rondina and Walker (2020) who also explore non-invertible equilibrium dynamics. In Kasa, Walker, and Whiteman (2014), the non-invertibility is due to the underlying fundamental itself being non-invertible, and in Rondina and Walker (2020), a prerequisite is that the endogenous signal is non-invertible when information is perfect. Besides providing a general condition on the determinant of equilibrium invertibility, our work differs from theirs in two ways: First, in our environment, endogenous variables always perfectly aggregate information in the absence of informational frictions, and the aggregation is imperfect only when information is dispersed.[4] Second, the equilibrium in their models follows a tractable finite-state process. In our paper, the equilibrium dynamics is more involved and we prove that it does not permit a finite-state representation, even with a square information structure.

This paper is also related to the large literature on monetary non-neutrality due to informational frictions. Lucas (1972, 1973) showed this is consistent with rational expectations. In that setup, real effects of monetary shocks were predicated on variations in aggregate nominal expenditure not being forecastable. It follows that, if data on past aggregates were available these effects could, at most, be transitory. Woodford (2003), following Sims (2001), argues that the main informational bottleneck is not its plain availability but the "limited capacity of private decision-makers to pay attention" to it. Then, assuming aggregates are never fully observed, he shows that monetary shocks can have persistent effects. One could also generate persistent effects by including additional exogenous aggregate shocks in the model or imposing rational inattention, as in Nimark (2008), Adam (2007), Maćkowiak and Wiederholt (2009), and Melosi (2017), or endogenous sentiment shocks, as in Acharya, Benhabib, and Huo (2017). In this paper, we show that even if past aggregates are perfectly observed,

[4]This is due to the fact that aggregate variables are observed with a lag in our environment, which allows the invertibility of aggregate variables to be endogenous to informational frictions.

monetary policy can have long-lasting effects which arise endogenously only if GE effects are strong enough.[5]

A common assumption in the literature is that information is static or the underlying states are perfectly revealed after one period, which allows one to focus on the within-period inference problem. Grossman (1976) and Hellwig (1980) studied to what extent prices can summarize multidimensional information in a static setting. Messner and Vives (2005), Angeletos and Pavan (2009), Amador and Weill (2010, 2012) and Vives (2014, 2017) study learning externalities associated with the observation of endogenous aggregates. A particular type of learning externality plays an important role in deviations from perfect information aggregation in our model: strategic complementarity leads agents to respond less to their private information, which has the negative side effect of reducing the informativeness of aggregates in general and prices in particular. In persistent information settings as the one considered in our paper, this crowding out can lead to non-invertibility and oscillatory dynamics. Angeletos, Iovino, and La'O (2020) explore optimal policy under this type of learning externality. They find that agents are better off when incentivized to respond more aggressively to their belief variations, which helps to achieve better information aggregation. Gaballo (2018) and Chahrour and Gaballo (2019) show that learning from prices can play either a propagating or a mitigating role. Their mechanism relies on agents receiving different types of signals, while our framework focuses on the classical case with a common information structure.

The rest of the paper is organized as follows: Section 2 sets up a simple monetary model with informational frictions. Section 3 establishes the main result about the two regimes of price dynamics. Section 4 introduces noisy price observation which allows us to bridge the predictions of the model with the ones from its exogenous-information counterpart. Section 5 extends the main results to more general information and payoff structures, and explores various applications. Section 6 concludes.

2. A Monetary Model

In this section, we introduce a simple model of monopolistic competition to illustrate the basic idea of information aggregation through endogenous actions. The model is largely borrowed from Woodford (2003)[6] and it is designed not to exhaust the implications of our main results but to be the simplest setup in which we can highlight the role of strategic complementarity in determining the information content of endogenous information. It allows us to examine the question of how efficiently prices and other macro variables aggregate the information that firms need to make their decisions.

Suppose the producer of good i chooses the current and future prices of their good, $\{P_{it+k}\}$, to

[5] An exception is Hellwig and Venkateswaran (2015). They identify special cases in which dispersed information is irrelevant for allocation.

[6] The main difference relative to Woodford (2003) is that we consider a more general information structure.

maximize the expected present value of future (real) profits

$$\mathbb{E}_{i,t}\left[\sum_{k=0}^{\infty}\beta^k m(Y_{t+k})\left[\frac{P_{it+k}}{P_{t+k}}Y_{i,t+k}-C(Y_{i,t+k},Y_{t+k})\right]\right]$$

subject, in every period $t+k$, to the demand function

$$Y_{i,t+k}=Y_{t+k}\left(\frac{P_{i,t+k}}{P_{t+k}}\right)^{-\theta}, \tag{2.1}$$

for some $\theta>1$. The producer takes the Dixit-Stiglitz index of real aggregate demand, Y_t, and the corresponding price index, P_t, as given. The real cost of producing is given by $C(Y_{it},Y_t)$ and depends not only on the amount produced by the firm, Y_{it}, but also on the aggregate output, Y_t, to account for its effect on factor prices. The firm weights different states by the stochastic discount factor, $m(Y_t)$, so that the expected discounted profits can be interpreted as a financial market valuation of the firm.

To focus on informational frictions we abstract from price frictions here. The firms can choose their prices every period independent of the past. Thus, in period t, firm i chooses P_{it} to maximize

$$\mathbb{E}_{it}\left[Y_t\left(\frac{P_{it}}{P_t}\right)^{1-\theta}-C\left(Y_t\left(\frac{P_{it}}{P_t}\right)^{-\theta},Y_t\right)\right],$$

where $\mathbb{E}_{it}$ denotes the expectation conditional on the information set of firm i in period t. The optimal pricing decision of firm i implies a first order condition which can be log-linearly approximated by

$$p_{it}=\mathbb{E}_{it}[(1-\alpha)y_t+p_t],$$

where lower-case variables denote log-deviations from the full-information steady-state versions of the corresponding upper-case variables, which satisfy $P_{it}=P_t$ and $Y_t=Y^*$. The parameter α can be written as a function of deep parameters.[7]

Changes in nominal aggregate demand, $Q_t=P_tY_t/Y^*$, can be decomposed into changes in the price level and in the real output,

$$q_t=p_t+y_t.$$

We assume that the nominal aggregate demand is determined exogenously by the central bank, following an AR(1) process[8]

$$q_t=\rho q_{t-1}+\eta_t,\quad \eta_t\sim\mathcal{N}(0,1).$$

Therefore, though the innovations to the nominal aggregate demand, η_t, can have a broader interpretation, from now on, we refer to them as monetary shocks.

[7]More specifically, Y^* solves $(1-\theta)+\theta C_y(Y^*,Y^*)=0$, and $\alpha\equiv 1-\frac{C_{yy}(Y^*,Y^*)Y^*+C_{yY}(Y^*,Y^*)Y^*}{C_y(Y^*,Y^*)+\theta C_{yy}(Y^*,Y^*)}$.

[8]We relax this assumption in Section 5.

From an individual firm's perspective, their pricing decision can be expressed as the best-response function of a standard static beauty-contest game, that is

$$p_{it} = (1-\alpha)\mathbb{E}_{it}[q_t] + \alpha\mathbb{E}_{it}[p_t], \quad \text{with} \quad p_t = \int p_{it}, \tag{2.2}$$

and where $\alpha \in (-1, 1)$ controls the degree of strategic complementarity between firms, or the strength of GE feedback effects—we use both terms interchangeably.

With incomplete information, the aggregate price depends not only on the firms' first-order expectations about the nominal aggregate demand but also on higher-order expectations as emphasized in Morris and Shin (2002) and Woodford (2003). We formalize this in the following lemma.

Lemma 1. *The aggregate price is given by*

$$p_t = (1-\alpha)\sum_{k=0}^{\infty}\alpha^k\overline{\mathbb{E}}_t^{k+1}[q_t], \tag{2.3}$$

where the higher-order expectations are defined recursively as follows

$$\overline{\mathbb{E}}_t^0[q_t] \equiv q_t, \quad \textit{and} \quad \overline{\mathbb{E}}_t^{k+1}[q_t] \equiv \int \mathbb{E}_{it}\left[\overline{\mathbb{E}}_t^k[q_t]\right].$$

The real effects of a monetary shock on output are captured by the gap between p_t and q_t. Condition (2.3) implies that this gap depends on the dynamics of all higher-order expectations, which in turn depend on the information structure faced by firms. To proceed, we first consider a benchmark case with dispersed but *exogenous* information, in which the information content is independent of equilibrium objects. In the next section, we turn to the *endogenous* information case, in which the information content is a function of equilibrium objects.

Exogenous Information Benchmark. Suppose that every period, t, firm i receives a new private signal, x_{it}, about the nominal aggregate demand[9]

$$x_{it} = q_t + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \tau^{-1}). \tag{2.4}$$

With $\tau \to \infty$, we return to the frictionless case where firms observe nominal aggregate demand, q_t, perfectly. Not only is there no first-order uncertainty about the fundamental, but also all higher-order expectations become common knowledge and collapse to the first-order expectation, that is, $\overline{\mathbb{E}}_t^k[q_t] = q_t$ for all $k \geq 0$. It follows that

$$p_t^* = q_t, \qquad y_t^* = 0. \tag{2.5}$$

[9]This information structure is similar to those used in Woodford (2003) and Angeletos and La'O (2010).

Thus, prices vary with the nominal aggregate demand in a one-to-one fashion, leaving aggregate output unchanged.

With $\tau < \infty$, information is incomplete, and the equilibrium outcome is shaped by the complicated dynamics of all the higher-order expectations in condition (2.3). We can, however, bypass this complexity by applying the single-agent result in Huo and Pedroni (2020), which directly yields a simple characterization of the price dynamics.

Proposition 2.1. *Under exogenous information, aggregate prices follow an AR(2) process,*

$$p_t = \left(1 - \frac{\vartheta}{\rho}\right) \frac{1}{1 - \vartheta L} q_t, \tag{2.6}$$

where $\vartheta \in (0, \rho)$ *is given by*

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

The aggregate price moves less than the nominal aggregate demand, captured by the term $1 - \vartheta/\rho$ in equation (2.6), which implies that real output co-moves with its nominal counterpart. In addition, it takes time for firms to learn the monetary shock; and the deviation of prices from the nominal aggregate demand is persistent, captured by the term $1/(1 - \vartheta L)$. This leads to long-lasting effects of monetary shocks on real output.[10]

3. Endogenous Information Aggregation

In this section, we allow the signals in firms' information set to include endogenous aggregate variables, so that the informativeness of the information is determined in equilibrium. We show that this change can lead to significantly different price and output dynamics, and affects how much can be learned from observing aggregate variables.

Information Structure. We assume that in any period t, signals in firms' information set include

1. private exogenous signals, $\{x_{i,t}, x_{i,t-1}, \ldots\}$ as in (2.4);
2. past prices, $\{p_{t-1}, p_{t-2}, \ldots\}$.

The first set of signals makes information dispersed as mentioned before, while the second set of signals allow agents to better coordinate and learn from each others' actions.

There are two main reasons why we assume agents do not observe the contemporaneous aggregate outcomes but rather their delayed values: First, from a practical point of view, firms and households

[10]In Angeletos and Huo (2018), aggregate outcomes follow similar dynamics in a forward-looking game, where the determination of ϑ is more complicated.

typically need to plan their production or purchasing with a period of time in advance of the realizations the fundamentals like demand shocks or income. There are lags in the publication of aggregate statistics as well, and nowcasts are necessary for most aggregate variables probably with the exception of frequently traded asset prices. Second, from a theoretical point of view, the simultaneity of making an individual decision while observing the aggregate outcomes is easier to be justified when the aggregate states are perfectly observable and agents can form rational expectations about the aggregate variables without frictions. In our economy with dispersed information, observing past outcomes is a more natural timing of decisions as aggregate states are observed with noise. Nevertheless, the period length is a priori not determined, so one could consider the smallest of lags and the analysis that follows still applies.

To measure the amount of information aggregated by price, we define price informativeness as

$$\chi \equiv 1 - \frac{\text{Var}\left[q_t - \mathbb{E}\left[q_t | p^t\right]\right]}{\text{Var}[q_t]}.$$

Notice that $\chi \in [0, 1]$. When prices perfectly reveal the nominal aggregate demand, the mean squared prediction error is zero and price informativeness reaches its maximum level, $\chi = 1$. When prices are not informative at all, the forecast $\mathbb{E}\left[q_t | p^t\right] = 0$ and $\chi = 0$.

3.1 Invertibility

We start with a thought experiment in which only a single firm can observe past prices. As this firm is infinitesimal, the aggregate price process is the same as in the exogenous information case. How much can this firm learn from the history of prices? To answer this question, we rearrange condition (2.6), which leads to

$$q_{t-1} = \left(1 - \frac{\vartheta}{\rho}\right)^{-1} (p_{t-1} - \vartheta p_{t-2}). \tag{3.1}$$

That is, by observing the history of prices, the uncertainty about *past* shocks is completely eliminated. Price informativeness is, then, at its maximal level, $\chi = 1$.

This is the basic logic behind the conventional wisdom that prices can effectively aggregate information and facilitate allocative efficiency. In the context of monetary neutrality, it would imply that when firms observe past prices, the monetary shock can only have transitory effects. However, this logic does not necessarily extend to the case in which all firms can observe past prices, since the different information structure leads to different equilibrium price dynamics. Importantly, the mapping from prices to underlying shocks obtained in condition (3.1) may not be possible. What determines if a price process is able to perfectly aggregate information is its invertibility, as we formalize in the next lemma.

Lemma 2. *A price process $p_t = g(L)\eta_t$ is invertible if $g(L)$ does not contain any inside root.*[11] *When invertible, prices perfectly aggregate information about the underlying shock, i.e.,*

$$\mathbb{E}[q_{t-k}|p^t] = \frac{g(L)^{-1}p_{t-k}}{1-\rho L} = q_{t-k}, \quad \text{for } k \geq 0. \tag{3.2}$$

Corollary 1. *With exogenous information, for all admissible parameters (α, τ, ρ), the price process is always invertible and the history of prices contains the same information.*

The aforementioned exogenous information case is an example of an invertible process. Perhaps somewhat surprisingly, even with very large noise (small τ), the history of prices still reveals all the information about the monetary shock, even though the magnitude of the price response can be arbitrarily small.

On the other hand, the following hypothetical price process is non-invertible:[12]

$$p_t = (L-\lambda)q_t, \quad \text{with } |\lambda| < 1. \tag{3.3}$$

Note that, if we attempt to use the formula in equation (3.2) to forecast nominal aggregate demand, we would run into the problem that $g(L)^{-1}$ contains L with negative powers.[13] Thus, to apply the formula we would need to use future realizations of prices which are not in the firms' information sets. It is important to note, however, that the actual equilibrium price dynamics share some similarities with this example but can be significantly more involved.

3.2 Two Regimes of Information Aggregation

When information is endogenous and dispersed, the extent to which prices efficiently aggregate information hinges both on the severity of informational frictions and, crucially, on the degree of strategic complementarity.

Proposition 3.1. *The equilibrium with endogenous information exists. The equilibrium price, p_t, perfectly aggregates information, that is, $\mathbb{E}[q_t|p^t] = q_t$, if and only if the triple (α, τ, ρ) is in the invertible region that satisfies*

$$\alpha < 1 - \frac{\rho}{\tau}. \tag{3.4}$$

Proof. See Appendix B. □

This proposition partitions the parameter space into two regions: the invertible region in which prices fully reveal the underlying fundamental, and the non-invertible region in which they do not. In the invertible region, p^t contains the same information as the nominal aggregate demand q^t,

[11] The function $g(L)$ is said to have an inside root if there exists a complex number, z, inside the unit circle of the complex plane, such that $g(z) = 0$.

[12] This example is similar to the ones used in Rondina and Walker (2020) and Acharya, Benhabib, and Huo (2017).

[13] To see this, notice that, if $|\lambda| < 1$, then $(L-\lambda)^{-1} = L^{-1}\sum_{k=0}^{\infty}(\lambda^{-1}L)^{-k}$.

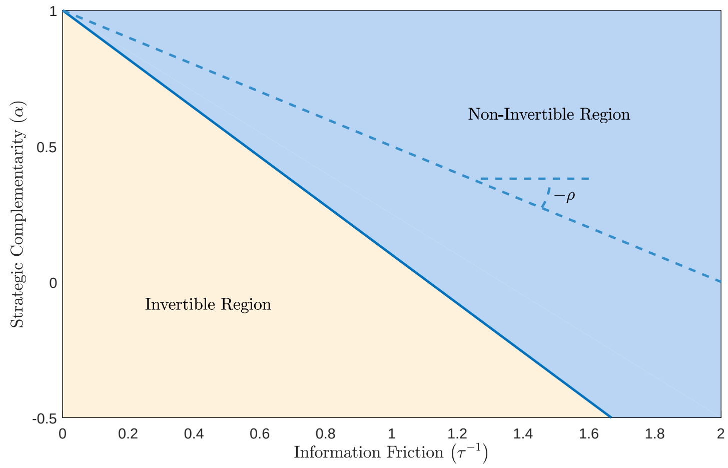


Figure 1: Two Regions of Price Informativeness

The solid line corresponds to $\rho = 0.9$ and the dashed line to $\rho = 0.5$.

and therefore we call this the *perfect revealing regime*. In the non-invertible region, p^t contains less information than q^t and firms are no longer able to infer the monetary shocks perfectly, which we refer to as the *confounding regime*.

As shown in Figure 1, the invertible region features relatively mild degrees of informational frictions and strategic complementarity. The opposite prevails in the non-invertible region. To understand this partition of the parameter space, we first go through a more "mechanical" explanation, and then we provide a more intuitive argument. Consider the impulse response function of prices to a monetary shock, and suppose that the process of p_t is invertible. We now explore under which conditions this conjecture can be verified. In period $t = 0$, firms observe the price history except for the current one. Since it is common knowledge that the monetary shocks in the past are equal to zero, to infer the current monetary shock firms rely solely on their private signal, $x_{i,0}$. The aggregate price level in this period, p_0, is determined the same way as in a static beauty-contest game with exogenous information. In period $t = 1$, firms observe p_0 and, by invertibility, q_0 becomes common knowledge. It follows that, in periods $t \geq 1$, prices are perfectly in line with the nominal aggregate demand, $p_t = q_t = \rho^t q_0$. Taking stock, the law of motion of price is given by

$$g(L) = \underbrace{p_0}_{\text{impact effect}} + \frac{\rho L}{1 - \rho L}, \tag{3.5}$$

where the first term, p_0, captures the impact effect, and the second term captures the response in the following periods when the monetary shock effectively becomes public.

The analysis above is based on the assumption that the process of p_t is invertible. To examine whether this is indeed the case, we need to check if any of the roots of $g(L)$ are inside the unit circle, which is true if and only if

$$p_0 = \frac{(1-\alpha)\tau}{1+(1-\alpha)\tau} > \frac{\rho}{1+\rho}. \tag{3.6}$$

That is, prices perfectly aggregate information if and only if the initial response is large enough. It is exactly when the informational friction is relatively small (large τ) and the coordination motive is relatively weak (small α) that the impact response is sizable. The former increases the responsiveness of first- and higher-order expectations, and the latter decreases the weight on generally more inertial higher-order expectations.

Intuitively, the information contained in prices ultimately comes from each individual firm's private information. The more firms rely on their private signals when making decisions, the more information can be potentially encoded in prices. With more private signal noise or higher degree of complementarity, firms put more weight on information in the public domain which reduces the usage of private signals. This intuition is reminiscent of the one in Amador and Weill (2010), where an increase of the public signal's precision may actually lower the information content of prices as the information contained in private signals is utilized less than a social planner would prefer.

This intuition is subject to the following caveat. Recall that, with exogenous information, independent of the informational friction and coordination motive, the price process is always invertible. With endogenous information, invertibility is obtained in a subset of the parameter space. This difference highlights that the informativeness of prices is affected not only by their overall responsiveness but also by their particular dynamic pattern. The importance of *dynamic* information aggregation also manifests itself via condition (3.4): with $\rho = 0$, information is always perfectly aggregated. Moreover, a higher ρ moves the fundamental process away from a static i.i.d. and increases the non-invertible region, as can be seen in Figure 1.

Proposition 3.1 already indicates that varying the degree of strategic complementarity affects the invertibility of prices. Figure 2 shows how price informativeness changes with the complementarity. A higher α diverts agents from the use of private information, which tends to lower the price informativeness. In the non-invertible region, this is indeed the case. Woodford (2003) emphasizes that higher strategic complementarity results in more inertia in inflation as higher-order expectations play a more important role. Our results imply that, when information is endogenous, there is an additional information channel through which the degree of complementarity shapes the aggregate outcome, a point we further elaborate on in Section 4.

3.3 Price Dynamics

We have characterized the invertible region in which prices perfectly aggregate information and the corresponding price dynamics. What does the stochastic process for prices look like if the parameters are in the non-invertible region? The following proposition shows that the price dynamics become significantly more complex.

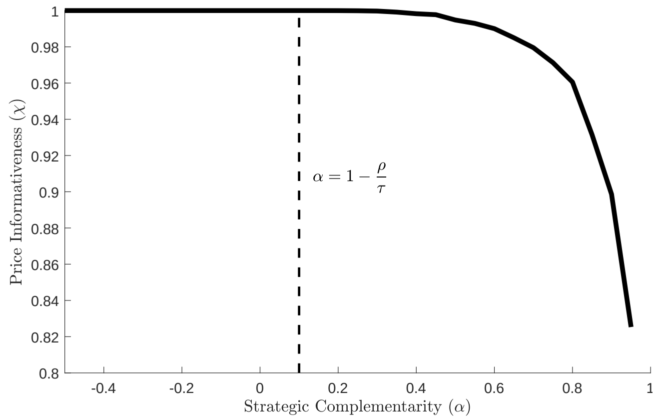


Figure 2: Price Informativeness

Parameters: $\rho = 0.9$, and $\tau = 1$.

Proposition 3.2. *1. In the invertible region, the law of motion of p_t is $p_t = q_t - \frac{1}{1+(1-\alpha)\tau}\eta_t$.*

*2. In the non-invertible region, the law of motion of p_t does **not** have a finite-state representation.*

Proof. See Appendix C. □

In the invertible region, the law of motion for prices differs from the nominal aggregate demand only at the impact response, following an ARMA(1,1) process. To visualize the process, the red dashed line in Figure 3 plots the impulse response of prices, p_t, with $\alpha = 0.1$, which is on the edge of the invertible region (we set $\rho = 0.9$, and $\tau = 1$). Except for the first period, prices are equal to nominal aggregate demand. Varying the parameters within the invertible region only changes the initial response. Recall that real output is equal to the difference between the nominal aggregate demand and the price level, $y_t = q_t - p_t$. So, in this case, the effect of the shock on real output is only transitory. This is in line with the conventional wisdom that the effects of a monetary shock to real aggregates are transitory when prices become public.

In the non-invertible region, the law of motion for p_t does not follow any finite ARMA process. In the literature, when the number of signals is the same as the number of shocks, one can typically obtain the equilibrium law of motion with a finite-state representation (see Kasa (2000), Kasa, Walker, and Whiteman (2014), Rondina and Walker (2020) for example). In contrast, in our environment, where the fundamental is invertible and the non-invertibility of equilibrium prices is obtained endogenously when the degree of strategic complementarity is high enough, no stochastic process with a finite number of inside roots can be supported as an equilibrium.

In Figure 3, the blue solid line shows the impulse response for p_t with $\alpha = 0.9$, which is in the non-invertible region. The response of p_t is initially dulled, and builds up gradually. Different from the invertible case, it does not coincide with the nominal demand, q_t, after period $t = 1$. Instead, it oscillates around q_t. This pattern is the result of the imperfect dynamic information aggregation. To understand this oscillation pattern, it is useful to consider the following two observations: First,

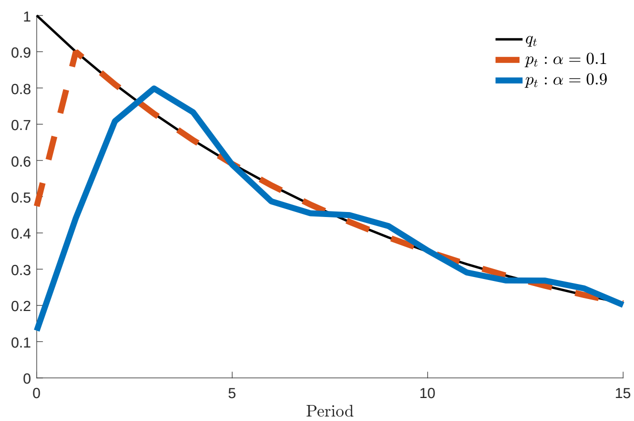


Figure 3: Response of Aggregate Price to Monetary Shock

Parameters: $\rho = 0.9$, and $\tau = 1$.

recall from Lemma 2 that prices are essentially the average forecasts about the fundamental, which cannot systematically deviate from the nominal demand. For example, with exogenous information, the impulse response of price is uniformly below that of the nominal aggregate demand, and this implies a persistent underestimation of the monetary shock and past prices. However, this type of persistent underestimation of past prices does not square with the fact that past prices are already part of agents' information set. Second, prices cannot be precisely equal to the nominal aggregate demand, as this would make them perfectly informative, contradicting the assumed non-invertibility of the price process. As a result, the only possible scenario is the waves of optimism and pessimism that we see here.

A direct implication of Proposition 3.2 is that the monetary shocks can induce persistent real output fluctuations. To allow long-lasting effects of monetary shocks on output, it is necessary that prices do not fully reveal the underlying shock. Woodford (2003) directly assumes that past prices cannot be observed, which can be rationalized by rational inattention, as in Maćkowiak and Wiederholt (2009). Another way to avoid perfect revelation is to add additional exogenous aggregate shocks, as in Nimark (2008) and Melosi (2017), or to add endogenous sentiment shocks, as in Acharya, Benhabib, and Huo (2017). Our results differ from those in the literature. The non-invertibility of the price process is an endogenous equilibrium outcome, and whether it is indeed the case depends on the underlying parameters. Prices efficiently aggregate information only when individual firms' responses to their *private* signals are sufficiently strong. When such response is relatively weak, monetary shocks have persistent real effects even if past prices are perfectly observed.

We conclude this section with a remark on uniqueness of equilibrium. In general, multiple equilibria may arise when information is endogenous (see Gaballo (2018) and Acharya, Benhabib, and Huo (2017) for example). In our setting, the equilibrium is unique in the invertible region, which we show by construction. Though we cannot rule out multiplicity in the non-invertible region, we think it is not likely for the following reason: the common cause of multiplicity is the strong complementarity between the responsiveness of the current action and the informativeness of endogenous signals. Since we assume that the aggregate outcome is observed with a one-period delay, the informativeness of newly arrived signals does not depend on agents' contemporaneous actions.[14] Lastly, regardless of whether the equilibrium is unique or not, all the theoretical results and analyses we have derived remain true.

4. Imperfect Price Observation

In this subsection, we explore to which extent previous results remain true when prices are not perfectly observed. The imperfect observation could arise when prices are recorded with measurement errors, or firms perceive published prices in an imperfect way due to costly contemplation. We show that the exogenous information case in Section 2 and the endogenous information case with perfect price observation in Section 3 serve as two useful benchmarks: the outcomes of the noisy price observation case live in the middle of these two extremes. In addition, we suggest a potential empirical footprint for endogenous information aggregation.

4.1 Noisy Prices: An Intermediate Case

Suppose firms observe past prices with idiosyncratic noise ν_{it},

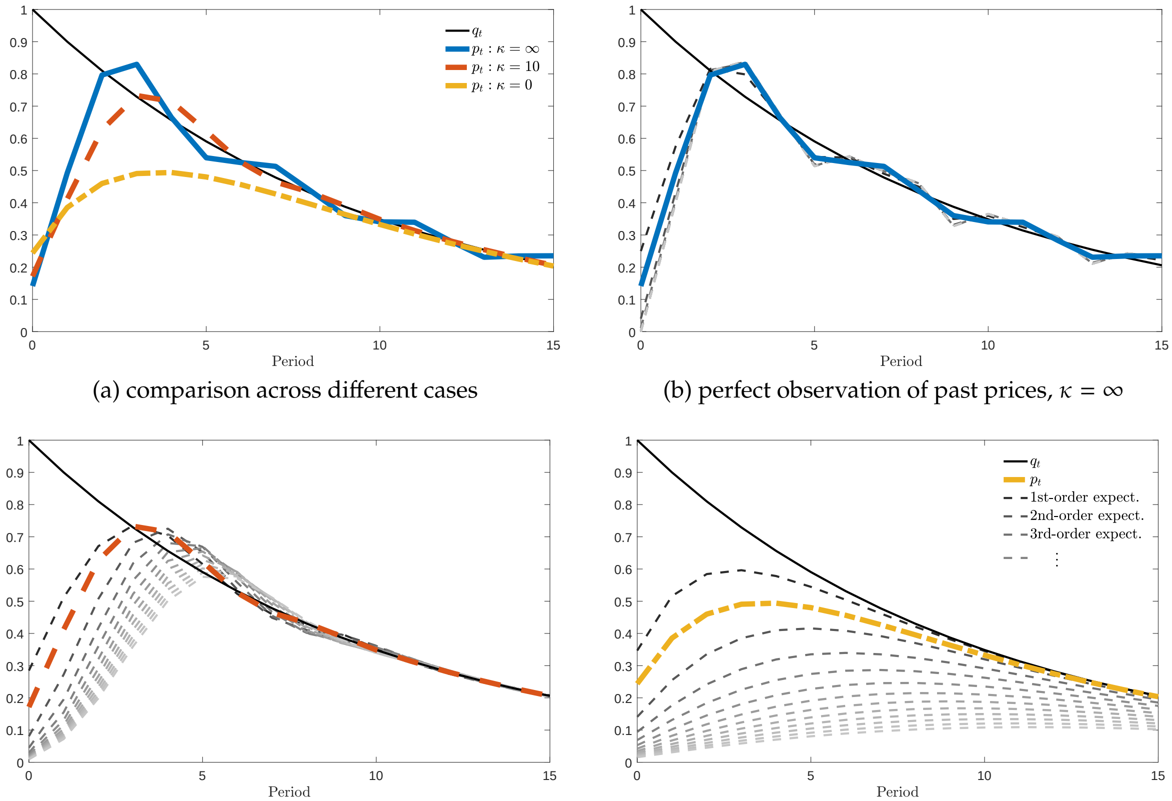
$$z_{it} = p_{t-1} + \nu_{it}, \quad \nu_{it} \sim \mathcal{N}(0, \kappa^{-1}). \tag{4.1}$$

The information set for firm i, then, becomes $\mathcal{I}_{it} = \{x_i^t, z_i^t\}$. This information structure nests the two benchmarks we have studied: by setting κ to infinity, we return to the case where prices are perfectly observed; by setting κ to zero, we obtain the other extreme where firms only have access to exogenous information, $\{x_i^t\}$. In the intermediate range, information is still endogenous as the informativeness of the price signal, z_{it}, is determined endogenously in equilibrium.[15]

We first show how the price dynamics changes when we vary the noise to the price observation. In Panel (b) of Figure 4, the blue solid line corresponds to the case in which prices are perfectly observed, $\kappa \to \infty$, and the parameters lie in the non-invertible region. When the magnitude of the private noise is positive but relatively small, by continuity, the price dynamics is similar to the limiting case. The red dashed line in Panel (c) confirms this logic: with $\kappa = 10$, it resembles the pattern of the blue solid line though the oscillation around the fundamental is less pronounced. When $\kappa = 0$ (the exogenous information case), there is no oscillation anymore. The yellow line in Panel (d) is always below the fundamental. In summary, our analysis in the previous two sections can serve as two baseline cases.

[14]If there is an initial period, it follows that the equilibrium is always unique no matter where it converges to. Moreover, though this is certainly not definitive, our numerical algorithm always yields a unique solution.

[15]To solve the intermediate case, one has to resort to numerical methods as, generally, no finite-state solution is available. See Huo and Takayama (2017) for a more detailed discussion.



(a) comparison across different cases

(b) perfect observation of past prices, $\kappa = \infty$

(c) noisy observation of past prices, $\kappa = 10$

(d) no price observation, $\kappa = 0$

Figure 4: Response of Aggregate Prices Varying Precision of Price Observation (κ)

Parameters: $\rho = 0.9$, $\alpha = 0.5$, and $\tau = 0.25$.

As κ changes from infinity to zero, the properties of the price dynamics shift from one extreme to the other.

A notable implication of dispersed information is that higher-order expectations differ from first-order expectations. Depending on the information structure, higher-order expectations display quite different patterns, as shown in Panels (b) to (d) in Figure 4. With only private exogenous information, there is little information in the public domain besides the common prior. As a result, higher-order expectations significantly depart from first-order expectations. On the other end of the spectrum, with perfect observation of past prices, the expectations about the fundamental are much better coordinated, resulting in a high level of synchronization among first-order and higher-order expectations. With moderate private noise on the price observation, the patterns of higher-order expectations again strike a balance between the two benchmarks.

The comparison above implies that the way in which the strategic complementarity shapes the outcomes hinges on whether information is endogenous or not. In particular, the effect of shifting the

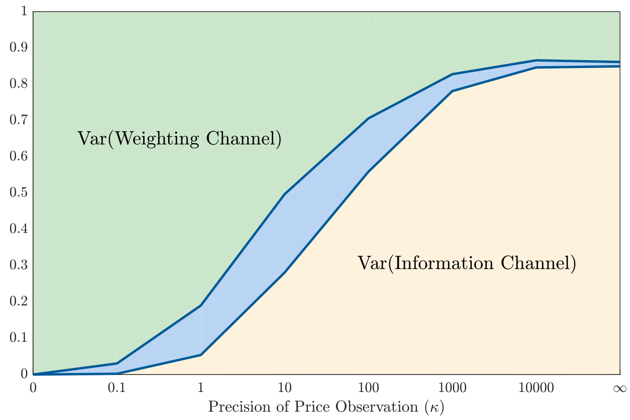


Figure 5: Impact of Changing α from 0 to 0.5.

degree of complementarity from α to $\widehat{\alpha}$ can be split into two parts,

$$p_t^{\widehat{\alpha}} - p_t^{\alpha} = (1-\widehat{\alpha}) \sum_{k=0}^{\infty} \widehat{\alpha}^k \left\{ \mathbb{E}_t^{k+1}[q_t \,;\, \widehat{\alpha}] - \mathbb{E}_t^{k+1}[q_t \,;\, \alpha] \right\} + \sum_{k=0}^{\infty} \left\{ (1-\widehat{\alpha})\widehat{\alpha}^k - (1-\alpha)\alpha^k \right\} \mathbb{E}_t^{k+1}[q_t \,;\, \alpha].$$

The first channel is associated with changes in higher-order expectations, we call this the information channel. A different α may modify the informativeness of prices which, in turn, affects higher-order expectations. The second channel is the weighting channel. It has to do only with the fact that, for higher levels of α, the equilibrium actions put relatively more weight on expectations of higher order. With exogenous information, only the second channel is at work. With perfect price observation, though qualitatively both channel are present, the impact of the second channel is quantitatively limited as higher-order expectations are close to first-order expectations. Figure 5 displays the variance decomposition of prices, p_t, into the information and weighting channels. With noisy price observation, both channels can have similar quantitative contributions.

4.2 Oscillatory Dynamics

The pattern of oscillation with endogenous information is not only theoretically interesting but can also be empirically relevant. In Figure 6, we reproduce the identified impulse responses of inflation and output to a monetary policy shock from Altig, Christiano, Eichenbaum, and Linde (2011). The literature has emphasized that the delayed peak of inflation is difficult to rationalize in standard DSGE models but is consistent with models with informational frictions. However, another feature of the identified impulse responses is that aggregate variables do not return to their steady-state values monotonically but display waves of over- and undershooting, which few models are able to generate.

To show that this pattern can be the result of endogenous information aggregation, we extend the baseline model in Section 2 in two ways: First, we allow the observation of past prices with noise as

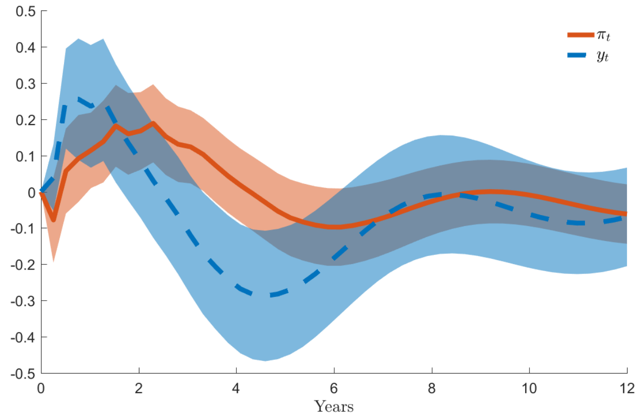


Figure 6: Estimated impulse responses of real GDP and inflation to an unexpected interest rate reduction. Source: Altig, Christiano, Eichenbaum, and Linde (2011).

in equation (4.1). Second, we allow the growth rate of the nominal aggregate demand to follow an AR(1) process as in Woodford (2003),

$$q_t - q_{t-1} = \rho(q_{t-1} - q_{t-2}) + \eta_t. \tag{4.2}$$

The left panel of Figure 7 shows what happens with dispersed but exogenous information. It reproduces the last panel in Figure 5 of Woodford (2003), which he uses to argue that the model can generate a hump-shaped response of inflation and output with a delayed peak for inflation. In this case, however, the responses of aggregate variables return to their steady-state values monotonically after their initial peak. In contrast, the right panel shows the impulse responses with noisy endogenous learning, introduced via an increase in the precision of the price observation, κ, from 0 to 1. As in the identified impulse responses, besides inflation peaking after output, both also display oscillation over time.

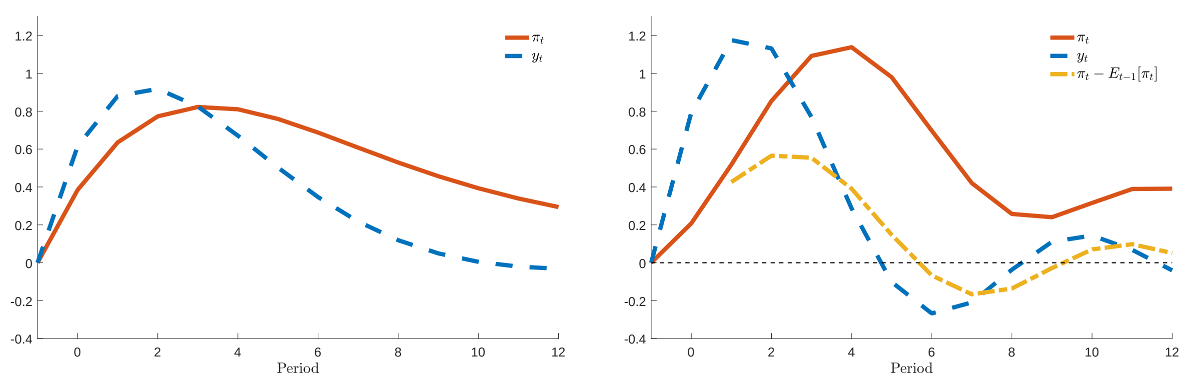


Figure 7: Comparison of impulse response functions predicted by the two models, for $\kappa = 0$ (Exogenous Information) and $\kappa = 1$ (Endogenous Information)

Other parameters, as in Woodford (2003): $\rho = 0.9$, $\alpha = 0.85$, and $\tau = 1/6$.

Moreover, the model yields interesting patterns for the forecasts of aggregate outcomes. The yellow dashed line in the right panel plots the one-period ahead inflation forecast errors. It is positive initially and turns to negative later on, implying that agents' average forecasts underreact initially and overreact later on. This pattern is an empirical regularity which is emphasized in Angeletos, Huo, and Sastry (2020) and also identified by Kucinskas and Peters (2018).[16] Angeletos, Huo, and Sastry (2020) rationalize this pattern by combining dispersed information with over-extrapolation, and here we find that it can be the result of imperfect information aggregation under rational expectations. A more detailed analysis of how well the model is able to match empirical evidence of this kind is left for future work.[17] But we think the model offers some promise of providing a satisfactory explanation to this type of dynamics.

Different from the concesus forecasts, Bordalo, Gennaioli, Ma, and Shleifer (2020) examine the correlation between forecast errors and forecast revisions for various variables at the *individual* level, and show that agents tend to overreact to news on average. Our benchmark model could not replicate this fact, since some deviation from rational expectations is required. Nevertheless, when augmented with either overconfidence as in Broer and Kohlhas (2019) or the diagnostic expectations as in Bordalo, Gennaioli, Ma, and Shleifer (2020), our main results on aggregate dynamics remain valid, while it would help to match the empirical evidence at the individual level at the same time.[18]

5. Extensions and Other Applications

In this section we show that the main insights developed in this paper extend to significantly more involved information structures, allowing for fundamentals that can follow any stochastic process, and multiple public and private signals with noise that also follows arbitrary processes. We also extend the results to environments with more sophisticated linkages among agents featuring forward and backward complementarities, and multivariate systems that can be viewed as a network game with incomplete information.

5.1 General Information Structure

To facilitate the analysis, we switch to a more general notation, since the applications of the following results encompass a wide range of settings that may differ from the one considered in the monetary model analyzed above.

[16]This oscillatory pattern of forecast errors could help reconcile the seemingly conflicting evidence documented by Coibion and Gorodnichenko (2015) and Kohlhas and Walther (2019)—the former finds underreaction of average forecasts while the latter finds overreaction, as explained in Angeletos, Huo, and Sastry (2020).

[17]Here, the law of motion (4.2) is too simple, and we have assumed that nominal aggregate demand is affected only by monetary shocks which is not consistent with the identified VAR.

[18]If, for instance, agents' perceived precision, $\hat{\tau}$, was higher than the actual precision, τ, our results on aggregate dynamics remain true, with the only exception of replacing τ by $\hat{\tau}$. However, at the individual level, econometricians would find agents overreact to their signals.

Best Response. Denote agent i's action in period t by a_{it}. Their best response function is given by

$$a_{it} = \varphi \mathbb{E}_{it} [\xi_{it}] + \alpha \mathbb{E}_{it} [a_t] , \tag{5.1}$$

where ξ_{it} denotes the individual fundamental, which can depend in a flexible way on the aggregate and idiosyncratic shocks, $\boldsymbol{\eta}_t$ and $\boldsymbol{u}_{it}$,

$$\xi_{it} = \boldsymbol{d}(L)\boldsymbol{\eta}_t + \boldsymbol{e}(L)\,\boldsymbol{u}_{it}, \quad \text{with} \quad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\eta^2), \quad \text{and} \quad \boldsymbol{u}_{it} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_u^2).$$

The lag-operator polynomial vectors, $\boldsymbol{d}(L)$ and $\boldsymbol{e}(L)$, are assumed to have square-summable coefficients. We also assume that $\boldsymbol{d}(L)$ is not a constant vector to rule out fundamentals that are i.i.d. on common shocks; in which case the equilibrium is always invertible, and we rule out redundant shocks by assuming that $\boldsymbol{\Sigma}_\eta$ and $\boldsymbol{\Sigma}_u$ have full rank.

Information Structure. Agents have perfect recall and, in each period t, observe three sets of signals: (1) the previous period aggregate action, a_{t-1}, (2) a vector of private signals, $\boldsymbol{x}_{it}$, (3) and a vector of public signals, $\boldsymbol{z}_t$, where

$$\boldsymbol{x}_{it} = \mathbf{A}(L)\boldsymbol{\eta}_t + \mathbf{B}(L)\boldsymbol{u}_{it}, \quad \text{and} \quad \boldsymbol{z}_t = \mathbf{C}(L)\boldsymbol{\eta}_t.$$

We make two assumptions: first, that there are as many common shocks as there are public signals (including a_{t-1}), that is $\dim(\boldsymbol{\eta}_t) = \dim(\boldsymbol{z}_t) + 1$; second, that the matrix $\mathbf{B}(L)$ is invertible. To see why we impose the first assumption, note that, when $\dim(\boldsymbol{\eta}_t) < \dim(\boldsymbol{z}_t) + 1$, there are more public signals than common shocks and past aggregate fundamentals are always perfectly revealed independently of the degree of strategic complementarity or of informational friction; when $\dim(\boldsymbol{\eta}_t) > \dim(\boldsymbol{z}_t) + 1$, there are more common shocks than public signals and the system is non-invertible by construction, as in Section 4. The interesting case for a discussion about invertibility is the one in which $\dim(\boldsymbol{\eta}_t) = \dim(\boldsymbol{z}_t) + 1$. The second assumption is more standard, it essentially excludes the cases in which non-invertibility is due to exogenously imposed shock processes. If $\mathbf{B}(L) = \mathbf{B}_0$, which is usually the case in most information structures considered in the literature, then the invertibility of $\mathbf{B}(L)$ is only violated if $\mathbf{B}_0$ is singular or not square in which case one of the private signals is redundant. Moreover, the polynomial matrices, $\mathbf{A}(L)$, $\mathbf{B}(L)$ and $\mathbf{C}(L)$, must have square-summable coefficients.

Invertibility. In the economy studied in Section 3 there is only one aggregate shock, the monetary shock, and invertibility is obtained if and only if the equilibrium process for the price index does not contain an inside root. When there are multiple aggregate shocks, the equilibrium process can be expressed as

$$a_t = \boldsymbol{g}(L)\boldsymbol{\eta}_t.$$

Driven by more than one aggregate shock, the aggregate outcome by itself can no longer reveal all underlying states. The relevant question becomes whether the history of signals, taken altogether, contains sufficient information.

Formally, we define an equilibrium process to be invertible if the history of the public signals, $\{a^t, \boldsymbol{z}^t\}$, contains the same information as the common shocks. The following lemma provides the corresponding criterion for invertibility.

Lemma 3. *If* $\det \begin{bmatrix} \boldsymbol{g}(L) & \mathbf{C}(L) \end{bmatrix}'$ *does not contain any inside root, the equilibrium is invertible. Then, the public signals and the aggregate outcomes perfectly aggregate information,*

$$\mathbb{E}\left[\boldsymbol{\eta}_t | a^t, \boldsymbol{z}^t\right] = \boldsymbol{\eta}_t, \quad \text{and} \quad \mathbb{E}\left[\xi_{it} | a^t, \boldsymbol{z}^t, \boldsymbol{x}_i^t\right] = \xi_{it}. \tag{5.2}$$

This result generalizes Lemma 2 to a multivariate system. The exogenously specified signal structure and the endogenously determined equilibrium process jointly determine whether agents can perfectly infer past shocks in the economy. Note that, once aggregate shocks are known, idiosyncratic shocks are also known since $\mathbf{B}(L)$ has been assumed to be invertible.

With the general information structure, we can no longer provide a simple partition of the parameter space into invertible and non-invertible regions as in Proposition 3.1. However, the basic insight derived in Section 3 remains true.

Theorem 1. *There exists* $\overline{\alpha} \in (-1, 1)$ *such that, if* $\alpha > \overline{\alpha}$*, the equilibrium is not invertible.*

Proof. See Appendix D. □

The exact threshold $\overline{\alpha}$, above which the equilibrium is not invertible, depends on the details of the information structure. Independent of these details, however, such a threshold always exists. As in Section 3, if the degree of strategic complementarity is high enough, the equilibrium is not invertible. Under this more general information structure, however, the aggregate action may contain information on the aggregate fundamental as well as the common noise. The degree of strategic complementarity affects the invertibility of the joint dynamics of aggregate actions and public signals.

There is a sense in which the information structure we set up in this section is too general. In the following example, we consider a simplified structure that encompasses many used in the literature. Directly, it allows for the introduction of public signals to the monetary model from Section 3. We also allow agents to have idiosyncratic fundamentals which they observe with arbitrary precision. Hence, it also encompasses the information structure in the business-cycle model from Angeletos and La'O (2010) (with the addition of the observation of past aggregate outcomes). With this simpler structure, we can characterize in more detail how public signals affect information aggregation with takeaways that are still broadly applicable.

Example: how public signals affect invertibility. Suppose the best response function is given by equation (5.1) with the individual fundamental, ξ_{it}, and the aggregate fundamental, ξ_t, satisfying

$$\xi_{it} = \xi_t + \omega_{it}, \quad \omega_{it} \sim \mathcal{N}(0, \tau_\omega^{-1}), \quad \text{and} \quad \xi_t = d(L)\eta_t, \quad \eta_t \sim \mathcal{N}(0, 1),$$

for some arbitrary $d(L)$ with the normalization $d_0 = 1$. For simplicity we set $\varphi = 1 - \alpha$. Every period, agent i observes last period's aggregate action, a_{t-1}, a private, and a public signal,

$$x_{it} = \xi_{it} + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \tau_u^{-1}), \quad \text{and} \quad z_t = \xi_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \tau_\varepsilon^{-1}).$$

In this setup, we can establish the following proposition which provides an explicit condition for invertibility. The proof of it also serves as a sketch of the proof of Theorem 1.

Proposition 5.1. *The equilibrium is invertible if and only if every root of the function* $\Gamma(z)$

$$\Gamma(z) \equiv d(z) - \frac{\alpha\tau_u + \tau_\omega}{\tau_u + \tau_\omega + (1-\alpha)\,\tau_u\,(\tau_\varepsilon + \tau_\omega)} \tag{5.3}$$

lies outside the unit circle, which is not the case for α *high enough.*

In particular, if the fundamental follows an AR(1) process, i.e. $d(L) = 1/(1-\rho L)$*, it is necessary and sufficient for invertibility that*

$$\alpha < 1 - \frac{\rho(\tau_u + \tau_\omega)}{\tau_u(1 + \rho + \tau_\varepsilon + \tau_\omega)}. \tag{5.4}$$

Proof. Suppose the equilibrium is invertible. Then, agents can infer past aggregate shocks perfectly and the effect of common shocks on the aggregate outcome can only be transitory. It follows that, the impulse response of the aggregate outcome to any common shock only differs from that of the aggregate fundamental, ξ_t, on impact. Accordingly, the law of motion of the aggregate outcome can be expressed as

$$a_t = \underbrace{g_\eta \eta_t + g_\varepsilon \varepsilon_t}_{\text{impact effect}} + (d(L) - 1)\eta_t.$$

To verify if the equilibrium is indeed invertible, we need to check the condition in Lemma 3, which, in this case, reduces to checking if any of the roots of

$$\Gamma(z) = d(z) - \frac{1 - g_\eta}{1 - g_\varepsilon}, \tag{5.5}$$

are inside the unit circle. The impact effects

$$g_\eta = \frac{(1-\alpha)\tau_u(1+\tau_\omega) + \tau_\varepsilon(\tau_\omega + \tau_u)}{(1-\alpha)\tau_u\tau_\omega + (1+\tau_\varepsilon)(\tau_u + \tau_\omega)}, \quad \text{and} \quad g_\varepsilon = \frac{\tau_\varepsilon(\tau_\omega + \alpha\tau_u)}{(1-\alpha)\tau_u\tau_\omega + (1+\tau_\varepsilon)(\tau_u + \tau_\omega)}, \tag{5.6}$$

can be obtained by solving a simple static forecasting problem. Then, equation (5.3) follows from substituting equation (5.6) into (5.5). Next notice that, as α increases towards 1, equation (5.3) converges to $d(z) - 1$ which has root zero, since $d_0 = 1$. Hence, since the roots of a polynomial are a continuous function of its coefficients, there exists some $\overline{\alpha} \in (-1, 1)$ such that, for all $\alpha > \overline{\alpha}$, this equation has an inside root and the equilibrium cannot be invertible. Moreover, condition (5.4) follows immediately from solving for the root of equation (5.3) and requiring it to be outside the unit circle. □

Two opposing forces. How does the precision of the public signal, τ_ε, affect invertibility? Note that in the proof of Proposition 5.1, the root of evil that causes non-invertibility is that $\frac{1-g_\eta}{1-g_\varepsilon} \to 0$ when the magnitude of α approaches to zero. In general, the equilibrium is more likely to be invertible, or $\Gamma(z)$ less likely to have an inside root when $\frac{1-g_\eta}{1-g_\varepsilon}$ is relatively large. This observation suggests that intuitively, a higher τ_ε has the following two opposing effects:

1. The extra precision leads to a better estimate of the fundamental and a stronger response to the fundamental, a higher g_η. This tends to make the equilibrium invertible.

2. The response to the common noise, g_ε, also increase since agents rely relatively more on the public signal z_t. This makes the information content of the aggregate action a_{t-1} closer to that of the public signal z_t. Put differently, there is less differential information contained in the aggregate outcome in comparison with the public signal, and this tends to make the equilibrium non-invertible.

To further appreciate this last point, consider the special case in which $d(L) = 1/(1 - \rho L)$ and $\tau_\omega \to \infty$. The information structure, then, reduces to the one from Section 3 with the addition of a public signal. In this case, the invertibility condition (5.4) becomes

$$\alpha < 1 - \frac{\rho}{\tau_u},$$

which is identical to condition (3.4). This may seem puzzling at first, as the precision of the public signal plays no role in determining invertibility. However, this is because the two forces discussed above exactly cancel each other. In the extreme, when τ_ε goes to infinity, agents can infer the fundamental almost perfectly using the public signal. On the other hand, they correspondingly discard their private signals, and the aggregate outcome contains no more information than the one already obtained with the public signal. These effects cancel, leaving open the possibility that the equilibrium is not invertible.

In contrast, when τ_ω is finite, agents always use their private signals to learn about their idiosyncratic fundamental. It follows that the aggregate outcome necessarily aggregates the information contained in private signals, which differentiates itself from the public signal z_t. The two forces do not cancel each other, and the precision of the public signal, τ_ε, does matter for the determination of invertibility.

5.2 Forward Complementarities

The best-response function in equation (5.1) only allows for static strategic complementarities, that is, agent i's action depends on the current aggregate action. Here, we extend the analysis to allow for arbitrary forward-looking complementarities, that is, agent i's action can depend on future aggregate actions or on their own future actions in a flexible way. We consider the following best-response function,

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_{it}] + \alpha \mathbb{E}_{it}[a_t] + \mathbb{E}_{it}[\gamma(L) a_t] + \mathbb{E}_{it}[\beta(L) a_{it}], \tag{5.7}$$

where

$$\gamma(L) \equiv \sum_{k=1}^{\infty} \gamma_k L^{-k}, \quad \beta(L) \equiv \sum_{k=1}^{\infty} \beta_k L^{-k}, \quad \text{and} \quad |\alpha| + \|\gamma(L)\| + \left\|\beta(L)\right\| < 1.$$

and we impose a relatively weak condition on the parameters that guarantees existence of the equilibrium.[19]

Even though the model structure is more sophisticated, it turns out that when the equilibrium is invertible, the general best-response function in condition (5.7) collapses to the static best response in condition (5.1) with a modified fundamental, as forward-looking higher-order expectations collapse to first-order expectations in this scenario. The following proposition formalizes the required transformation.

Proposition 5.2. *If the equilibrium is invertible, then the actions under best response (5.7) are observationally equivalent to those under the following transformed best response*

$$a_{it} = \varphi \mathbb{E}_{it}[\widetilde{\xi}_{it}] + \alpha \mathbb{E}_{it}[a_t], \quad \textit{with} \quad \widetilde{\xi}_{it} \equiv \frac{1-\alpha}{1-\alpha-\gamma(L)-\beta(L)}\, \xi_{it}.$$

Proof. See Appendix E. □

With this transformation, Theorem 1 can be applied to the general class of best-response functions described by condition (5.7). The fact that there are forward complementaries does not change the fact that there always exists a threshold level for the degree of *static* strategic complementarity, $\overline{\alpha}$, such that, if $\alpha \geq \overline{\alpha}$, the equilibrium is not invertible.

The effects of forward complementarities on the invertibility of the equilibrium, however, are not as simple as in the static case. In Appendix H we provide an analysis of the invertiblility condition with an ARMA (1,1) fundamental process, which is fairly complicated as it hinges on the interaction between the fundamental and the coordination structure. In Appendix I we show how Theorem 1 can be further extended to economies with both forward and backward complementarities which encompasses many environments considered in the DSGE literature.

In what follows, we focus on a stylized forward-looking game, which allows us to explore how price stickiness and MPC affect the invertibility of inflation and output, respectively.

[19] In these conditions, $\|\cdot\|$ denotes the operator norm.

Example: effects of price stickiness and MPC on invertibility. Consider the following best response

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_t] + \alpha \mathbb{E}_{it}[a_t] + \beta \mathbb{E}_{it}[a_{it}].$$

This setting can nest the two building blocks of the New Keynesian model, the Dynamic IS curve, and the New Keynesian Philips curve (NKPC). Relaxing the common-knowledge foundations of the New Keynesian model along the lines of Angeletos and Lian (2018) yields[20]

$$\pi_{it} = \kappa \mathbb{E}_{it}[mc_t] + (1-\theta)\mathbb{E}_{it}[\pi_t] + \delta\theta \mathbb{E}_{it}[\pi_{it+1}], \tag{5.8}$$

$$c_{it} = -\varsigma(1-\texttt{mpc})\mathbb{E}_{it}[r_t] + \texttt{mpc}\ \mathbb{E}_{it}[c_t] + (1-\texttt{mpc})\ \mathbb{E}_{it}[c_{i,t+1}]. \tag{5.9}$$

In the NKPC, mc_t is the real marginal cost, θ is the Calvo parameter, and δ is the discount rate. In the dynamic IS curve, r_t is the real rate, ς is the intertemporal elasticity of substitution, $\texttt{mpc}$ is the marginal propensity to consume. With complete information, these two conditions reduce to their familiar textbook versions. Crucially, the price stickiness θ and the MPC control the degree of static strategic complementarity in firms' pricing decision and consumers' saving-consumption decision.

Here, we treat the marginal cost, mc_t, and the real rate, r_t, as exogenous fundamentals, and therefore the supply block (5.8) and the demand block (5.9) do not interact with each other and can be treated separately, as in Nimark (2008) and Angeletos and Huo (2018). As in our baseline specification, we assume that the fundamentals mc_t and r_t follow AR(1) processes. Every period, firms and consumers observe an exogenous signal about the fundamental with private noises, and the past realization of π_t and c_t, respectively.

Applying Proposition 5.2 and Theorem 1 leads to the following observation.

Corollary 2. *For a fixed level of private noise, inflation is noninvertible when prices are sufficiently flexible (lower θ), and aggregate consumption is noninvertible when MPC is sufficiently high (higher $\texttt{mpc}$).*

Proof. See Appendix F. □

The underlying logic for this result can be understood as follows: a lower θ implies more frequent price adjustments. As a result, for the firms that can adjust their prices, there is a greater need to worry about other firms' price-setting decisions. A higher MPC makes consumption more sensitive to income changes, which leads to a stronger dependence on aggregate demand. Both of these effects map to a higher degree of static strategic complementarity when the best responses are transformed according to Proposition 5.2.

5.3 Multivariate System: effect of aggressiveness of nominal interest rate response

In the previous example, we treat inflation and output separately. In this section, we show that the main insights extend to the case where the supply and demand block of the economy interact with

[20] Appendix F presents a derivation of these equations together with the proof of Corollary 2.

each other. Particularly, we show that when the central bank follows a standard Taylor rule, a less aggressive response of the nominal interest rate to inflation tends to reduce the amount of information aggregated by output and inflation, yielding a non-invertible equilibrium.

We extend the model in the following way,

$$\pi_{it} = \kappa \mathbb{E}_{it}[c_t + \xi_t^s] + (1-\theta)\mathbb{E}_{it}[\pi_t] + \delta\theta\mathbb{E}_{it}[\pi_{it+1}], \tag{5.10}$$

$$c_{it} = -\varsigma(1-\texttt{mpc})\mathbb{E}_{it}[i_t - \pi_{t+1} + \xi_t^d] + \texttt{mpc}\ \mathbb{E}_{it}[c_t] + (1-\texttt{mpc})\ \mathbb{E}_{it}[c_{i,t+1}], \tag{5.11}$$

$$i_t = \phi_\pi \pi_t. \tag{5.12}$$

The first equation is the NKPC, modified to to allow a real marginal cost proportional to aggregate output, and subject to a cost-push shock ξ^s. The second equation is the dynamic IS curve replacing real interest rate by the difference of nominal interest rates and expected future inflation, and augmented with a preference shock ξ_t^d. The third equation is the Taylor rule for the nominal interest rate, where ϕ_π controls how aggressive monetary policy is against inflation.

In terms of information, we allow individual firms and consumers to observe private signals about the demand and supply shocks, $x_{i,t}^s = \xi_t^s + u_{i,t}^s$, $x_{i,t}^d = \xi_t^d + u_{i,t}^d$, as well as past inflation and output realizations. To simplify the analysis, we assume that both ξ_t^s and ξ_t^d follows an AR(1) process with persistence ρ, and that $u_{i,t}^s$ and $u_{i,t}^d$ are both i.i.d shocks with variance τ^{-1}.[21]

System (5.10)-(5.12) effectively consists of a forward-looking network game. As a result, the strength of GE effects, or the degree strategic complementarity, no longer depends on a single scalar but on all the relevant structural parameters that control within-group and cross-group dependences. Here, we highlight the role of ϕ_π. When the nominal interest rate is passive (such as in the zero-lower bound), higher inflation induces higher demand, which further pushes up marginal costs and inflation. A higher ϕ_π suppresses aggregate demand's response and mutes this reinforcing mechanism, which implies a higher degree of strategic *substitutability*.

With perfect information, varying ϕ_π only matters for the magnitude of the responses of output and inflation to shocks. With incomplete information and observation about endogenous outcomes, varying ϕ_π also changes the information content of c_t and π_t. As a higher ϕ_π implies higher strategic substitutability, it makes agents rely more on their private signals. Thus, the economy is more likely to be invertible in the sense that, by observing $\{c^t, \pi^t\}$, one can perfectly infer the demand and supply shocks, i.e.,[22]

$$\mathbb{E}[\xi_t^s | c^t, \pi^t] = \xi_t^s, \quad \text{and} \quad \mathbb{E}[\xi_t^d | c^t, \pi^t] = \xi_t^d.$$

This argument is formalized in the following proposition.

[21] The main conclusion of our results on the comparative statics on ϕ_π does not depend on the symmetry assumption.

[22] Technically, invertibility requires the equilibrium process $\begin{bmatrix} c_t \\ \pi_t \end{bmatrix} = \mathbf{g}(L) \begin{bmatrix} \eta_t^s \\ \eta_t^d \end{bmatrix}$ to satisfy the condition that $\det[\mathbf{g}(L)]$ does not contain any inside root.

Proposition 5.3. *The outcomes* $\{\pi^t, c^t\}$ *perfectly aggregate information only if* ϕ_π *exceeds the threshold* Φ *given by*

$$\Phi = \begin{cases} \frac{\rho^2-\theta(1-mpc)\tau^2}{\kappa\varsigma(1-mpc)\tau^2}, & \text{when } (1-\theta-mpc)^2 < 4\kappa\varsigma(1-mpc)\phi_\pi, \\ \frac{\tau\rho(1+\theta-mpc)-\rho^2-\theta(1-mpc)\tau^2}{\kappa\varsigma(1-mpc)\tau^2}, & \text{otherwise.} \end{cases} \tag{5.13}$$

Proof. See Appendix G. □

The left panel of Figure 8 shows how equilibrium invertibility depends on the aggressiveness of the nominal interest rate response, ϕ_π, and the precision of the exogenous signals received by the agents, τ. Following Proposition 5.3, a higher ϕ_π makes endogenous outcomes more likely to be invertible. The right panel presents an example of an impulse response of inflation to the supply shock ξ_t^s in the non-invertible region. In the invertible region, the IRF of inflation differs from its perfect information counterpart only in the initial period, that is,

$$\pi_t = \pi_t^* - \varphi\eta_t^s,$$

where π_t^* is the perfect information outcome, and φ is a constant modifying the response on impact.[23]

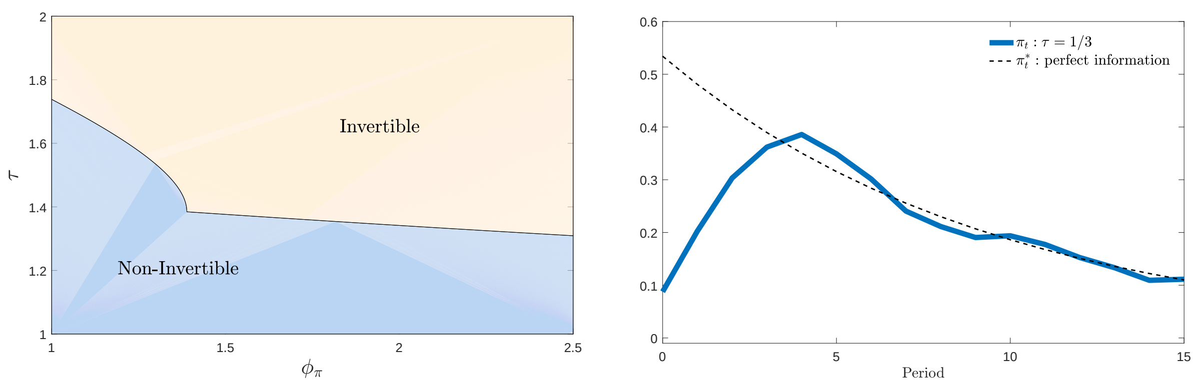


Figure 8: Invertibility region and Example of a non-invertible impulse response function

Parameters: $\kappa = 0.05$, $\theta = 0.4$, $\delta = 0.99$, $\varsigma = 1$, mpc $= 0.1$, and $\rho = 0.9$. For the right panel, $\phi_\pi = 1$ is used to generate the IRF.

In contrast, in the non-invertible region, inflation displays the now familiar early underreaction and the following oscillatory pattern. This suggests that the central bank's policy rule matters not only for the magnitude of aggregate outcome's response to shocks but also disciplines their entire dynamic patterns. Interestingly, according to our theory, a more hawkish monetary policy, such as that followed in the post-Volker era, is predicted to eliminate the additional persistence of inflation due to imperfect information aggregation, which is consistent with the empirical findings in Fuhrer (2010) and Cogley, Primiceri, and Sargent (2010).

[23] The formulas for π_t^* and φ can be found in Appendix G.

6. Conclusion

We show that, with dispersed information, even when past aggregate outcomes are perfectly observed, the underlying states may not be fully revealed. The extent to which endogenous outcomes help aggregate information is increasing in the precision of private signals about economic fundamentals and is decreasing in the strength of general equilibrium feedback effects. With imperfect information aggregation, the equilibrium dynamics oscillate around the economic fundamental, and the forecast errors display an initial underreaction and a delayed overreaction which is consistent with the empirical evidence in Angeletos, Huo, and Sastry (2020). Together with the solution to the exogenous information environment, our results provide a second benchmark to characterize properties of equilibrium under incomplete information.

When the theory is applied to familiar macro models, it leads to several applied implications. In monetary models à la Woodford (2003), the GE feedback effect comes from firms' pricing complementarity. When it is strong enough, monetary shocks can have persistent effects even when firms observe past prices. In standard NKPC models, a less aggressive monetary policy induces a stronger GE feedback effect between inflation and output, and can result in hump-shaped and oscillatory responses of aggregate variables.

Left outside this paper is the extension beyond the standard linear-Gaussian framework. In an asset pricing setting, Albagli, Hellwig, and Tsyvinski (2015) allow a much more flexible payoff structure and shock distribution. Hassan and Mertens (2015) also incorporate dispersed information into a non-linear DSGE model, and quantify the informational role of stock prices. These works allow agents to observe past shocks, and therefore the learning is essentially static. It would be interesting to explore the extent to which agents can learn from the past in a non-linear environment.

References

Acharya, S., J. Benhabib, and Z. Huo (2017): "The anatomy of sentiment-driven fluctuations," Discussion paper, National Bureau of Economic Research.

Adam, K. (2007): "Optimal monetary policy with imperfect common knowledge," *Journal of Monetary Economics*, 54(2), 267–301.

Albagli, E., C. Hellwig, and A. Tsyvinski (2015): "A Theory of Asset Prices based on Heterogeneous Information," Discussion paper, Yale Mimeo.

Altig, D., L. Christiano, M. Eichenbaum, and J. Linde (2011): "Firm-Specific Capital, Nominal Rigidities and the Business Cycle," *Review of Economic Dynamics*, 14(2), 225–247.

Amador, M., and P.-O. Weill (2010): "Learning from Prices: Public Communication and Welfare," *Journal of Political Economy*, 118(5), 866 – 907.

——— (2012): "Learning from Private and Public Observations of Others Actions," *Journal of Economic Theory*, 147(3), 910–940.

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

Angeletos, G.-M., Z. Huo, and K. Sastry (2020): "Imperfect Macroeconomic Expectations: Evidence and Theory," in *NBER Macroeconomics Annual 2020, volume 35*, NBER Chapters. National Bureau of Economic Research, Inc.

Angeletos, G.-M., L. Iovino, and J. La'O (2020): "Learning over the Business Cycle: Policy Implications," .

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.

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.

——— (2009): "Policy with Dispersed Information," *Journal of the European Economic Association*, 7(1), 11–60.

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

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

Broer, T., and A. Kohlhas (2019): "Forecaster (Mis-)Behavior," *IIES miméo*.

Chahrour, R., and G. Gaballo (2019): "Learning from House Prices: Amplification and Business Fluctuations," CEPR Discussion Papers 14120, C.E.P.R. Discussion Papers.

Christiano, L., M. Eichenbaum, and C. Evans (2005): "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*.

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010): "Inflation-gap persistence in the US," *American Economic Journal: Macroeconomics*, 2(1), 43–69.

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

Fuhrer, J. C. (2010): "Inflation persistence," in *Handbook of monetary economics*, vol. 3, pp. 423–486. Elsevier.

Gaballo, G. (2018): "Price Dispersion, Private Uncertainty, and Endogenous Nominal Rigidities," *Review of Economic Studies*, 85(2), 1070–1110.

Grossman, S. J. (1976): "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information," *Journal of Finance*, 31(2), 573–585.

Hassan, T. A., and T. M. Mertens (2015): "Information Aggregation in a Dynamic Stochastic General Equilibrium Model," *NBER Macroeconomics Annual*, 29(1), 159–207.

Hayek, F. A. (1945): "The Use of Knowledge in Society," *The American Economic Review*, 35(4), 519–530.

Hellwig, C., and V. Venkateswaran (2015): "Dispersed Information, Sticky Prices and Monetary Business Cycles: A Hayekian Perspective," Discussion paper.

Hellwig, M. F. (1980): "On the aggregation of information in competitive markets," *Journal of Economic Theory*, 22(3), 477–498.

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

Huo, Z., and N. Takayama (2017): "Rational Expectations Models with Higher Order Beliefs," Manuscript, Yale University.

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

Kasa, K., T. B. Walker, and C. H. Whiteman (2014): "Heterogeneous Beliefs and Tests of Present Value Models," *Review of Economic Studies*, 81(3), 1137–1163.

Kohlhas, A., and A. Walther (2019): "Asymmetric Attention," *IIES miméo*.

Kucinskas, S., and F. Peters (2018): "Measuring Biases in Expectation Formation," Tinbergen Institute Discussion Papers 18-058/IV, Tinbergen Institute.

Lucas, Robert E, J. (1973): "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63(3), 326–334.

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

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

Melosi, L. (2017): "Signalling effects of monetary policy," *The Review of Economic Studies*, 84(2), 853–884.

Messner, S., and X. Vives (2005): "Allocative and Productive Efficiency in REE with Asymmetric Information," Discussion paper, IESE mimeo.

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

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

Nimark, K. P. (2014): "Man-Bites-Dog Business Cycles," *American Economic Review*, 104(8), 2320–67.

Rondina, G., and T. B. Walker (2020): "Confounding Dynamics," Working paper.

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

Sims, C. A. (2001): "Solving Linear Rational Expectations Models," *Journal of Computational Economics*, 20, 1–20.

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

Vives, X. (2014): "On The Possibility Of Informationally Efficient Markets," *Journal of the European Economic Association*, 12(5), 1200–1239.

——— (2017): "Endogenous Public Information and Welfare in Market Games," *Review of Economic Studies*, 84(2), 935–963.

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

Appendix

A. Proof of Existence

This proof of existence is based on the one contained in Huo and Takayama (2017).

Proposition A.1. *If the best-response function can be written as*

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_t] + \alpha \mathbb{E}_{it}[a_t],$$

with $|\alpha| < 1$*, then an equilibrium exists.*

Proof. Let $\mathcal{D} \subset \ell^2$ be a bounded subset such that

$$\lim_{N\to\infty} \sup\{||0, 0, \ldots, x_N, x_{N+1}, \ldots|| : x \in \mathcal{D}\} = 0.$$

Let $g \in \mathcal{D}$ and $h \in \mathcal{D}$ be an arbitrary policy rule such that

$$a_{it} = g(L)\eta_t + h(L)u_{it}, \quad \text{and} \quad a_t = g(L)\eta_t.$$

With the norm

$$||\{g, h\}|| = \sqrt{\sum_{k=0}^{\infty} \sigma_\eta^2 g_k^2 + \sigma_u^2 h_k^2},$$

the set of such policy rules is a Banach space. Define $\{\widehat{g}_\xi, \widehat{h}_\xi\}$, and $\{\widehat{g}_a, \widehat{h}_a\}$ such that

$$\mathbb{E}_{it}[\xi_{it}] = \widehat{g}_\xi(L)\eta_t + \widehat{h}_\xi(L)u_{it}, \quad \text{and} \quad \mathbb{E}_{it}[a_t] = \widehat{g}_a(L)\eta_t + \widehat{h}_a(L)u_{it}.$$

Note that in the inference problem, the loading on past signals must converge to zero as signals realized in the infinite past contain no information about current shocks. This implies that $\{\widehat{g}_\xi, \widehat{h}_\xi\}$ and $\{\widehat{g}_a, \widehat{h}_a\}$ are also in set $\mathcal{D}$.

Hence, for $\{g, h\}$ to be an equilibrium it must be that

$$g(L)\eta_t + h(L)u_{it} = \varphi\left(\widehat{g}_\xi(L)\eta_t + \widehat{h}_\xi(L)u_{it}\right) + \alpha\left(\widehat{g}_a(L)\eta_t + \widehat{h}_a(L)u_{it}\right).$$

Define the operator $\mathcal{T} : \mathcal{D} \times \mathcal{D} \to \mathcal{D} \times \mathcal{D}$ as

$$\mathcal{T}\left(\{g, h\}\right) = \left\{\varphi\widehat{g}_\xi + \alpha\widehat{g}_a\,,\ \varphi\widehat{h}_\xi + \alpha\widehat{h}_a\right\}.$$

The equilibrium is a fixed point of the operator $\mathcal{T}$, and existence follows from Schaefer's fixed point theorem. To apply this theorem, we need to show that $\mathcal{T}$ is a continuous mapping and it maps to compact set. Notice that if $\mathcal{D}$ is bounded, then it is totally bounded. Since the space ℓ^2 is complete, and therefore the set $\mathcal{D}$ is compact. In below we show $\mathcal{D}$ is bounded.

By iteration on the best-response function and the definition of the aggregate action we obtain the higher-order expectation representation of the agents' actions,

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_{it}] + \varphi \sum_{k=1}^{\infty} \alpha^k \mathbb{E}_{it}^k[\xi_{it}].$$

By the law of total variance, the variances of $\mathbb{E}_{it}[\xi_{it}]$ and $\mathbb{E}_{it}^k[\xi_{it}]$ are bounded by the variance of ξ_{it}. Further, we have that

$$\begin{aligned}
\mathrm{Var}\left(\sum_{k=1}^{\infty} \alpha^k \mathbb{E}_{it}^k[\xi_{it}]\right) &= \sum_{k=1}^{\infty} \alpha^{2k}\, \mathrm{Var}\left(\mathbb{E}_{it}^k[\xi_{it}]\right) + 2 \sum_{1 \le k \le q \le \infty} \alpha^{kq}\, \mathrm{Cov}\left(\mathbb{E}_{it}^k[\xi_{it}], \mathbb{E}_{it}^q[\xi_{it}]\right) \\
&\le \sum_{k=1}^{\infty} \alpha^{2k}\, \mathrm{Var}\left(\mathbb{E}_{it}^k[\xi_{it}]\right) + 2 \sum_{1 \le k \le q \le \infty} \alpha^{kq} \sqrt{\mathrm{Var}\left(\mathbb{E}_{it}^k[\xi_{it}]\right) \mathrm{Var}\left(\mathbb{E}_{it}^q[\xi_{it}]\right)} \\
&= \left(\sum_{k=1}^{\infty} \alpha^k \sqrt{\mathrm{Var}\left(\mathbb{E}_{it}^k[\xi_{it}]\right)}\right)^2.
\end{aligned}$$

Therefore, no matter what the signal process is, the variance of a_{it} is bounded. It follows that $\mathcal{T}$ is a bounded mapping. The fact that $\mathcal{T}$ is continuous follows from the continuity of expectation operator. This completes the proof. □

B. Proof of Proposition 3.1

Suppose that the stochastic process for p_t is invertible, then observing $\{p_k\}_{k=-\infty}^{t-1}$ perfectly reveals the underlying aggregate shocks $\{\eta_k\}_{k=-\infty}^{t-1}$ and, therefore, $\{q_k\}_{k=-\infty}^{t-1}$, so that the only shock the firms are uncertain about is the current η_t. Guess that the equilibrium policy function has state variables x_{it} and q_{t-1}, that is guess that firm i's policy function can be written as

$$p_{it} = \phi_x x_{it} + \phi_q q_{t-1},$$

for some scalars ϕ_x and ϕ_q. It follows that, in aggregate terms,

$$p_t = \phi_x q_t + \phi_q q_{t-1}.$$

To verify the guess, notice that, since $x_{it} - \rho q_{t-1}$ is a noisy signal about η_t with precision τ, we have that

$$\mathbb{E}_{it}[q_t] = \frac{\rho}{1+\tau} q_{t-1} + \frac{\tau}{1+\tau} x_{it}.$$

Substituting these results into the best-response function (2.2) we obtain

$$p_{it} = \frac{\tau((1-\alpha) + \alpha\phi_x)}{1+\tau} x_{it} + \left(\frac{\rho((1-\alpha)+\alpha\phi_x)}{1+\tau} + \alpha\phi_q\right) q_{t-1},$$

which implies the following consistency requirement

$$\phi_q = \frac{\rho}{1+(1-\alpha)\tau}, \qquad \text{and} \qquad \phi_x = \frac{(1-\alpha)\tau}{1+(1-\alpha)\tau}.$$

Hence,for p_t to indeed follow an invertible process it is necessary and sufficient that $g(L) = \frac{\phi_x+\phi_q L}{1-\rho L}$ not have an inside root, or that $|\phi_q/\phi_x| < 1$ which implies the result.

C. Proof of Proposition 3.2

The first claim follows directly from the proof of Proposition 3.1. To establish the second claim we follow a significantly more involved argument. To facilitate reading it we include the proof of the necessary lemmas at the end of this section.

For a contradiction, suppose there is a finite-state representation, then the law of motion of the aggregate action can be written as

$$p_t = g(L)\eta_t = C(L)h(L)\eta_t,$$

where $h(L)$ is analytic and does not contain any inside root, and $C(z)$ is given by

$$C(z) = \prod_{i=1}^{n} (z - \lambda_i).$$

where $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ are inside roots of $g(z)$. The signal structure can be written as follows

$$\begin{bmatrix} x_{it} \\ p_{t-1} \end{bmatrix} = \mathbf{\Gamma}(L) \begin{bmatrix} u_{it} \\ \eta_t \end{bmatrix} \equiv \begin{bmatrix} \tau^{-1/2} & \frac{1}{1-\rho L} \\ 0 & Lg(L) \end{bmatrix} \begin{bmatrix} u_{it} \\ \eta_t \end{bmatrix}.$$

The determinant of $\mathbf{\Gamma}(L)$ is

$$\det[\mathbf{\Gamma}(L)] = \tau^{-1/2} L C(L) h(L),$$

and it contains inside roots $\{\lambda_1, \ldots, \lambda_n, \lambda_{n+1}\}$, with $\lambda_{n+1} \equiv 0$. Denote the Blaschke matrix by

$$\mathbf{B}(L; \lambda) = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1-\lambda L}{L-\lambda} \end{bmatrix},$$

and let the fundamental representation of the signal process be given by

$$\mathbf{\Gamma}^*(L)\varepsilon_{it} = \mathbf{\Gamma}(L) \begin{bmatrix} u_{it} \\ \eta_t \end{bmatrix},$$

where

$$\varepsilon_{it} \equiv \mathbf{A}(L) \begin{bmatrix} u_{it} \\ \eta_t \end{bmatrix}, \quad \mathbf{A}(L) \equiv \mathbf{B}'\left(L^{-1}; \lambda_{n+1}\right) \mathbf{W}'_{\lambda_{n+1}} \cdots \mathbf{B}'\left(L^{-1}; \lambda_1\right) \mathbf{W}'_{\lambda_1}, \quad \mathbf{\Gamma}^*(L) \equiv \mathbf{\Gamma}(L) \mathbf{A}'(L^{-1}).$$

and $\{\mathbf{W}_{\lambda_i}\}$ are the rotation matrices that satisfy $\mathbf{W}_{\lambda_i}\mathbf{W}'_{\lambda_i} = \mathbf{I}$. Next, define the following matrices recursively

$$\mathbf{\Gamma}_0(L) \equiv \mathbf{\Gamma}(L),$$
$$\mathbf{\Gamma}_k(L) \equiv \mathbf{\Gamma}_{k-1}(L)\mathbf{W}_{\lambda_k}\mathbf{B}(L; \lambda_k).$$

The following lemma characterizes useful properties of $\mathbf{\Gamma}_k(L)$.

Lemma 4. *The matrix $\mathbf{\Gamma}_k(L)$ is given by*

$$\mathbf{\Gamma}_k(L) = \begin{bmatrix} \gamma_1^k(L) & \gamma_2^k(L) \\ \gamma_3^k(L)\, g(L) & \gamma_4^k(z)\, g(L) \end{bmatrix}$$

with all $\gamma_i^k(L)$ independent of $g(L)$. Moreover,

$$\delta_k(z) \equiv \frac{\gamma_1^k(z)}{\gamma_2^k(z)}$$

satisfies the following recursive structure:

$$\delta_0(z) = \tau^{-1/2}(1-\rho z), \tag{C.1}$$

$$\delta_k(z) = \frac{1+\delta_{k-1}(\lambda_k)\,\delta_{k-1}(z)}{\delta_{k-1}(\lambda_k)-\delta_{k-1}(z)}\,\frac{z-\lambda_k}{1-\lambda_k z}. \tag{C.2}$$

Finally, for $k \geq 2$, there exists some constant d_k such that

$$\delta_{k-1}(z) = \delta_{k-1}(\lambda_k) + d_k \frac{z-\lambda_k}{1-\lambda_{k-1} z}. \tag{C.3}$$

Using the recursive structure of $\delta_k(z)$, it is straightforward to verify that $\mathbf{A}(z)$ can be written as

$$\mathbf{A}(z) = c_1 \mathbf{\Phi}(z) = \mathbf{A}_{n+1}(z)\,\mathbf{A}_n(z)\ldots\mathbf{A}_2(z)\,\mathbf{A}_1(z),$$

with

$$c_1 = \prod_{k=1}^{n+1} \frac{\gamma_2^{k-1}(\lambda_k)}{\sqrt{\gamma_1^{k-1}(\lambda_k)^2 + \gamma_2^{k-1}(\lambda_k)^2}},$$

$$\mathbf{A}_k(z) = \frac{\sqrt{\gamma_1^{k-1}(\lambda_k)^2 + \gamma_2^{k-1}(\lambda_k)^2}}{\gamma_2^{k-1}(\lambda_k)}\mathbf{B}'\left(z^{-1};\lambda_k\right)\mathbf{W}'_{\lambda_k} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{z-\lambda_k}{1-\lambda_k z} \end{bmatrix} \begin{bmatrix} \delta_{k-1}(\lambda_k) & 1 \\ -1 & \delta_{k-1}(\lambda_k) \end{bmatrix}.$$

The following lemma characterizes a useful property of $\mathbf{\Phi}(z)$.

Lemma 5. *The elements of $\mathbf{\Phi}(z)$ satisfy*

$$\Phi_{12}(z) = -\frac{\prod_{k=1}^{n+1}\left(\delta_{k-1}(\lambda_k)^2+1\right)}{c_2}\frac{\delta_0(z^{-1})}{H(z^{-1})} + \Phi_{22}(z)\delta_{n+1}(z^{-1}),$$

$$\Phi_{22}(z) = c_2 \frac{H(z)\delta_0(z^{-1})\delta_{n+1}(z)+1}{\delta_0(z)\delta_0(z^{-1})+1},$$

where $H(z) \equiv \prod_{k=1}^{n+1}\frac{z-\lambda_k}{1-\lambda_k z}$, and c_2 is some constant.

The equivalence result in Huo and Pedroni (2020) implies that

$$p_{it} = \widetilde{\mathbb{E}}_{it}[q_t],$$

where $\widetilde{\mathbb{E}}_{it}$ is the expectation conditional on the same information set but with a precision of private signals misperceived to be $\tau \equiv (1-\alpha)\tau$. Then, the aggregate action is

$$p_t = \int \widetilde{\mathbb{E}}_{it}[q_t].$$

Since the Hansen-Sargent formula implies

$$\widetilde{\mathbb{E}}_{it}\left[q_t\right] = \frac{1}{\rho}\left(\frac{\mathbf{\Gamma}(L)}{L} - \frac{\mathbf{\Gamma}^*(0)\,\mathbf{A}(L)}{L}\right)_{1^{st}\text{ row}}\begin{bmatrix} u_{it} \\ \eta_t \end{bmatrix},$$

we obtain the following fixed point problem,

$$g(z) = \frac{1}{\rho}\left[\frac{\mathbf{\Gamma}(z)}{z} - \frac{\mathbf{\Gamma}^*(0)\,\mathbf{A}(z)}{z}\right]_{12} = \frac{1}{\rho z}\left(\frac{1}{1-\rho z} - S(z)\right),$$

or

$$h(z) = \frac{1}{\rho}\frac{1}{zC(z)}\left(\frac{1}{(1-\rho z)} - S(z)\right),$$

where

$$S(z) \equiv \begin{bmatrix} 1 & 0 \end{bmatrix}\mathbf{\Gamma}^*(0)\,\mathbf{A}(z)\begin{bmatrix} 0 \\ 1 \end{bmatrix} = c_1\Big(\gamma_1^{n+1}(0)\Phi_{12}(z) + \gamma_2^{n+1}(0)\Phi_{22}(z)\Big).$$

For $h(z)$ to be an equilibrium, it has to be that

$$\frac{1}{(1-\rho\lambda_i)} = S(\lambda_i), \quad \text{for all } i \in \{1,\dots,n\}, \tag{C.4}$$

so that the poles in $C(z)$ can be removed.

Next, we show that with a finite number of inside roots, there does not exist an equilibrium.

Case 1: $n = 1$.. First let $x \equiv \rho + \frac{1+\tau}{\rho}$ and notice that $|\rho| < 1$ and $\tau > 0$ imply $|x| > 2$. In this case $S(z)$ takes a simple form and we can calculate

$$S(\lambda_1) - \frac{1}{1-\rho\lambda_1} = \frac{\tau\lambda_1}{(1-\lambda_1\rho)}\frac{x-\lambda_1}{(1-\lambda_1^2) + (x-2\lambda_1)(\rho-x)},$$

so that equation (C.4) implies that $\lambda_1 = x$ which is outside the unit circle.

Case 2: $n = 2$.. Suppose that $\lambda_1 \neq \lambda_2$, then

$$\frac{S(\lambda_1)}{S(\lambda_2)} - \frac{1-\rho\lambda_2}{1-\rho\lambda_1} = \frac{\tau(\lambda_2-\lambda_1)}{(1-\rho\lambda_1)}\frac{(x-\lambda_1)(x-\lambda_2)-1}{(1-\lambda_1^2)(\rho-x+\lambda_2) - (x-2\lambda_1)((x-\lambda_2)(\rho-x)+1)},$$

and equation (C.4) for $i = \{1,2\}$ implies that $(x-\lambda_1)(x-\lambda_2) = 1$, which implies that either $|\lambda_1| > 1$ or $|\lambda_2| > 1$. Next, if $\lambda_1 = \lambda_2 = \lambda$, we have that

$$S(\lambda) - \frac{1}{1-\rho\lambda} = \frac{\tau\lambda}{1-\lambda\rho}\frac{-3\lambda^2 + 3x\lambda - (x^2-1)}{(1-\lambda^2)(4\lambda - x(1+\lambda^2)) + (4x\lambda - x^2 + \lambda^4 - 6\lambda^2 + 1)(\rho-x)},$$

and equation (C.4) implies that $3\lambda^2 - 3x\lambda + x^2 = 1$. Notice that the discriminant of this quadratic equation on λ is $9x^2 - 12(x^2-1)$ and that it is negative whenever $|x| > 2$. Therefore, the solutions are complex. Complex λ's are allowed but necessitate a conjugate which is not possible in this case since we have assumed $\lambda_1 = \lambda_2$.

Case 3: $n > 2$.. From the definition of $S(z)$ and Lemma 5, it follows that

$$S(\lambda_i) = c_1 c_2 \gamma_2^{n+1}(0) \frac{1 + \delta_{n+1}(0)\delta_{n+1}(\lambda_i^{-1})}{1 + \delta_0(\lambda_i)\delta_0(\lambda_i^{-1})}.$$

Equation (C.3) together with the fact that $\lambda_{n+1} = 0$ implies that

$$\delta_{n+1}(z) = \delta_{n+1}(0) + d_{n_1} z,$$

for some constant d_{n+1}. Thus, we can rewrite $S(\lambda_i)$ as

$$S(\lambda_i) = c_1 c_2 \gamma_2^{n+1}(0) \frac{1 + \delta_{n+1}(0)(\delta_{n+1}(0) + d_{n+1}\lambda_i^{-1})}{1 + \delta_0(\lambda_i)\delta_0(\lambda_i^{-1})}.$$

Suppose that the solution to the system of equations (C.4) includes λ_i, λ_j, λ_k different from each other and all inside the unit circle. It follows that

$$\frac{S(\lambda_i)}{S(\lambda_j)} = \frac{1 - \rho\lambda_j}{1 - \rho\lambda_i}, \quad \text{or} \quad \frac{1 + \delta_{n+1}(0)(\delta_{n+1}(0) + d_{n+1}\lambda_i^{-1})}{1 + \delta_0(\lambda_i)\delta_0(\lambda_i^{-1})} = \frac{1 + \delta_{n+1}(0)(\delta_{n+1}(0) + d_{n+1}\lambda_j^{-1})}{1 + \delta_0(\lambda_j)\delta_0(\lambda_j^{-1})},$$

which can be written as

$$\delta_{n+1}(0)\left(\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_j\right)\right) + \left(\left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_j\right) + \lambda_i\lambda_j\tau\right)\rho\left(1 + \delta_{n+1}^2(0)\right) = 0.$$

Suppose that $\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_j\right) \neq 0$, then

$$\frac{\left(1 - \lambda_i\lambda_j\right)\tau}{\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_j\right)} = 1 - \frac{d_{n+1}\delta_{n+1}(0)}{\rho\left(1 + \delta_{n+1}^2(0)\right)}.$$

Similarly, if $\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_k\right) \neq 0$, then

$$\frac{\left(1 - \lambda_i\lambda_k\right)\tau}{\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_k\right)} = 1 - \frac{d_{n+1}\delta_{n+1}(0)}{\rho\left(1 + \delta_{n+1}^2(0)\right)}.$$

Combining the two conditions above, we have

$$\frac{\left(1 - \lambda_i\lambda_j\right)}{\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_j\right)} = \frac{\left(1 - \lambda_i\lambda_k\right)}{\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_k\right)},$$

which implies

$$\tau\lambda_i + (1 - \lambda_i\rho)(1\lambda_i - \rho) = 0 \quad \Rightarrow \quad 1 + \delta_0(\lambda_i)\delta_0(\lambda_i^{-1}) = 0.$$

This cannot be the case since $S(\lambda_i)$ must be a finite number. Therefore, if there exists a solution, it has to be that

$$\tau + \left(1 - \rho\lambda_i\right)\left(1 - \rho\lambda_j\right) = 0.$$

This condition, however, implies that λ_i and λ_j cannot be simultaneously within the unit circle.

Next, suppose there exits a solution with $\lambda_i = \lambda$, for all $i \in \{1, \ldots, n\}$ with $|\lambda| < 1$. Let $\{\lambda_{1,k}\}_{k=0}^{\infty}, \ldots, \{\lambda_{n,k}\}_{k=0}^{\infty}$ be n sequences such that $\lim_{k\to\infty} \lambda_{i,k} = \lambda$ and $\lambda_{i,k} \neq \lambda_{j,k}$ for all $i, j \in \{1, \ldots, n\}$ and all $k \geq 0$. Define

$$\omega_k \equiv \sum_{i=1}^{n} \left| S(\lambda_{i,k}) - \frac{1}{1-\rho\lambda_{i,k}} \right|.$$

By continuity of $S(\cdot)$, ω_k approaches 0 as k goes to infinity, since $\lambda_{i,k}$ approaches to λ for all $i \in \{1, \ldots, n\}$. However, since $\lambda_{i,k} \neq \lambda_{j,k}$, as established above, only if some $\lambda_{i,k}$ are outside the unit circle, can ω_k approach 0. Since $|\lambda| < 1$, as k goes to infinity, all $|\lambda_{i,k}| < \delta$ for any $\delta < 1$, which implies that ω_k cannot be close to zero. This is a contradiction. The case where all λ_i equal to each other, except for one, can be dealt with in a similar way, which concludes the proof.

C.1 Proof of Lemma 4

We prove this lemma by induction. For $k = 0$, this is clearly the case. For $k \geq 1$, suppose that

$$\mathbf{\Gamma}_k(z) = \begin{bmatrix} \gamma_1^k(z) & \gamma_2^k(z) \\ \gamma_3^k(z)\, g(z) & \gamma_4^k(z)\, g(z) \end{bmatrix}.$$

Then, it follows that

$$\mathbf{\Gamma}_{k+1}(z) = \mathbf{\Gamma}_{k-1}(z)\, \mathbf{W}_{\lambda_{k+1}} \mathbf{B}(z; \lambda_{k+1}),$$

with

$$\mathbf{W}_{\lambda_{k+1}} = \frac{1}{\sqrt{\gamma_1^k(\lambda_{k+1})^2 + \gamma_2^k(\lambda_{k+1})^2}} \begin{bmatrix} \gamma_1^k(\lambda_{k+1}) & -\gamma_2^k(\lambda_{k+1}) \\ \gamma_2^k(\lambda_{k+1}) & \gamma_1^k(\lambda_{k+1}) \end{bmatrix}.$$

So that

$$\begin{aligned} \mathbf{\Gamma}_{k+1}(z) &= \frac{1}{\sqrt{\gamma_1^k(\lambda_{k+1})^2 + \gamma_2^k(\lambda_{k+1})^2}} \begin{bmatrix} \gamma_1^k(z) & \gamma_2^k(z) \\ \gamma_3^k(z)\, g(z) & \gamma_4^k(z)\, g(z) \end{bmatrix} \begin{bmatrix} \gamma_1^k(\lambda_{k+1}) & -\gamma_2^k(\lambda_{k+1}) \\ \gamma_2^k(\lambda_{k+1}) & \gamma_1^k(\lambda_{k+1}) \end{bmatrix} \\ &= \frac{\begin{bmatrix} \left(\gamma_1^k(\lambda_{k+1})\,\gamma_1^k(z) + \gamma_2^k(\lambda_{k+1})\,\gamma_2^k(z)\right) & \left(\gamma_1^k(\lambda_{k+1})\,\gamma_2^k(z) - \gamma_2^k(\lambda_{k+1})\,\gamma_1^k(z)\right) \frac{1-z\lambda_{k+1}}{z-\lambda_{k+1}} \\ \left(\gamma_1^k(\lambda_{k+1})\,\gamma_2^k(z) + \gamma_2^k(\lambda_{k+1})\,\gamma_4^k(z)\right) g(z) & \left(\gamma_1^k(\lambda_{k+1})\,\gamma_4^k(z) - \gamma_2^k(\lambda_{k+1})\,\gamma_3^k(z)\right) \frac{1-z\lambda_{k+1}}{z-\lambda_{k+1}}\, g(z) \end{bmatrix}}{\sqrt{\gamma_1^k(\lambda_{k+1})^2 + \gamma_2^k(\lambda_{k+1})^2}} \end{aligned}$$

which has the desired structure. Note, moreover, that γ_1^{k+1} and γ_2^{k+1} satisfy

$$\begin{aligned} \gamma_1^{k+1}(z) &= \frac{\gamma_1^k(\lambda_{k+1})\,\gamma_1^k(z) + \gamma_2^k(\lambda_{k+1})\,\gamma_2^k(z)}{\sqrt{\gamma_1^k(\lambda_{k+1})^2 + \gamma_2^k(\lambda_{k+1})^2}}, \\ \gamma_2^{k+1}(z) &= \frac{\gamma_1^k(\lambda_{k+1})\,\gamma_2^k(z) - \gamma_2^k(\lambda_{k+1})\,\gamma_1^k(z)}{\sqrt{\gamma_1^k(\lambda_{k+1})^2 + \gamma_2^k(\lambda_{k+1})^2}} \frac{1-z\lambda_{k+1}}{z-\lambda_{k+1}}. \end{aligned}$$

It follows from these recursions that $\delta_k(z)$ satisfies equation (C.2). To prove equation (C.3), first notice that, for $k = 1$, it follows from (C.1) and (C.2) that

$$\delta_1(z) = \delta_1(\lambda_1) + \frac{(1-\rho\lambda_1)(\lambda_1-\rho)+\lambda_1\tau}{\sqrt{\tau}\rho\left(1-\lambda_1^2\right)}\frac{z-\lambda_1}{1-\lambda_1 z}.$$

Next, suppose that there exists d_k such that

$$\delta_k(z) = \delta_k(\lambda_{k+1}) + d_k\frac{z-\lambda_{k+1}}{1-\lambda_k z},$$

then, equation (C.2) implies that

$$\delta_{k+1}(z) = \delta_{k+1}(\lambda_{k+1}) + \frac{\left(1+\delta_k(\lambda_{k+1})^2\right)(\lambda_k-\lambda_{k+1}) - d_k\left(1-\lambda_{k+1}^2\right)\delta_k(\lambda_{k+1})}{d_k\left(1-\lambda_{k+1}^2\right)}\frac{z-\lambda_{k+1}}{1-\lambda_{k+1}z},$$

which, again by induction, establishes the result. □

C.2 Proof of Lemma 5

From the definition of $\mathbf{A}_k(z)$ and equation (C.2), it follows that

$$\mathbf{A}_k(z)\begin{bmatrix}1\\-\delta_{k-1}(z)\end{bmatrix} = (\delta_{k-1}(\lambda_k)-\delta_{k-1}(z))\begin{bmatrix}1\\-\delta_k(z)\end{bmatrix}.$$

Define

$$H(z) \equiv \prod_{k=1}^{n+1}\frac{z-\lambda_k}{1-\lambda_k z}, \quad \text{and} \quad G(z) \equiv \prod_{k=1}^{n+1}(\delta_{k-1}(\lambda_k)-\delta_{k-1}(z)).$$

Since $\mathbf{A}(z) = c_1\mathbf{\Phi}(z)$, it follows that

$$G(z) = c_2(1-z\lambda_{n+1})H(z) = c_2H(z), \tag{C.5}$$

for some constant c_2, and that

$$\mathbf{\Phi}(z)\begin{bmatrix}1\\-\delta_0(z)\end{bmatrix} = G(z)\begin{bmatrix}1\\-\delta_{n+1}(z)\end{bmatrix}. \tag{C.6}$$

For a function $f(z)$, define the tilde operator as $\widetilde{f}(z) = f(1/z)$. Then,

$$\widetilde{\mathbf{A}}_k(z)'\mathbf{A}_k(z) = \left(\delta_{k-1}(\lambda_k)^2+1\right)\begin{bmatrix}1&0\\0&1\end{bmatrix},$$

and, therefore,

$$\widetilde{\mathbf{\Phi}}(z)'\mathbf{\Phi}(z) = c_3\begin{bmatrix}1&0\\0&1\end{bmatrix}, \quad \text{where} \quad c_3 = \prod_{k=1}^{n+1}\left(\delta_{k-1}(\lambda_k)^2+1\right).$$

Next, apply the tilde transformation to equation (C.6) to obtain

$$\widetilde{\mathbf{\Phi}}(z)\begin{bmatrix}1\\-\widetilde{\delta}_0(z)\end{bmatrix}=\widetilde{G}(z)\begin{bmatrix}1\\-\widetilde{\delta}_{n+1}(z)\end{bmatrix}.$$

Transposing and multiplying from the right with $\mathbf{\Phi}(z)$ yields

$$c_3\begin{bmatrix}1 & -\widetilde{\delta}_0(z)\end{bmatrix}=\begin{bmatrix}1 & -\widetilde{\delta}_0(z)\end{bmatrix}\widetilde{\mathbf{\Phi}}(z)'\mathbf{\Phi}(z)=\widetilde{G}(z)\begin{bmatrix}1 & -\tilde{\delta}_{n+1}(z)\end{bmatrix}\mathbf{\Phi}(z).$$

Together with the equation (C.6) we obtain four linear equations for the four entries of $\mathbf{\Phi}(z)$,

$$\begin{aligned}\Phi_{11}(z)-\Phi_{12}(z)\delta_0(z)&=G(z)\\ \Phi_{21}(z)-\Phi_{22}(z)\delta_0(z)&=-\delta_{n+1}(z)G(z)\\ (\Phi_{11}(z)-\Phi_{21}(z)\widetilde{\delta}_{n+1}(z))\widetilde{G}(z)&=c_3\\ (\Phi_{12}(z)-\Phi_{22}(z)\widetilde{\delta}_{n+1}(z))\widetilde{G}(z)&=-c_3\widetilde{\delta}_0(z).\end{aligned}$$

The rank of the system is 3. Use the first, second and fourth equations to express $\Phi_{11}(z), \Phi_{12}(z), \Phi_{21}(z)$ in terms of $\Phi_{22}(z)$. The third equation does not allow to solve for $\Phi_{22}(z)$, rather it collapses to

$$\left(1+\delta_{n+1}(z)\widetilde{\delta}_{n+1}(z)\right)G(z)\widetilde{G}(z)=\left(1+\delta_0(z)\widetilde{\delta}_0(z)\right)c_3. \tag{C.7}$$

We can, now, determine $\Phi_{22}(z)$ from[24]

$$\Phi_{11}(z)\Phi_{22}(z)-\Phi_{12}(z)\Phi_{21}(z)=\det(\mathbf{\Phi}(z))=c_3H(z),$$

which implies

$$\Phi_{22}(z)=\frac{(G(z)\widetilde{\delta}_0(z)\delta_{n+1}(z)+\widetilde{G}(z)H(z))c_3}{(1+\delta_{n+1}(z)\widetilde{\delta}_{n+1}(z))G(z)\widetilde{G}(z)}.$$

Together with (C.7), this can be simplified to

$$\Phi_{22}(z)=\frac{G(z)\widetilde{\delta}_0(z)\delta_{n+1}(z)+\widetilde{G}(z)H(z)}{\delta_0(z)\widetilde{\delta}_0(z)+1}.$$

Applying the tilde operation to equation (C.5) yields

$$\widetilde{G}(z)=c_2\widetilde{H}(z).$$

Finally, it follows from the definition of $H(z)$ that $\widetilde{H}(z)H(z)=1$, and therefore

$$\Phi_{22}(z)=\frac{c_2H(z)\widetilde{\delta}_0(z)\delta_{n+1}(z)+c_2}{\delta_0(z)\widetilde{\delta}_0(z)+1}. \qquad \square$$

[24] Notice that

$$\det(\mathbf{A}_k)=\left(1+\delta_{k-1}(\lambda_k)^2\right)\frac{z-\lambda_k}{1-z\lambda_k}.$$

D. Proof of Theorem 1

Suppose that when $\alpha = 0$ the equilibrium is invertible, otherwise the result is trivial. Section D.1 characterizes the equilibrium assuming invertibility. Using this characterization, Section D.2 takes the limit as α increases to 1 and shows that in it the equilibrium cannot be invertible.

D.1 Solution Assuming Invertibility

Suppose that the equilibrium is invertible, then the information set of agent i in period t is given by $\mathcal{I}_{it} \equiv \{\mathbf{x}_{i\tau}, \mathbf{z}_{\tau}, \boldsymbol{\eta}_{\tau-1}, \boldsymbol{u}_{i\tau-1}\}_{\tau=-\infty}^{t}$. Therefore,

$$\mathbb{E}_{it}[\xi_{it}] = \mathbb{E}[\boldsymbol{d}(L)\boldsymbol{\eta}_t + \boldsymbol{e}(L)\boldsymbol{u}_{it}|\mathcal{I}_{it}] = (\boldsymbol{d}(L)-\boldsymbol{d}_0)\boldsymbol{\eta}_t + \boldsymbol{d}_0\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] + (\boldsymbol{e}(L)-\boldsymbol{e}_0)\boldsymbol{u}_{it} + \boldsymbol{e}_0\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}],$$

$$\mathbb{E}_{it}[a_t] = \mathbb{E}[\boldsymbol{g}(L)\boldsymbol{\eta}_t|\mathcal{I}_{it}] = (\boldsymbol{g}(L)-\boldsymbol{g}_0)\boldsymbol{\eta}_t + \boldsymbol{g}_0\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}].$$

Moreover, since

$$\mathbf{x}_{it} - (\mathbf{A}(L)-\mathbf{A}_0)\boldsymbol{\eta}_t - (\mathbf{B}(L)-\mathbf{B}_0)\boldsymbol{u}_{it} = \mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}_0\boldsymbol{u}_{it}, \quad \text{and} \quad \mathbf{z}_t - (\mathbf{C}(L)-\mathbf{C}_0)\boldsymbol{\eta}_t = \mathbf{C}_0\boldsymbol{\eta}_t,$$

it follows that $\mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}_0\boldsymbol{u}_{it}$ is a noisy signal about $\boldsymbol{\eta}_t$ and $\boldsymbol{u}_{it}$, and $\mathbf{C}_0\boldsymbol{\eta}_t$ is a noisy signal about $\boldsymbol{\eta}_t$, which allows us to calculate

$$\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] = \begin{bmatrix}\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}_0\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top \\ \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}_0\boldsymbol{u}_{it} \\ \mathbf{C}_0\boldsymbol{\eta}_t\end{bmatrix},$$

$$\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}] = \begin{bmatrix}\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}_0\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top \\ \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}_0\boldsymbol{u}_{it} \\ \mathbf{C}_0\boldsymbol{\eta}_t\end{bmatrix}.$$

Therefore,

$$a_{it} = \varphi\left((\boldsymbol{d}(L)-\boldsymbol{d}_0)\boldsymbol{\eta}_t + \boldsymbol{d}_0\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] + (\boldsymbol{e}(L)-\boldsymbol{e}_0)\boldsymbol{u}_{it} + \boldsymbol{e}_0\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}]\right) + \alpha\left((\boldsymbol{g}(L)-\boldsymbol{g}_0)\boldsymbol{\eta}_t + \boldsymbol{g}_0\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}]\right),$$

which can be reorganized as

$$a_{it} = \left[\varphi(\boldsymbol{d}(L)-\boldsymbol{d}_0) + \alpha(\boldsymbol{g}(L)-\boldsymbol{g}_0)\right]\boldsymbol{\eta}_t + \varphi(\boldsymbol{e}(L)-\boldsymbol{e}_0)\boldsymbol{u}_{it} + \left(\varphi\boldsymbol{d}_0 + \alpha\boldsymbol{g}_0\right)\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] + \varphi\boldsymbol{e}_0\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}].$$

Consistency requires, in particular, that

$$\boldsymbol{g}(L) = \varphi(\boldsymbol{d}(L)-\boldsymbol{d}_0) + \alpha(\boldsymbol{g}(L)-\boldsymbol{g}_0) + \left(\varphi\boldsymbol{d}_0 + \alpha\boldsymbol{g}_0\right)\boldsymbol{k}_1\boldsymbol{\Omega} + \varphi\boldsymbol{e}_0\boldsymbol{k}_2\boldsymbol{\Omega},$$

where

$$\boldsymbol{k}_1 \equiv \begin{bmatrix}\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}, \quad \boldsymbol{k}_2 \equiv \begin{bmatrix}\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{0}\end{bmatrix}, \quad \text{and} \quad \boldsymbol{\Omega} \equiv \begin{bmatrix}\mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}_0\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top \\ \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}_0 \\ \mathbf{C}_0\end{bmatrix}.$$

We can rewrite the equation above as

$$(1-\alpha)\,\boldsymbol{g}(L) = \varphi\,\boldsymbol{d}(L) + \left(\varphi \boldsymbol{d}_0 + \alpha \boldsymbol{g}_0\right)(\boldsymbol{k}_1 \boldsymbol{\Omega} - \mathbf{I}) + \varphi \boldsymbol{e}_0 \boldsymbol{k}_2 \boldsymbol{\Omega}$$

It is useful to replace the lag operator in this equation with an arbitrary complex number z. Evaluating this equation at $z = 0$, for instance, implies the following equilibrium condition,

$$\boldsymbol{g}_0 = \varphi\,(\boldsymbol{d}_0 \boldsymbol{k}_1 + \boldsymbol{e}_0 \boldsymbol{k}_2)\,\boldsymbol{\Omega}\,(\mathbf{I} - \alpha \boldsymbol{k}_1 \boldsymbol{\Omega})^{-1}.$$

Notice that, using the block-matrix inversion formula, $\boldsymbol{g}_0$ can be rewritten as

$$\boldsymbol{g}_0 = \frac{\varphi}{1-\alpha}\boldsymbol{d}_0 \mathbf{E} + \varphi\,(\boldsymbol{d}_0\,(\mathbf{I}-\mathbf{E})\,\mathbf{D} + \boldsymbol{e}_0\,\mathbf{F}) \sum_{j=0}^{\infty} \alpha^j \left[(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})\right]^j, \tag{D.1}$$

where

$$\begin{aligned}
\mathbf{D} &\equiv \boldsymbol{\Sigma}_\eta^2 \mathbf{A}_0^\top \left(\mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2 \mathbf{A}_0^\top + \mathbf{B}_0 \boldsymbol{\Sigma}_u^2 \mathbf{B}_0^\top\right)^{-1} \mathbf{A}_0, \\
\mathbf{E} &\equiv \boldsymbol{\Sigma}_\eta^2 \mathbf{C}_0^\top \left(\mathbf{C}_0 \boldsymbol{\Sigma}_\eta^2 \mathbf{C}_0^\top\right)^{-1} \mathbf{C}_0, \\
\mathbf{F} &\equiv \boldsymbol{\Sigma}_u^2 \mathbf{B}_0^\top \left(\mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2 \mathbf{A}_0^\top + \mathbf{B}_0 \boldsymbol{\Sigma}_u^2 \mathbf{B}_0^\top\right)^{-1} \mathbf{A}_0,
\end{aligned}$$

and we have used the fact that $\mathbf{E}$ is idempotent. Finally, substituting the expression for $\boldsymbol{g}_0$ into the equation for $\boldsymbol{g}(z)$ we obtain

$$\boldsymbol{g}(z) = \boldsymbol{g}_0 + \frac{\varphi}{1-\alpha}(\boldsymbol{d}(z) - \boldsymbol{d}_0).$$

D.2 Taking the Limit as α Increases to 1

Let

$$\boldsymbol{g}(z) = \frac{1-\alpha}{\varphi}\boldsymbol{g}_0 + \boldsymbol{d}(z) - \boldsymbol{d}_0,$$

so that

$$\boldsymbol{g}(z) = \frac{\varphi}{1-\alpha}\boldsymbol{g}(z),$$

and notice that this is well defined for all $\alpha < 1$ and that, if $\lim_{\alpha\to1^-} \det \begin{bmatrix} \mathbf{C}(z) & \boldsymbol{g}(z) \end{bmatrix}^\top$ has an inside root, then there exists $\alpha < 1$ high enough such that $\det \begin{bmatrix} \mathbf{C}(z) & \boldsymbol{g}(z) \end{bmatrix}^\top$ is well defined and has an inside root. It is, in fact, sufficient to show that

$$\lim_{\alpha\to1^-} \det \begin{bmatrix} \mathbf{C}_0 & \boldsymbol{g}_0 \end{bmatrix}^\top = 0.$$

Accordingly, using equation (D.1) we have that

$$\lim_{\alpha\to1^-} \boldsymbol{g}_0 = \lim_{\alpha\to1^-} \frac{1-\alpha}{\varphi}\boldsymbol{g}_0 = \boldsymbol{d}_0 \mathbf{E} + \lim_{\alpha\to1^-}(1-\alpha)\,(\boldsymbol{d}_0\,(\mathbf{I}-\mathbf{E})\,\mathbf{D} + \boldsymbol{e}_0 \mathbf{F}) \sum_{j=0}^{\infty} \alpha^j \left[(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})\right]^j.$$

In order to proceed, it is useful to establish the following lemma.

Lemma 6. *The matrix* $(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})$ *has all eigenvalues in* $[0, 1)$.

Proof. Notice that $(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2$ is symmetric, so that $\left((\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2\right)^\top = (\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2$. Let $\mathbf{M} = \mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta$ and $\mathbf{N} = \mathbf{B}_0\boldsymbol{\Sigma}_u$, then it follows that

$$\begin{aligned}
\mathbf{M}\mathbf{M}^\top &= \mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta\left((\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta\right)^\top\mathbf{A}_0^\top \\
&= \mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta\left((\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2\boldsymbol{\Sigma}_\eta^{-1}\right)^\top\mathbf{A}_0^\top \\
&= \mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta\boldsymbol{\Sigma}_\eta^{-1}\left((\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2\right)^\top\mathbf{A}_0^\top \\
&= \mathbf{A}_0\,(\mathbf{I}-\mathbf{E})\,\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top.
\end{aligned}$$

Therefore

$$(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E}) = \boldsymbol{\Sigma}_\eta\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\mathbf{M}\boldsymbol{\Sigma}_\eta^{-1},$$

so that $(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})$ is similar to $\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\mathbf{M}$. Let $n \equiv \dim(\boldsymbol{u}_{it})$ and $m \equiv \dim(\boldsymbol{\eta}_t)$, then, it follow that $\mathbf{N}$ is $n \times n$ and $\mathbf{M}$ is $m-1 \times m$. If $n \geq m$,

$$\operatorname{spectrum}\left(\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\mathbf{M}\right) \cup \{0_1,\ldots,0_{n-m}\} = \operatorname{spectrum}\left(\mathbf{M}\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\right)$$

while if $n \leq m$,

$$\operatorname{spectrum}\left(\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\mathbf{M}\right) = \operatorname{spectrum}\left(\mathbf{M}\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\right) \cup \{0_1,\ldots,0_{m-n}\}.$$

The matrix $\mathbf{M}\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}$ is the product of a positive semi-definite with a positive definite matrix, so must have positive eigenvalues while $\mathbf{N}\mathbf{N}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}$ is the product of two positive definite matrices, and, therefore, has strictly positive eigenvalues. Finally, since

$$\mathbf{M}\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1} = \mathbf{I} - \mathbf{N}\mathbf{N}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1},$$

it follows that $\mathbf{M}\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}$ must have eigenvalues lower than 1. Hence,

$$\begin{aligned}
\operatorname{spectrum}\left((\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})\right) &= \operatorname{spectrum}\left(\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\mathbf{M}\right) \\
&= \operatorname{spectrum}\left(\mathbf{M}\mathbf{M}^\top\left(\mathbf{M}\mathbf{M}^\top+\mathbf{N}\mathbf{N}^\top\right)^{-1}\right) \subset [0,1).
\end{aligned}$$

□

The lemma implies that

$$\sum_{j=0}^{\infty}\alpha^j\left[(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})\right]^j$$

is well defined and finite, so that

$$\lim_{\alpha\to1^-}(1-\alpha)\left(\boldsymbol{d}_0\,(\mathbf{I}-\mathbf{E})\,\mathbf{D}+\boldsymbol{e}_0\mathbf{F}\right)\sum_{j=0}^{\infty}\alpha^j\left[(\mathbf{I}-\mathbf{E})\,\mathbf{D}\,(\mathbf{I}-\mathbf{E})\right]^j = \mathbf{0}.$$

Therefore,

$$\lim_{\alpha\to1^-}\boldsymbol{g}_0 = \boldsymbol{d}_0\mathbf{E},$$

and, using the definition of $\mathbf{E}$,

$$\boldsymbol{g}_0 = \boldsymbol{d}_0 \boldsymbol{\Sigma}_\eta^2 \mathbf{C}_0^\top \left(\mathbf{C}_0 \boldsymbol{\Sigma}_\eta^2 \mathbf{C}_0^\top\right)^{-1} \mathbf{C}_0 = \boldsymbol{a}\mathbf{C}_0,$$

for some vector $\boldsymbol{a}$. Finally, notice that

$$\det \begin{bmatrix} \mathbf{C}_0 \\ \boldsymbol{g}_0 \end{bmatrix} = \det \begin{bmatrix} \mathbf{C}_0 \\ \boldsymbol{a}\mathbf{C}_0 \end{bmatrix} = 0,$$

which implies that $z = 0$ is a root of $\det \begin{bmatrix} \mathbf{C}(z) & \boldsymbol{g}(z) \end{bmatrix}^\top$ and, therefore, $\begin{bmatrix} \mathbf{C}(L) & \boldsymbol{g}(L) \end{bmatrix}^\top$ is not invertible for α close enough to zero (from below).

E. Proof of Proposition 5.2

The following lemma establishes a type of law of iterated expectations.

Lemma 7. *If $\mathcal{I}_{it} \supseteq \{\boldsymbol{\eta}_\tau, \boldsymbol{u}_{i\tau}\}_{\tau=-\infty}^{t-1}$, then, for any stochastic variable $y_{i,t+j} = \boldsymbol{f}(L)\boldsymbol{\eta}_{t+j} + \boldsymbol{g}(L)\boldsymbol{u}_{i,t+j} = \sum_{s=0}^{\infty} \boldsymbol{f}_s \boldsymbol{\eta}_{t+j-s} + \sum_{s=0}^{\infty} \boldsymbol{g}_s \boldsymbol{u}_{i,t+j-s}$,*

$$\mathbb{E}_{it}[\overline{\mathbb{E}}_{t+k}[y_{i,t+j}]] = \mathbb{E}_{it}[y_{i,t+j}], \qquad \textit{for all } k \geq 1.$$

Proof. Let $\boldsymbol{f}_s \equiv \mathbf{0}$, and $\boldsymbol{g}_s \equiv \mathbf{0}$, for all $s < 0$ and note that

$$\mathbb{E}_{i,t+k}[y_{i,t+j}] = \sum_{s=1}^{\infty} (\boldsymbol{f}_{j-k+s}\boldsymbol{\eta}_{t+k-s} + \boldsymbol{g}_{j-k+s}\boldsymbol{u}_{i,t+k-s}) + \boldsymbol{f}_{j-k}(\mathbf{H}_\eta \boldsymbol{\eta}_{t+k} + \mathbf{H}_u \boldsymbol{u}_{i,t+k}) + \boldsymbol{g}_{j-k}(\mathbf{P}_\eta \boldsymbol{\eta}_{t+k} + \mathbf{P}_u \boldsymbol{u}_{i,t+k}),$$

where we have used the fact that: (1) $\mathbb{E}_{i,t+k}[\boldsymbol{\eta}_{t+j}] = \mathbb{E}_{i,t+k}[\boldsymbol{u}_{i,t+j}] = \mathbf{0}$, for $j > k$; (2) $\mathbb{E}_{i,t+k}[\boldsymbol{\eta}_{t+j}] = \boldsymbol{\eta}_{t+j}$, and $\mathbb{E}_{i,t+k}[\boldsymbol{u}_{i,t+j}] = \boldsymbol{u}_{i,t+j}$, for $j < k$; and (3) $\mathbb{E}_{i,t+k}[\boldsymbol{\eta}_{t+k}] = \mathbf{H}_\eta \boldsymbol{\eta}_{t+k} + \mathbf{H}_u \boldsymbol{u}_{it+k}$, and $\mathbb{E}_{i,t+k}[\boldsymbol{u}_{i,t+k}] = \mathbf{P}_\eta \boldsymbol{\eta}_{t+k} + \mathbf{P}_u \boldsymbol{u}_{it+k}$, for some constant matrices $\mathbf{H}_\eta$, $\mathbf{H}_u$, $\mathbf{P}_\eta$, and $\mathbf{P}_u$.

In aggregate,

$$\overline{\mathbb{E}}_{t+k}[y_{i,t+j}] = \sum_{s=1}^{\infty} \boldsymbol{f}_{j-k+s}\boldsymbol{\eta}_{t+k-s} + \boldsymbol{f}_{j-k}\mathbf{H}_\eta \boldsymbol{\eta}_{t+k} + \boldsymbol{g}_{j-k}\mathbf{P}_\eta \boldsymbol{\eta}_{t+k}.$$

Consider agent i's the inference in period t,

$$\mathbb{E}_{it}\left[\overline{\mathbb{E}}_{t+k}[y_{i,t+j}]\right] = \sum_{s=1}^{\infty} \boldsymbol{f}_{j-s}\boldsymbol{\eta}_{t-s} + \boldsymbol{f}_j(\mathbf{H}_\eta \boldsymbol{\eta}_t + \mathbf{H}_u \boldsymbol{u}_{it}) = \mathbb{E}_{it}[y_{i,t+j}],$$

where the last equality follows from the same three facts listed above (with $k = 0$). □

Next, substituting the best-response (5.7) into itself, using the law of iterated expectations, we get

$$a_{it} = \mathbb{E}_{it}[\varphi\xi_{it} + (\alpha + \gamma(L))a_t + \beta(L)(\varphi\xi_{it} + (\alpha + \gamma(L))a_t + \beta(L)a_{it})],$$

and iterating on this procedure, using the fact that $\|\beta(L)\| < 1$ in the operator norm, leads to

$$a_{it} = \mathbb{E}_{it}[\varphi\widehat{\xi}_{it} + (\alpha + \kappa(L))a_t], \quad \text{where} \quad \widehat{\xi}_{it} \equiv \frac{1}{1-\beta(L)}\xi_{it} \quad \text{and} \quad \kappa(L) \equiv \frac{\gamma(L) + \alpha\beta(L)}{1-\beta(L)}. \tag{E.1}$$

Notice that $\kappa_0 = 0$. Aggregating implies

$$a_t = \overline{\mathbb{E}}_t[\varphi \widehat{\xi}_{it} + (\alpha + \kappa(L))a_t].$$

Multiplying both sides by $\kappa(L)$ and considering the inference of agent i in period t, we have that,

$$\mathbb{E}_{it}[\kappa(L)a_t] = \mathbb{E}_{it}\left[\kappa(L)\overline{\mathbb{E}}_t[\varphi \widehat{\xi}_{it} + (\alpha + \kappa(L))a_t]\right].$$

Since the equilibrium is invertible, $\mathcal{I}_{it} \supseteq \{\boldsymbol{\eta}_\tau, \boldsymbol{u}_{i\tau}\}_{\tau=-\infty}^{t-1}$, so that, using Lemma 7, it follows that,

$$\mathbb{E}_{it}[\kappa(L)a_t] = \mathbb{E}_{it}[\varphi\kappa(L)\widehat{\xi}_{it} + \kappa(L)(\alpha + \kappa(L))a_t],$$

and, iterating on this procedure, using the fact that $||\alpha + \kappa(L)|| < 1$,[25] we obtain

$$\mathbb{E}_{it}[\kappa(L)a_t] = \mathbb{E}_{it}\left[\varphi \frac{\kappa(L)}{1 - (\alpha + \kappa(L))}\widehat{\xi}_{it}\right].$$

The result follows from substituting this fact and the definitions of $\widehat{\xi}_{it}$ and $\kappa(L)$ into equation (E.1).

F. Invertibility in New Keynesian Model with Incomplete Information

F.1 Firms: The New Keynesian Phillips Curve

The optimal reset price solves the following problem:

$$P^*_{i,t} = \arg\max_{P_{i,t}} \sum_{k=0}^{\infty} (\delta\theta)^k \, \mathbb{E}_{i,t}\left[Q_{t|t+k}\left(P_{i,t}Y_{i,t+k|t} - P_{t+k}mc_{t+k}Y_{i,t+k|t}\right)\right]$$

subject to the demand equation, $Y_{i,t+k} = \left(\frac{P_{i,t}}{P_{t+k}}\right)^{-\varepsilon} Y_{t+k}$, where $Q_{t|t+k}$ is the stochastic discount factor between t and $t+k$, Y_{t+k} and P_{t+k} are, respectively, aggregate income and the aggregate price level in period $t+k$, $P_{i,t}$ is the firm's price, as set in period t, $Y_{i,t+k|t}$ is the firm's quantity in period $t+k$, conditional on not having changed the price since t, and mc_{t+k} is the real marginal cost in period $t+k$. The firm's discount factor is δ, and θ is the Calvo parameter (probability of not resetting price).

Taking the first-order condition and log-linearizing around a steady state with no shocks and zero inflation, we get the following, familiar, characterization of the optimal reset price:

$$p^*_{i,t} = (1-\delta\theta)\sum_{k=0}^{\infty}(\delta\theta)^k \, \mathbb{E}_{i,t}\left[mc_{t+k} + p_{t+k}\right]. \tag{F.1}$$

[25]To see this, notice that

$$\|\alpha + \kappa(L)\| = \left\|\frac{\alpha + \gamma(L)}{1 - \beta(L)}\right\| \leq \|\alpha + \gamma(L)\|\left\|(1 - \beta(L))^{-1}\right\| \leq \||\alpha| + \gamma(L)\|\left\|1 - \beta(L)\right\|^{-1} \leq \frac{\alpha + \|\gamma(L)\|}{1 - \left\|\beta(L)\right\|} < 1,$$

where the first inequality follows from Cauchy-Schwarz, the second from the fact that, for any operator T, $\|T\|^{-1} \leq \left\|T^{-1}\right\|$, and the fourth from the assumptions that $\left\|\beta(L)\right\| < 1$, and $|\alpha| + \|\gamma(L)\| + \left\|\beta(L)\right\| < 1$.

Suppose that firms observe the aggregate prices up to period $t-1$, that is, they observe p^{t-1}, then we can restate condition (F.1) as

$$p_{i,t}^{*} - p_{t-1} = (1-\delta\theta)\sum_{k=0}^{\infty}(\delta\theta)^{k}\,\mathbb{E}_{i,t}\left[mc_{t+k}\right] + \sum_{k=0}^{\infty}(\delta\theta)^{k}\,\mathbb{E}_{i,t}\left[\pi_{t+k}\right]. \tag{F.2}$$

Since only a fraction $1-\theta$ of the firms adjust their prices each period, the price level in period t is given by $p_t = (1-\theta)\int p_{i,t}^{*}di + \theta p_{t-1}$, and inflation is given by

$$\pi_t \equiv p_t - p_{t-1} = (1-\theta)\int\left(p_{i,t}^{*} - p_{t-1}\right).$$

Define the firm specific inflation rate to be

$$\pi_{i,t} \equiv (1-\theta)\left(p_{i,t}^{*} - p_{t-1}\right).$$

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

$$\pi_{i,t} = (1-\theta)\,\mathbb{E}_{i,t}\left[(1-\delta\theta)\,mc_t + \pi_t\right] + \delta\theta\mathbb{E}_{i,t}\left[(1-\theta)\sum_{k=0}^{\infty}(\delta\theta)^{k}\,\mathbb{E}_{i,t+1}\left[(1-\delta\theta)\,mc_{t+1+k} + \pi_{t+1+k}\right]\right],$$

and we obtain the following beauty contest game, which includes equation (5.8),

$$\pi_{i,t} = (1-\theta)(1-\delta\theta)\,\mathbb{E}_{i,t}\left[mc_t\right] + (1-\theta)\,\mathbb{E}_{i,t}\left[\pi_t\right] + \delta\theta\mathbb{E}_{i,t}\left[\pi_{i,t+1}\right], \quad \text{with } \pi_t = \int \pi_{i,t}.$$

F.2 Households: The Dynamic IS

The consumer's problem is

$$\max_{\{C_{i,t},B_{i,t+1}\}} \mathbb{E}_{i,t}\left[\sum_{k=0}^{\infty}\beta^{k}\frac{C_{i,t+k}^{1-\frac{1}{\varsigma}}}{1-\frac{1}{\varsigma}}\right]$$

subject to

$$C_{i,t} + B_{i,t+1} = R_{t-1}B_{i,t} + Y_t,$$

where R_t and W denote real interest rates and income. At any state, the life-time budget constraint can be written as

$$\sum_{k=0}^{\infty}\frac{C_{i,t+k}}{\prod_{j=1}^{k}R_{t+j-1}} = R_{t-1}B_{i,t} + \sum_{k=0}^{\infty}\frac{Y_t}{\prod_{j=1}^{k}R_{t+j-1}},$$

which can be log-linearized into

$$\sum_{k=0}^{\infty}\beta^{k}c_{i,t+k} = \sum_{k=0}^{\infty}\beta^{k}y_{t+k}.$$

Combining this with the log-linearized version of the households' Euler equation,

$$c_{i,t} = \mathbb{E}_{i,t}[c_{i,t+1}] - \varsigma\mathbb{E}_{i,t}[r_t],$$

and using the market clearing condition, $y_t = c_t$, we obtain

$$c_{i,t} = -\varsigma\beta \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{i,t}[r_{t+k}] + (1-\beta) \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{i,t}[c_{t+k}].$$

Finally, notice this is implied by the following beauty-contest game,

$$c_{i,t} = -\varsigma\beta \mathbb{E}_{i,t}[r_t] + (1-\beta)\mathbb{E}_{i,t}[c_t] + \beta \mathbb{E}_{i,t}[c_{i,t+1}], \quad \text{with } c_t = \int c_{i,t}.$$

Letting $\texttt{mpc} = 1 - \beta$ yields equation (5.9).

F.3 Invertibility: Proof of Corollary 2

Suppose that ξ_t follows an AR(1) process

$$\xi_t = \rho \xi_{t-1} + \eta_t + \theta \eta_{t-1}, \quad \text{with } \eta_t \sim N(0,1),$$

and that, besides past prices, firms only observe a private signal about it with precision τ_u,

$$x_{i,t} = \xi_t + u_{i,t}, \quad \text{with } u_{i,t} \sim N\left(0, \tau_u^{-1}\right).$$

Then, it follows from Corollary 3 that for a best response

$$a_{i,t} = \varphi \mathbb{E}_{i,t}[\xi_t] + \alpha \mathbb{E}_{i,t}[a_t] + \gamma \mathbb{E}_{i,t}[a_{t+1}] + \beta \mathbb{E}_{i,t}[a_{i,t+1}]$$

the equilibrium is invertible if and only if

$$C(\alpha) \equiv \left| \frac{1-\alpha}{\rho\alpha(1-H)} \right| > 1, \quad \text{where} \quad H \equiv \frac{\tau_u}{1+\tau_u}.$$

For positive α and ρ, the relevant case here, it is easy to see that $C(\alpha)$ is a decreasing function. Hence, higher degrees of static strategic complementarity make it less likely that the equilibrium is invertible. Since this controlled by $1-\theta$ in the NKPC and $\texttt{mpc} \equiv 1-\beta$ in the Dynamic IS, invertibility is less likely for lower θ and higher $\texttt{mpc}$.

G. Proof of Proposition 5.3

Suppose the equilibrium is invertible and guess

$$\pi_t = \alpha_0 \xi_t^s + \alpha_1 \xi_t^d + \alpha_2 \eta_t^s + \alpha_3 \eta_t^d, \qquad \text{and} \qquad c_t = \beta_0 \xi_t^s + \beta_1 \xi_t^d + \beta_2 \eta_t^s + \beta_3 \eta_t^d.$$

From the best-response functions, it follows that,

$$
\begin{aligned}
\pi_t = & \left\{\kappa\left(1+\beta_0\right)+(1-\theta)\,\alpha_0+\delta\theta\alpha_0\rho\right\}\rho\xi^s_{t-1}+\left\{\kappa\beta_1+(1-\theta)\,\alpha_1+\delta\theta\alpha_1\rho\right\}\rho\xi^d_{t-1} \\
& +\left\{\kappa\beta_2+(1-\theta)\,\alpha_2+\kappa\left(1+\beta_0\right)+(1-\theta)\,\alpha_0+\delta\theta\alpha_0\rho\right\}E_t\left[\eta^s_t\right] \\
& +\left\{\kappa\beta_3+(1-\theta)\,\alpha_3+\kappa\beta_1+(1-\theta)\,\alpha_1+\delta\theta\alpha_1\rho\right\}E_t\left[\eta^d_t\right],
\end{aligned}
$$

and

$$
\begin{aligned}
c_t = & \left\{\mathtt{mpc}\beta_0-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_0-\rho\alpha_0\right)-\rho\beta_0\left(\mathtt{mpc}-1\right)\right\}\rho\xi^s_{t-1} \\
& +\left\{\mathtt{mpc}\beta_1-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_1-\rho\alpha_1+1\right)-\rho\beta_1\left(\mathtt{mpc}-1\right)\right\}\rho\xi^d_{t-1} \\
& +\left\{\mathtt{mpc}\beta_0-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_0-\rho\alpha_0\right)+\mathtt{mpc}\beta_2-\rho\beta_0\left(\mathtt{mpc}-1\right)-\varsigma(1-\mathtt{mpc})\phi\alpha_2\right\}E_t\left[\eta^s_t\right] \\
& +\left\{\mathtt{mpc}\beta_1-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_1-\rho\alpha_1+1\right)+\mathtt{mpc}\beta_3-\rho\beta_1\left(\mathtt{mpc}-1\right)-\varsigma(1-\mathtt{mpc})\phi\alpha_3\right\}E_t\left[\eta^d_t\right].
\end{aligned}
$$

Let $\lambda=\frac{\tau}{1+\tau}$, then

$$
E_t\left[\eta^s_t\right]=\lambda\eta^s_t,\;\; \text{and}\;\; E_t\left[\eta^d_t\right]=\lambda\eta^d_t.
$$

Matching coefficients, we get a system of equations that can be split into a system for α_0, β_0, α_1, and β_1,

$$
\begin{aligned}
\alpha_0 &= \left(\kappa\left(1+\beta_0\right)+(1-\theta)\,\alpha_0+\delta\theta\alpha_0\rho\right), \\
\beta_0 &= \left(\mathtt{mpc}\beta_0-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_0-\rho\alpha_0\right)-\rho\beta_0\left(\mathtt{mpc}-1\right)\right), \\
\alpha_1 &= \left(\kappa\beta_1+(1-\theta)\,\alpha_1+\delta\theta\alpha_1\rho\right), \\
\beta_1 &= \left(\mathtt{mpc}\beta_1-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_1-\rho\alpha_1+1\right)-\rho\beta_1\left(\mathtt{mpc}-1\right)\right),
\end{aligned}
$$

a system for α_2, and β_2,

$$
\begin{aligned}
\left(\alpha_0+\alpha_2\right) &= \left(\kappa\beta_2+(1-\theta)\,\alpha_2+\kappa\left(1+\beta_0\right)+(1-\theta)\,\alpha_0+\delta\theta\alpha_0\rho\right)\lambda, \\
\left(\beta_0+\beta_2\right) &= \left(\mathtt{mpc}\beta_0-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_0-\rho\alpha_0\right)+\mathtt{mpc}\beta_2-\rho\beta_0\left(\mathtt{mpc}-1\right)-\varsigma(1-\mathtt{mpc})\phi\alpha_2\right)\lambda,
\end{aligned}
$$

and a system for α_3, and β_3,

$$
\begin{aligned}
\left(\alpha_1+\alpha_3\right) &= \left(\kappa\beta_3+(1-\theta)\,\alpha_3+\kappa\beta_1+(1-\theta)\,\alpha_1+\delta\theta\alpha_1\rho\right)\lambda, \\
\left(\beta_1+\beta_3\right) &= \left(\mathtt{mpc}\beta_1-\varsigma(1-\mathtt{mpc})\left(\phi\alpha_1-\rho\alpha_1+1\right)+\mathtt{mpc}\beta_3-\rho\beta_1\left(\mathtt{mpc}-1\right)-\varsigma(1-\mathtt{mpc})\phi\alpha_3\right)\lambda.
\end{aligned}
$$

Solving the system yields

$$
\begin{aligned}
\alpha_0 &= \frac{\kappa\left(1-\mathtt{mpc}\right)\left(1-\rho\right)}{\left(1-\mathtt{mpc}\right)\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)}, \\
\alpha_1 &= -\frac{\kappa\varsigma(1-\mathtt{mpc})}{\left(1-\mathtt{mpc}\right)\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)}, \\
\beta_0 &= -\frac{\kappa\varsigma(1-\mathtt{mpc})\left(\phi-\rho\right)}{\left(1-\mathtt{mpc}\right)\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)}, \\
\beta_1 &= -\frac{\theta\varsigma(1-\mathtt{mpc})\left(1-\delta\rho\right)}{\left(1-\mathtt{mpc}\right)\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)},
\end{aligned}
$$

and

$$\alpha_2 = \frac{\kappa(1-\lambda)\left((\mathtt{mpc}-1)\left(\rho-1\right)(\lambda\mathtt{mpc}-1)+\kappa\varsigma(1-\mathtt{mpc})\lambda\phi-\kappa\varsigma(1-\mathtt{mpc})\lambda\rho\right)}{\left((1-\lambda\mathtt{mpc})(1-\lambda(1-\theta))+\kappa\varsigma(1-\mathtt{mpc})\phi\lambda^2\right)\left((1-\mathtt{mpc})\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)\right)},$$

$$\beta_2 = \frac{\kappa\varsigma(1-\mathtt{mpc})(1-\lambda)\left(\phi\left(1+\lambda\left(\theta-\mathtt{mpc}-\rho+\mathtt{mpc}\rho\right)\right)-\rho(1-\lambda(1-\theta))\right)}{\left((1-\lambda\mathtt{mpc})(1-\lambda(1-\theta))+\kappa\varsigma(1-\mathtt{mpc})\phi\lambda^2\right)\left((1-\mathtt{mpc})\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)\right)},$$

$$\alpha_3 = \frac{\kappa\varsigma(1-\mathtt{mpc})(1-\lambda)\left(1+\lambda\left(\theta-\mathtt{mpc}-\theta\delta\rho\right)\right)}{\left((1-\lambda\mathtt{mpc})(1-\lambda(1-\theta))+\kappa\varsigma(1-\mathtt{mpc})\phi\lambda^2\right)\left((1-\mathtt{mpc})\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)\right)},$$

$$\beta_3 = \frac{\varsigma(1-\mathtt{mpc})(1-\lambda)\left(\theta\left(1-\delta\rho\right)(1-\lambda(1-\theta))-\kappa\varsigma(1-\mathtt{mpc})\phi\lambda\right)}{\left((1-\lambda\mathtt{mpc})(1-\lambda(1-\theta))+\kappa\varsigma(1-\mathtt{mpc})\phi\lambda^2\right)\left((1-\mathtt{mpc})\left(1-\rho\right)\left(1-\delta\rho\right)\theta-\kappa\varsigma(1-\mathtt{mpc})\left(\rho-\phi\right)\right)}.$$

This solution validates the guess. Next, to guarantee that the equilibrium is indeed invertible, we need the following system to be invertible,

$$\begin{bmatrix} \pi_t \\ c_t \end{bmatrix} = \begin{bmatrix} \frac{\alpha_0}{1-\rho L}+\alpha_2 & \frac{\alpha_1}{1-\rho L}+\alpha_3 \\ \frac{\beta_0}{1-\rho L}+\beta_2 & \frac{\beta_1}{1-\rho L}+\beta_3 \end{bmatrix} \begin{bmatrix} \eta_t^s \\ \eta_t^d \end{bmatrix}.$$

For that, we need both roots of the determinant,

$$\Delta(L) = \left(\frac{\alpha_0}{1-\rho L}+\alpha_2\right)\left(\frac{\beta_1}{1-\rho L}+\beta_3\right)-\left(\frac{\alpha_1}{1-\rho L}+\alpha_3\right)\left(\frac{\beta_0}{1-\rho L}+\beta_2\right),$$

to be outside the unit circle. The roots are given by

$$r_1 \equiv -\frac{\tau(1+\theta-\mathtt{mpc})+\tau\sqrt{(1-\theta-\mathtt{mpc})^2-4\kappa\varsigma(1-\mathtt{mpc})\phi}}{2\rho},$$

$$r_2 \equiv -\frac{\tau(1+\theta-\mathtt{mpc})-\tau\sqrt{(1-\theta-\mathtt{mpc})^2-4\kappa\varsigma(1-\mathtt{mpc})\phi}}{2\rho}.$$

If $(1-\theta-\mathtt{mpc})^2-4\kappa\varsigma(1-\mathtt{mpc})\phi<0$, the roots are complex and their magnitude is above 1 if

$$\phi > \frac{\rho^2-\theta(1-\mathtt{mpc})\tau^2}{\kappa\varsigma(1-\mathtt{mpc})\tau^2}.$$

Otherwise, the roots are real and $|r_1|>|r_2|$. So we need to show that $|r_2|>1$, and for that we need to consider two cases: If $\tau(1+\theta-\mathtt{mpc})<2\rho$, then $|r_2|$ is always less than 1, otherwise, it is a necessary and sufficient condition that

$$\phi > \frac{\tau\rho(1+\theta-\mathtt{mpc})-\rho^2-\theta(1-\mathtt{mpc})\tau^2}{\kappa\varsigma(1-\mathtt{mpc})\tau^2}.$$

H. Forward Complementarity Example

Consider the following simple version of equation (5.7),

$$a_{it} = \varphi\mathbb{E}_{it}[\xi_{it}]+\alpha\mathbb{E}_{it}[a_t]+\gamma\mathbb{E}_{it}[a_{t+1}]+\beta\mathbb{E}_{it}[a_{it+1}],$$

and suppose that the equilibrium is invertible. Then, Proposition 5.2 implies that this is equivalent to the static best response

$$a_{it} = \varphi \mathbb{E}_{it}[\widetilde{\xi}_{it}] + \alpha \mathbb{E}_{it}[a_t], \quad \text{with} \quad \widetilde{\xi}_{it} \equiv \frac{1-\alpha}{1-\alpha-(\gamma+\beta)L^{-1}}\,\xi_{it}.$$

As noted above, we can see that forward aggregate and individual complementarities, controlled by γ and β have interchangeable effects. This property allows us to simplify the analysis since we do not need to distinguish between the two types of forward complementarities focusing only on their sum.

Next, to be more concrete, suppose that ξ_{it} does not depend on idiosyncratic shocks (so we suppress the notation i) and follows an ARMA(1,1) process,

$$\xi_t = \rho \xi_{t-1} + \eta_t + \theta \eta_{t-1}, \quad \eta_t \sim \mathcal{N}(0,1),$$

and that, every period, agent i observes last period's aggregate action, a_{t-1}, and a private signal,

$$x_{it} = \xi_t + u_{it}, \quad u_{it} \sim \mathcal{N}(0, \tau_u^{-1}).$$

Then, we can establish the following corollary to Proposition 5.2.

Corollary 3. *The equilibrium is invertible if and only if*

$$\left| \frac{1-\alpha+\theta(\gamma+\beta)}{\rho(1-\alpha+\theta(\gamma+\beta))-(\rho+\theta)(1-\alpha H)} \right| > 1, \quad \textit{where} \quad H \equiv \frac{\tau_u}{1+\tau_u}. \tag{H.1}$$

Proof. It follows from Proposition 5.2 that, for any $d(L)$ such that $\xi_t = d(L)\eta_t$, the law of motion for the aggregate action satisfies,

$$g(L) = \varphi \left[\frac{(1-\alpha)L}{(1-\alpha)L-(\gamma+\beta)} d(L) \right]_+ + \alpha(g(L)-g_0) + \alpha g_0 H.$$

where $[\cdot]_+$ denotes the annihilator operator, and $H \equiv \tau_u/(1+\tau_u)$. Let $\kappa \equiv (\gamma+\beta)/(1-\alpha)$, replace the lag operator in this equation with an arbitrary complex number z, and rearrange to get

$$(1-\alpha)g(z) = \varphi \frac{zd(z)-\kappa d(\kappa)}{z-\kappa} + \alpha g_0(H-1).$$

We can obtain $g_0 = \varphi d(\kappa)(1-\alpha H)^{-1}$ by evaluating this equation at $z=0$, and, then,

$$g(z) = g_0 + \frac{\varphi}{1-\alpha}\frac{d(z)-d(\kappa)}{z-\kappa} z.$$

Using the fact that $d(z) = (1+\theta z)/(1-\rho z)$, we obtain

$$g(z) = \frac{\varphi}{1-\rho\kappa}\left(\frac{1+\theta\kappa}{1-\alpha H} + \frac{1}{1-\alpha}\frac{(\rho+\theta)z}{1-\rho z} \right),$$

which has root

$$z^* = \frac{(1-\alpha)(1+\theta\kappa)}{\rho(1-\alpha)(1+\theta\kappa)-(\rho+\theta)(1-\alpha H)}.$$

The equilibrium is invertible if and only if this root to be outside the unit circle, that is $|z^*| > 1$. $\square$

It is easy to see from equation (H.1) that the effect of forward complementarities, $\gamma+\beta$, on the invertibility of the equilibrium is ambiguous and depends both on the autoregressive and moving-average parameters, ρ and θ. To interpret this condition, it is useful to consider some particular cases. If, for instance, the fundamental follows an AR(1) process, with $\theta = 0$, the inequality simplifies to

$$\left| \frac{1-\alpha}{\alpha\rho(1-H)} \right| > 1,$$

and forward complementarities actually do not matter for invertibility.

To understand this, first let $[\cdot]_+$ denote the annihilator operator which sets negative powers of the lag operator to zero. Then, because the expected value of future shocks is always zero, we have that, for any stochastic variable $y_{i,t+j}$ and any j, $\mathbb{E}_{it}[y_{i,t+j}] = \mathbb{E}_{it}[[y_{i,t+j}]_+]$. When $\theta = 0$, we have that

$$[\widetilde{\xi}_t]_+ = \left[\frac{1-\alpha}{1-\alpha-(\gamma+\beta)L^{-1}} \frac{1}{1-\rho L} \eta_t \right]_+ = \frac{1-\alpha}{1-\alpha-\rho(\gamma+\beta)} \frac{1}{1-\rho L} \eta_t.$$

Thus, a change in $\gamma + \beta$ affects only the variance of the fundamental but not the autoregressive coefficient. Loosely speaking, when $\beta + \gamma$ increases, the agent puts relatively more weight on the next period fundamental. But since $\mathbb{E}_{it}[\xi_{t+1}] = \rho\mathbb{E}_{it}[\xi_t]$ this amounts to a proportional increase in the aggregate action in every period which does not affect invertibility.

Things are different if ξ_t follows an MA(1) process, that is, when $\rho = 0$ and $\theta \neq 0$. In that case, the inequality simplifies to

$$\left| \frac{1-\alpha+\theta(\gamma+\beta)}{\theta(1-\alpha H)} \right| > 1,$$

so that, if $\theta > 0$ (< 0) the equilibrium is non-invertible when $\gamma + \beta$ is low (high) enough.[26] In this case, we have that

$$[\widetilde{\xi}_t]_+ = \left[\frac{1-\alpha}{1-\alpha-(\gamma+\beta)L^{-1}} (1+\theta L)\eta_t \right]_+ = \left(1 - \frac{\theta(\gamma+\beta)}{1-\alpha} + \theta L \right) \eta_t.$$

Here, it is useful to consider the response of the aggregate action, a_t, to a shock to the fundamental, η_t, in period $t = 0$, assuming the equilibrium is invertible. Also, for simplicity, suppose that $\theta > 0$. Then, an increase in $\gamma + \beta$ decreases the response of a_0 by an amount proportional to θ. On the other hand, it leaves a_1 unchanged since the impulse response of ξ_{t+k} for $k \geq 2$ is zero, so that forward-looking complementarities do not affect a_1 or the action in any further period. It follows that a_0/a_1 decreases, which reduces the signal-to-noise ratio and, therefore, the informativeness of the observation of a_0 to forecast a_1. This, in turn, makes it less likely that the equilibrium is indeed invertible.

Figure 9 shows how when the sign of the moving average parameter, θ, flips the effect of an increase in the degree of forward-looking complementarities, $\beta + \gamma$, on invertibility. It also illustrates, in accordance with Theorem 1 and Proposition 5.2, that it is always the case that for a high enough degree of *static* complementarity, α, the equilibrium is non-invertible.

[26]More specifically, for $\theta > 0$, the equilibrium is non-invertible if $\gamma + \beta < (1-\alpha H) - (1-\alpha)/\theta$, and, for $\theta < 0$, if $\gamma + \beta > -(1-\alpha H) - (1-\alpha)/\theta$.

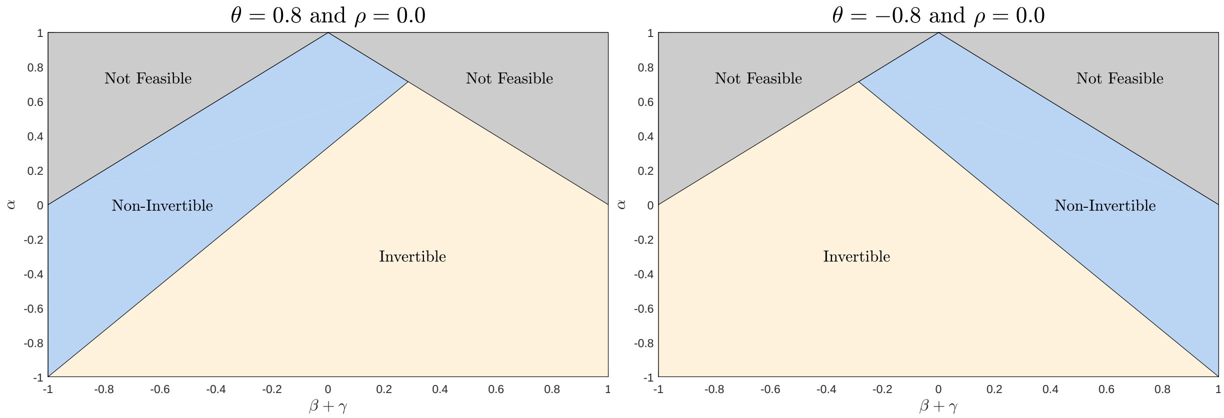


Figure 9: Regions of Invertibility with Forward-Looking Complementarities

The only free parameter, τ_u, is set to 1.

I. Backward and Forward Complementarities

Section 5.2 considers a best response function

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_{it}] + \alpha \mathbb{E}_{it}[a_t] + \delta(L)\mathbb{E}_{it}[a_t] + \lambda(L)\mathbb{E}_{it}[a_{it}], \tag{I.1}$$

with forward looking complementarities, that is, assuming that $\delta(L)$ and $\lambda(L)$ are functions only of negative powers of the lag operator L. This section handles the cases in which backward complementarities as well. First, Section I.1 discusses the case with only static and backward complementarities. Then, Section I.2 deals with the case in which there are both backward and forward complementarities.

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_{it}] + \alpha \mathbb{E}_{it}[a_t] + \delta(L)\mathbb{E}_{it}[a_t] + \lambda(L)\mathbb{E}_{it}[a_{it}],$$

I.1 Backward Complementarities

Consider the best response in equation I.1 with only backward complementarities, that is, such that $\delta(L)$ and $\lambda(L)$ only have positive powers of the lag polynomial. Since past aggregate actions have been assumed to be in agent's information sets, it immediately follows that

$$\mathbb{E}_{it}[a_{t-k}] = a_{t-k}, \qquad \text{for all } k \geq 1. \tag{I.2}$$

Assume that the perfect information equilibrium,

$$a_t = \frac{\varphi}{1 - \alpha - \delta(L) - \lambda(L)} \xi_{it}$$

is well defined, that is, that $\|\alpha + \delta(L) + \lambda(L)\| < 1$ in the operator norm.

Proposition I.1. *The equilibrium is invertible with the best response*

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_{it}] + \alpha \mathbb{E}_{it}[a_t],$$

if and only if it is invertible with the best response in equation I.1 with

$$\delta(L) = \sum_{k=1}^{\infty} \delta_k L^k, \quad \textit{and} \quad \lambda(L) = \sum_{k=1}^{\infty} \lambda_k L^k.$$

Proof. It follows from equation (I.2) that the best response with backward complementarities can be rewritten as

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_{it}] + \alpha \mathbb{E}_{it}[a_t] + \delta(L) a_t + \lambda(L) a_{it}.$$

Analogously to the steps in the proof of Theorem 1, we obtain the following consistency requirement for the law of motion of the aggregate action,

$$\boldsymbol{g}(L) = \varphi\left(\boldsymbol{d}(L) - \boldsymbol{d}_0\right) + \alpha\left(\boldsymbol{g}(L) - \boldsymbol{g}_0\right) + \left(\varphi \boldsymbol{d}_0 + \alpha \boldsymbol{g}_0\right) \boldsymbol{k}_1 \boldsymbol{\Omega} + \varphi \boldsymbol{e}_0 \boldsymbol{k}_2 \boldsymbol{\Omega} + (\delta(L) + \lambda(L)) \boldsymbol{g}(L),$$

which can be rewritten as

$$(1 - \alpha - \delta(L) - \lambda(L)) \boldsymbol{g}(L) = \varphi\, \boldsymbol{d}(L) + \left(\varphi \boldsymbol{d}_0 + \alpha \boldsymbol{g}_0\right)(\boldsymbol{k}_1 \boldsymbol{\Omega} - \mathbf{I}) + \varphi \boldsymbol{e}_0 \boldsymbol{k}_2 \boldsymbol{\Omega}$$

Since $\delta_0 = \lambda_0 = 0$, we have that, just as in the proof of Theorem 1,

$$\boldsymbol{g}_0 = \varphi\left(\boldsymbol{d}_0 \boldsymbol{k}_1 + \boldsymbol{e}_0 \boldsymbol{k}_2\right) \boldsymbol{\Omega} \left(\mathbf{I} - \alpha \boldsymbol{k}_1 \boldsymbol{\Omega}\right)^{-1}.$$

and it follows that

$$\boldsymbol{g}(L) = \frac{(1-\alpha)\boldsymbol{g}_0 + \varphi(\boldsymbol{d}(L) - \boldsymbol{d}_0)}{1 - \alpha - \delta(L) - \lambda(L)}.$$

Let

$$\boldsymbol{g}(z) = \frac{1-\alpha}{\varphi} \boldsymbol{g}_0 + \boldsymbol{d}(z) - \boldsymbol{d}_0,$$

so that

$$\boldsymbol{g}(z) = \frac{\varphi}{1 - \alpha - \delta(z) - \lambda(z)} \boldsymbol{g}(z),$$

and notice that $\boldsymbol{g}(z)$ is the same as in the proof of Theorem 1, so that, if the equilibrium is invertible in the static best response, it remains invertible with any feasible $\delta(L)$ and $\lambda(L)$ since, by assumption, $\|\alpha + \delta(L) + \lambda(L)\| < 1$. If it is non-invertible it remains non-invertible for the same reason. □

It follows that the result in Theorem 1 immediately generalizes to settings with backward-looking complementarities. So that, regardless of these complementarities, if the static complementarity, α, is large enough the equilibrium is not invertible.

I.2 Interacting Backward and Forward Complementarities

Next, consider the following best-response function which encompasses most environments considered in the literature including, for instance, the Euler equation in a New-Keynesian model with capital,

$$a_{it} = \varphi \mathbb{E}_{it}[\xi_t] + \alpha \mathbb{E}_{it}[a_t] + \gamma \mathbb{E}_{it}[a_{t+1}] + \beta \mathbb{E}_{it}[a_{it+1}] + \delta \mathbb{E}_{it}[a_{t-1}] + \lambda \mathbb{E}_{it}[a_{it-1}].$$

Perfect Information Benchmark. It is easy to see that, if agents observe every shock up to the current period perfectly, the equilibrium must satisfy the following consistency requirement

$$\boldsymbol{g}(L) = \varphi \boldsymbol{d}(L) + \alpha \boldsymbol{g}(L) + (\gamma + \beta)\left(\frac{\boldsymbol{g}(L) - \boldsymbol{g}_0}{L}\right) + (\delta + \lambda)\boldsymbol{g}(L)L,$$

which, replacing the lag operator with an arbitrary complex number z, can be rewritten as

$$\left[-(\delta + \lambda)z^2 + (1 - \alpha)z - (\gamma + \beta)\right] \boldsymbol{g}(z) = \varphi \boldsymbol{d}(z)z - (\gamma + \beta)\boldsymbol{g}_0.$$

In order for this equilibrium to exist and be unique, the polynomial on the left-hand side of this equation must have exactly one inside root, an assumption that we maintain throughout. By inside root we mean that the root is inside the unit circle in the complex plane. The right-hand side of the equation must be zero at any inside root of the polynomial at the left-hand side to avoid poles inside the unit circle. This condition is used to determine $\boldsymbol{g}_0$. With two outside roots, $\boldsymbol{g}_0$ is indeterminate so that there are multiple equilibria, and with two inside roots, $\boldsymbol{g}_0$ is over-determined so that, in general, an equilibrium does not exist. Let κ_1 and κ_2 be the inside and outside roots respectively, then, we would have that, the unique perfect-information equilibrium satisfies

$$\boldsymbol{g}(L) = \frac{\varphi}{\delta + \lambda} \frac{\kappa_2^{-1}}{1 - \kappa_2^{-1}L} \frac{\boldsymbol{d}(L)L - \boldsymbol{d}(\kappa_1)\kappa_1}{L - \kappa_1}.$$

In what follows we only consider the set of parameters in which this perfect-information equilibrium exists and is unique, that is, such that $|\kappa_1| < 1$ and $|\kappa_2| > 1$. We refer to parameters that do not satisfies these conditions as infeasible. In this setup, we can establish the following result.

Theorem 2. *Suppose that $\boldsymbol{e}(L) = \boldsymbol{e}$ and $\boldsymbol{B}(L) = \boldsymbol{B}$. For any $\omega_1, \omega_2 \neq 0$ such that $|\omega_1| < 1$ and $|\omega_2| < 1$, there exists $\varepsilon > 0$ low enough such that if $\alpha = 1 - (1 + \omega_1\omega_2)\varepsilon$, $\beta + \gamma = \omega_1\varepsilon$, and $\delta + \lambda = \omega_2\varepsilon$, the equilibrium is not invertible.*

Proof. The proof is presented in the next Section I.3. □

This theorem extends the result in Theorem 1 for the case in which there are both forward and backward complementarities under the restriction that $\boldsymbol{e}(L) = \boldsymbol{e}$ and $\boldsymbol{B}(L) = \boldsymbol{B}$. This restriction is not particularly relevant since most environments considered in the literature do satisfy it. The theorem implies that there always a region in the space of feasible complementarity-parameters $(\alpha, \gamma, \beta, \delta, \lambda)$ such that the equilibrium is non-invertible. The reason why it is not enough to take the limit as the static degree of complementarity, α, increases to 1, as in Theorem 1, is because, depending on the starting point, that might lead into an infeasible set of parameters. Hence, the limit must be taken in a careful enough way to guarantee that the region of non-invertibility is reached without violating feasibility.

I.3 Proof of Theorem 2

This proof follows very similar steps to the ones in the proof of Theorem 1, for clarity we closely follow the argument of that proof. Suppose that when $\alpha = 0$ the equilibrium is invertible, otherwise the result is trivial. Section I.3.1 characterizes the equilibrium assuming invertibility. Using this characterization, Section I.3.2 takes the appropriate limit and shows that, in it, the equilibrium cannot be invertible.

I.3.1 Solution Assuming Invertibility

Suppose that the equilibrium is invertible, then the information set of agent i in period t is given by $\mathcal{I}_{it} \equiv \{\boldsymbol{\eta}_{\tau-1}, \boldsymbol{u}_{i\tau-1}, \boldsymbol{x}_{i\tau}, \boldsymbol{z}_{\tau}\}_{\tau=-\infty}^{t}$. We guess (and verify below) that the individual policy function takes the form $a_{it} = \boldsymbol{g}(L)\boldsymbol{eta}_t + \boldsymbol{h}\boldsymbol{u}_{it}$. Therefore,

$$\mathbb{E}_{it}[\xi_{it}] = \mathbb{E}[\boldsymbol{d}(L)\boldsymbol{\eta}_t + \boldsymbol{e}\boldsymbol{u}_{it}|\mathcal{I}_{it}] = (\boldsymbol{d}(L) - \boldsymbol{d}_0)\boldsymbol{\eta}_t + \boldsymbol{d}_0\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] + \boldsymbol{e}\,\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}],$$

$$\mathbb{E}_{it}[a_t] = \mathbb{E}[\boldsymbol{g}(L)\boldsymbol{\eta}_t|\mathcal{I}_{it}] = (\boldsymbol{g}(L) - \boldsymbol{g}_0)\boldsymbol{\eta}_t + \boldsymbol{g}_0\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}],$$

$$\mathbb{E}_{it}[a_{it+1}] = \mathbb{E}[\boldsymbol{g}(L)\boldsymbol{\eta}_{t+1} + \boldsymbol{h}\boldsymbol{u}_{it+1}|\mathcal{I}_{it}] = \left(\frac{\boldsymbol{g}(L) - (\boldsymbol{g}_0 + \boldsymbol{g}_1 L)}{L}\right)\boldsymbol{\eta}_t + \boldsymbol{g}_1\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}],$$

$$\mathbb{E}_{it}[a_{t+1}] = \mathbb{E}[\boldsymbol{g}(L)\boldsymbol{\eta}_{t+1}|\mathcal{I}_{it}] = \left(\frac{\boldsymbol{g}(L) - (\boldsymbol{g}_0 + \boldsymbol{g}_1 L)}{L}\right)\boldsymbol{\eta}_t + \boldsymbol{g}_1\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}],$$

$$\mathbb{E}_{it}[a_{it-1}] = \mathbb{E}[\boldsymbol{g}(L)\boldsymbol{\eta}_{t-1} + \boldsymbol{h}\boldsymbol{u}_{it-1}|\mathcal{I}_{it}] = \boldsymbol{g}(L)\boldsymbol{\eta}_{t-1} + \boldsymbol{h}\boldsymbol{u}_{it-1},$$

$$\mathbb{E}_{it}[a_{t-1}] = \mathbb{E}[\boldsymbol{g}(L)\boldsymbol{\eta}_{t-1}|\mathcal{I}_{it}] = \boldsymbol{g}(L)\boldsymbol{\eta}_{t-1}.$$

Moreover, since

$$\boldsymbol{x}_{it} - (\mathbf{A}(L) - \mathbf{A}_0)\boldsymbol{\eta}_t = \mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}\boldsymbol{u}_{it}, \quad \text{and} \quad \boldsymbol{z}_t - (\mathbf{C}(L) - \mathbf{C}_0)\boldsymbol{\eta}_t = \mathbf{C}_0\boldsymbol{\eta}_t,$$

it follows that $\boldsymbol{x}_{it} - (\mathbf{A}(L) - \mathbf{A}_0)\boldsymbol{\eta}_t$ is a noisy signal about $\boldsymbol{\eta}_t$ and $\boldsymbol{u}_{it}$, and $\boldsymbol{z}_t - (\mathbf{C}(L) - \mathbf{C}_0)\boldsymbol{\eta}_t$ is a noisy signal about $\boldsymbol{\eta}_t$, which allows us to calculate

$$\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] = \begin{bmatrix}\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}\boldsymbol{\Sigma}_u^2\mathbf{B}^\top & \mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top \\ \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}\boldsymbol{u}_{it} \\ \mathbf{C}_0\boldsymbol{\eta}_t\end{bmatrix},$$

$$\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}] = \begin{bmatrix}\boldsymbol{\Sigma}_u^2\mathbf{B}^\top & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}\boldsymbol{\Sigma}_u^2\mathbf{B}^\top & \mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top \\ \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}_0\boldsymbol{\eta}_t + \mathbf{B}\boldsymbol{u}_{it} \\ \mathbf{C}_0\boldsymbol{\eta}_t\end{bmatrix}.$$

Therefore,

$$\begin{aligned}a_{it} &= \varphi\left((\boldsymbol{d}(L) - \boldsymbol{d}_0)\boldsymbol{\eta}_t + \boldsymbol{d}_0\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] + \boldsymbol{e}\,\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}]\right) + \alpha\left((\boldsymbol{g}(L) - \boldsymbol{g}_0)\boldsymbol{\eta}_t + \boldsymbol{g}_0\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}]\right) \\ &\quad + \gamma\left(\left(\frac{\boldsymbol{g}(L) - (\boldsymbol{g}_0 + \boldsymbol{g}_1 L)}{L}\right)\boldsymbol{\eta}_t + \boldsymbol{g}_1\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}]\right) + \beta\left(\left(\frac{\boldsymbol{g}(L) - (\boldsymbol{g}_0 + \boldsymbol{g}_1 L)}{L}\right)\boldsymbol{\eta}_t + \boldsymbol{g}_1\,\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}]\right) \\ &\quad + (\delta + \lambda)\boldsymbol{g}(L)L\boldsymbol{\eta}_t + \lambda\boldsymbol{h}L\boldsymbol{u}_{it},\end{aligned}$$

which can be reorganized as

$$\begin{aligned}a_{it} &= \left[\varphi\left(\boldsymbol{d}(L) - \boldsymbol{d}_0\right) + \alpha\left(\boldsymbol{g}(L) - \boldsymbol{g}_0\right) + \left(\gamma + \beta\right)\left(\frac{\boldsymbol{g}(L) - (\boldsymbol{g}_0 + \boldsymbol{g}_1 L)}{L}\right) + (\delta + \lambda)\boldsymbol{g}(L)L\right]\boldsymbol{\eta}_t \\ &\quad + \lambda\boldsymbol{h}L\boldsymbol{u}_{it} + \left(\varphi\boldsymbol{d}_0 + \alpha\boldsymbol{g}_0 + \left(\gamma + \beta\right)\boldsymbol{g}_1\right)\mathbb{E}[\boldsymbol{\eta}_t|\mathcal{I}_{it}] + \varphi\boldsymbol{e}\mathbb{E}[\boldsymbol{u}_{it}|\mathcal{I}_{it}].\end{aligned}$$

Consistency requires, in particular, that

$$\begin{aligned}\boldsymbol{g}(L) = \varphi\left(\boldsymbol{d}(L) - \boldsymbol{d}_0\right) + \alpha\left(\boldsymbol{g}(L) - \boldsymbol{g}_0\right) + \left(\gamma + \beta\right)\left(\frac{\boldsymbol{g}(L) - (\boldsymbol{g}_0 + \boldsymbol{g}_1 L)}{L}\right) + (\delta + \lambda)\boldsymbol{g}(L)L \\ + \left(\varphi \boldsymbol{d}_0 + \alpha \boldsymbol{g}_0 + \left(\gamma + \beta\right)\boldsymbol{g}_1\right)\boldsymbol{k}_1\boldsymbol{\Omega} + \varphi \boldsymbol{e}\boldsymbol{k}_2\boldsymbol{\Omega},\end{aligned}$$

where

$$\boldsymbol{k}_1 \equiv \begin{bmatrix}\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}, \quad \boldsymbol{k}_2 \equiv \begin{bmatrix}\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{0}\end{bmatrix}, \quad \text{and} \quad \boldsymbol{\Omega} \equiv \begin{bmatrix}\mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}_0\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top & \mathbf{A}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top \\ \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top & \mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{A}_0 \\ \mathbf{C}_0\end{bmatrix}.$$

The guess for the policy function can be verified by collecting the terms associated with the idiosyncratic shocks and noticing that they are zero for any period other than the current one. Thus, we can rewrite the equation above as

$$\begin{aligned}(1-\alpha)\left(L - \frac{\gamma+\beta}{1-\alpha} - \frac{\delta+\lambda}{1-\alpha}L^2\right)\boldsymbol{g}(L) = \varphi\,\boldsymbol{d}(L)L - \left(\gamma+\beta\right)\boldsymbol{g}_0 \\ + \left(\varphi \boldsymbol{d}_0 + \alpha \boldsymbol{g}_0 + \left(\gamma+\beta\right)\boldsymbol{g}_1\right)\left(\boldsymbol{k}_1\boldsymbol{\Omega} - \mathbf{I}\right)L + \varphi \boldsymbol{e}\boldsymbol{k}_2\boldsymbol{\Omega}L \qquad \text{(I.3)}\end{aligned}$$

It is useful to replace the lag operator in this equation with an arbitrary complex number z. Evaluating this equation at different values of z implies conditions that allow for the characterization of the equilibrium.

Solving for $\boldsymbol{g}(z)$. The right-hand side of equation (I.3) must be equal to $\mathbf{0}$ when evaluated at the inside root, κ_1, of the second-order polynomial on the left-hand side of the equation; we denote the outside root by κ_2.[27] Moreover, the equation must be consistent with the values of $\boldsymbol{g}_0$ and $\boldsymbol{g}_1$. Consistency at $z = 0$, i.e. $\boldsymbol{g}(0) = \boldsymbol{g}_0$ is automatic. Next, set $z = \kappa_1$ to get

$$(1-\alpha)\,\boldsymbol{g}_0 = \varphi \boldsymbol{d}(\kappa_1) + \left(\varphi \boldsymbol{d}_0 + \alpha \boldsymbol{g}_0 + \left(\gamma+\beta\right)\boldsymbol{g}_1\right)\left(\boldsymbol{k}_1\boldsymbol{\Omega} - \mathbf{I}\right) + \varphi \boldsymbol{e}\boldsymbol{k}_2\boldsymbol{\Omega}.$$

It follows that

$$\frac{\boldsymbol{g}(z) - \boldsymbol{g}_0}{z} = -\frac{\varphi(\boldsymbol{d}(z) - \boldsymbol{d}(\kappa_1))}{(\delta+\gamma)(z-\kappa_1)(z-\kappa_2)} - \frac{\boldsymbol{g}_0}{z-\kappa_2},$$

and since, by definition,

$$\boldsymbol{g}_1 = \left.\frac{\boldsymbol{g}(z) - \boldsymbol{g}_0}{z}\right|_{z=0},$$

we obtain

$$\boldsymbol{g}_1 = \frac{\boldsymbol{g}_0}{\kappa_2} + \frac{\varphi\left(\boldsymbol{d}(\kappa_1) - \boldsymbol{d}_0\right)}{\gamma+\beta},$$

Putting these results together we obtain that

$$(\kappa_2 - z)\boldsymbol{g}(z) = \kappa_2 \boldsymbol{g}_0 + \frac{\varphi}{\gamma+\beta}\frac{\boldsymbol{d}(z) - \boldsymbol{d}(\kappa_1)}{z-\kappa_1}z,$$

[27] Explicitly,

$$\kappa_1 = \frac{1-\alpha-\sqrt{(1-\alpha)^2 - 4(\gamma+\beta)(\delta+\lambda)}}{2(\delta+\lambda)}, \quad \text{and} \quad \kappa_2 = \frac{1-\alpha+\sqrt{(1-\alpha)^2 - 4(\gamma+\beta)(\delta+\lambda)}}{2(\delta+\lambda)}.$$

where

$$\boldsymbol{g}_0 = \varphi\left(\boldsymbol{d}(\kappa_1)\boldsymbol{k}_1 + \boldsymbol{e}\boldsymbol{k}_2\right)\boldsymbol{\Omega}\left(\mathbf{I} - (\alpha + (\delta+\lambda)\kappa_1)\boldsymbol{k}_1\boldsymbol{\Omega}\right)^{-1}.$$

Notice that, using the block-matrix inversion formula, $\boldsymbol{g}_0$ can be rewritten as

$$\begin{aligned}\boldsymbol{g}_0 = {} & \frac{\varphi}{1-(\alpha+(\delta+\lambda)\kappa_1)}\boldsymbol{d}(\kappa_1)\mathbf{E} \\ & + \varphi\left(\boldsymbol{d}(\kappa_1)\left(\mathbf{I}-\mathbf{E}\right)\mathbf{D} + \boldsymbol{e}\,\mathbf{F}\right)\sum_{j=0}^{\infty}(\alpha+(\delta+\lambda)\kappa_1)^j\left[\left(\mathbf{I}-\mathbf{E}\right)\mathbf{D}\left(\mathbf{I}-\mathbf{E}\right)\right]^j,\end{aligned} \tag{I.4}$$

where

$$\begin{aligned}\mathbf{D} &\equiv \boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top\left(\mathbf{A}_0\left(\mathbf{I}-\mathbf{E}\right)\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}_0\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top\right)^{-1}\mathbf{A}_0,\\ \mathbf{E} &\equiv \boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\left(\mathbf{C}_0\boldsymbol{\Sigma}_\eta^2\mathbf{C}_0^\top\right)^{-1}\mathbf{C}_0,\\ \mathbf{F} &\equiv \boldsymbol{\Sigma}_u^2\mathbf{B}^\top\left(\mathbf{A}_0\left(\mathbf{I}-\mathbf{E}\right)\boldsymbol{\Sigma}_\eta^2\mathbf{A}_0^\top + \mathbf{B}_0\boldsymbol{\Sigma}_u^2\mathbf{B}_0^\top\right)^{-1}\mathbf{A}_0,\end{aligned}$$

and we have used the fact that $\mathbf{E}$ is idempotent.

I.3.2 Taking the Appropriate Limit

Suppose that $\alpha = 1-(1+\omega_1\omega_2)\varepsilon$, $\gamma+\beta = \omega_1\varepsilon$, and $\delta+\lambda = \omega_2\varepsilon$. It follows that, for all $\varepsilon > 0$,

$$\kappa_1 = \omega_1, \quad \kappa_2 = \omega_2^{-1}, \quad \text{and} \quad \alpha + (\delta+\lambda)\kappa_1 = 1-\varepsilon.$$

To establish the claim, we consider the limit of

$$\boldsymbol{g}(z) = \frac{\varphi}{(1-\omega_2 z)\varepsilon}\left(\frac{\omega_2}{\omega_1}\frac{\boldsymbol{d}(z)-\boldsymbol{d}(\omega_1)}{z-\omega_1}z + \frac{\varepsilon}{\varphi}\boldsymbol{g}_0\right),$$

as ε decreases towards 0. Let

$$\boldsymbol{g}(z) = \frac{\omega_2}{\omega_1}\frac{\boldsymbol{d}(z)-\boldsymbol{d}(\omega_1)}{z-\omega_1}z + \frac{\varepsilon}{\varphi}\boldsymbol{g}_0,$$

so that

$$\boldsymbol{g}(z) = \frac{\varphi}{(1-\omega_2 z)\varepsilon}\boldsymbol{g}(z),$$

and notice that this is well defined for all $\varepsilon > 0$ and that, if $\lim_{\varepsilon\to 0^+}\det\left(\begin{bmatrix}\mathbf{C}(z) & \boldsymbol{g}(z)\end{bmatrix}^\top\right)$ has an inside root, then there exists ε low enough such that $\det\left(\begin{bmatrix}\mathbf{C}(z) & \boldsymbol{g}(z)\end{bmatrix}^\top\right)$ is well defined and has an inside root. Recall that $|\omega_2| < 1$. It is, in fact, sufficient to show that

$$\lim_{\varepsilon\to 0^+}\det\left(\begin{bmatrix}\mathbf{C}(0) & \boldsymbol{g}_0\end{bmatrix}^\top\right) = 0.$$

Accordingly, using equation (I.4) we have that

$$\lim_{\varepsilon \to 0^+} \boldsymbol{g}_0 = \lim_{\varepsilon \to 0^+} \frac{\varepsilon}{\varphi} \boldsymbol{g}_0 = \boldsymbol{d}(\omega_1)\mathbf{E} + \lim_{\varepsilon \to 0^+} \varepsilon \left(\boldsymbol{d}(\omega_1) \left(\mathbf{I} - \mathbf{E}\right) \mathbf{D} + \boldsymbol{e}\mathbf{F}\right) \sum_{j=0}^{\infty} (1 - \varepsilon)^j \left[(\mathbf{I} - \mathbf{E}) \mathbf{D} \left(\mathbf{I} - \mathbf{E}\right)\right]^j .$$

It follows from Lemma 6 that

$$\sum_{j=0}^{\infty} (1 - \varepsilon)^j \left[(\mathbf{I} - \mathbf{E}) \mathbf{D} \left(\mathbf{I} - \mathbf{E}\right)\right]^j$$

is well defined and finite, so that

$$\lim_{\varepsilon \to 0^+} \varepsilon \left(\boldsymbol{d}(\omega_1) \left(\mathbf{I} - \mathbf{E}\right) \mathbf{D} + \boldsymbol{e}\mathbf{F}\right) \sum_{j=0}^{\infty} (1 - \varepsilon)^j \left[(\mathbf{I} - \mathbf{E}) \mathbf{D} \left(\mathbf{I} - \mathbf{E}\right)\right]^j = \mathbf{0}.$$

Therefore,

$$\lim_{\varepsilon \to 0^+} \boldsymbol{g}_0 = \boldsymbol{d}(\omega_1)\mathbf{E}$$

and, using the definition of $\mathbf{E}$,

$$\boldsymbol{g}_0 = \boldsymbol{d}_0 \boldsymbol{\Sigma}_\eta^2 \mathbf{C}_0^\top \left(\mathbf{C}_0 \boldsymbol{\Sigma}_\eta^2 \mathbf{C}_0^\top\right)^{-1} \mathbf{C}_0 = \boldsymbol{a}\mathbf{C}_0$$

for some vector $\boldsymbol{a}$. Finally, notice that

$$\det \begin{bmatrix} \mathbf{C}_0 \\ \boldsymbol{g}_0 \end{bmatrix} = \det \begin{bmatrix} \mathbf{C}_0 \\ \boldsymbol{a}\mathbf{C}_0 \end{bmatrix} = 0,$$

which implies that $z = 0$ is a root of $\det \begin{bmatrix} \mathbf{C}(z) & \boldsymbol{g}(z) \end{bmatrix}^\top$ and, therefore, $\begin{bmatrix} \mathbf{C}(L) & \boldsymbol{g}(L) \end{bmatrix}^\top$ is not invertible for ε close enough to zero.