HANK beyond FIRE*

José-Elías Gallegos[†]

IIES Stockholm

This version: March 30, 2022

Abstract

The transmission channel of monetary policy in the benchmark New Keynesian (NK) framework relies on the counterfactual Full–Information Rational–Expectations (FIRE) assumption, both at the partial and general equilibrium (GE) dimensions. We relax the Full–Information assumption and build a Heterogeneous-Agents NK model under dispersed information. We find that the amplification multiplier is dampened. This result is explained by the lessened and lagged role of GE effects in our framework. We then conduct the standard full-fledged NK analysis: we find that the determinacy region is widened as a result of *as if* aggregate myopia and show that our framework beyond FIRE does not suffer from the forward guidance puzzle. Finally, we find that transitory "animal spirits" shocks generate persistent effects in output and inflation.

Keywords: Imperfect Information; New Keynesian; Heterogeneous Agents; Monetary Policy.

JEL Classifications: E31, E43, E52, E71.

*I am grateful to Tobias Broer, Alex Kohlhas and Per Krusell for their advice and support, and to Florin Bilbiie and Rajssa Mechelli for discussing my paper at the Oxford NuCamp conference. Further, I would like to thank Mattias Almgren, Gualtiero Azzalini, José Luis Dago Elorza, Richard Foltyn, John Hassler, Philipp Hochmuth, Markus Kondziella, John Kramer, Kurt Mitman, Markus Peters, Claire Thurwächter, Magnus Åhl and seminar participants at the IIES Macro Group and the IIES IMDb for useful feedback and comments. I am indebted to Marios Angeletos for the title of the paper, which I took from his MIT graduate course "Macroeconomics beyond FIRE". I thank Zhen Huo for graciously sharing his teaching materials.

[†]Institute for International Economic Studies, Stockholm University, Universitetsvägen 10, 106 91 Stockholm, Sweden. E-mail: jose.elias.gallegos@iies.su.se

1 Introduction

There is mounting evidence that inequality and information frictions are quantitatively relevant and matter for the transmission of aggregate shocks. On the one hand, the share of financially restricted agents is 34% in the U.S., in an upward trend since 2001, and around 31% in Europe with some countries exhibiting values greater than 40%.[1] Recent theoretical and empirical evidence suggests that economies with a larger degree of inequality respond more to fiscal and monetary shocks.[2] On the other hand, surveys of expectations to consumers, firms, professional forecasters and central bankers suggest that agents' expectations do not satisfy the Full Information Rational Expectations (FIRE) assumption. In particular, there is evidence of aggregate underreaction to news in average forecasts.[3] At the same time, empirical evidence suggests that households' and firms' aggregate underreaction reduces the effect of aggregate shocks; and that the role of GE effects after a monetary policy shock is dampened initially.[4]

To understand in a clean and transparent manner the mechanism of the interaction of these two forces, we build a tractable heterogeneous agents New Keynesian (HANK) model, based on Bilbiie (2019b). Despite its simplicity, this framework captures the key micro-heterogeneity inputs of the quantitative literature: cyclical inequality, idiosyncratic risk and precautionary savings, which together generate heterogeneous marginal propensities to consume (MPCs). In the benchmark FIRE setup, economies with larger inequality react more to exogenous shocks under plausible assumptions. This amplification result arises from the higher MPCs of financially constrained households, and depends on the FIRE assumption at the GE dimension.[5] In this paper we are interested in exploring whether this

1. We consider a household to be financially restricted if it has no liquid savings to self-insure against adverse shocks. The figures reported in the text are taken from Kaplan et al. (2014) and Almgren et al. (2018).

2. See Brinca et al. (2016) for the fiscal policy case, and Almgren et al. (2018) and Bilbiie (2008) for the monetary policy case.

3. Coibion and Gorodnichenko (2015) find that agents when agents, on average, revise their forecasts on unemployment and inflation upwards, they systematically undershoot the realization. These results suggest a rejection of the FIRE assumption as a whole.

4. See Angeletos et al. (2020) and Holm et al. (2021).

5. The mechanism at play is the following. Consider an economy with financially constrained households and optimizers. A monetary shock affects consumption through a substitution effect mandated by the optimizers' Euler condition, which we denote as the partial equilibrium (PE) effect. Households' consumption demand is affected, firms adapt to the new demand schedule and wages (endogenous to labor demand and supply) in turn change. This income effect through wages affects financially constrained agents, which exhibit large MPCs and will magnify the effects of monetary policy. We denote this second roundas the general equilibrium (GE) effect. The transmission channel relies heavily on the FIRE assumption: not only are agents (households and firms) perfectly aware that an aggregate shock has occurred, but are also certain that others have observed it, that others are aware that others have observed it, *ad infinitum*. In particular,

result is robust to a data-consistent deviation from the FIRE assumption. We couple the *HA* dimension with a deviation from the benchmark FIRE assumption in which we assume that agents have imperfect and dispersed information about the state of nature. We find that this framework is consistent with available evidence on the aggregate underreaction to news (Coibion and Gorodnichenko 2015) and the lagged response of GE effects after a monetary policy shock (Holm et al. 2021).

This paper quantifies the amplification multiplier from the *HA* literature away from the counterfactual FIRE assumption. We use our setting to pursue the standard positive and normative analysis in the NK tradition: study determinacy with interest rate rules, where imperfect information relaxes the lower bound on the monetary authority dovishness; and solving the forward-guidance puzzle (FGP). Finally, we study the different effect of a pure monetary policy shock vs. a belief or "animal spirits" shock.

In the standard FIRE setting agents face no uncertainty on the exogenous fundamental, and since information sets are homogenous across individuals, on others' actions. In this paper we accommodate such doubts. At the individual level, agents need to forecast not only the exogenous fundamental (the monetary policy shock) but also aggregate variables that are endogenous to individual actions (output and inflation). As a result, an agent needs to predict other agents' actions. We show that the economy can be described as a pair of *across–group* dynamic beauty contests between consumers and firms (the inflation-spending NK multiplier), with each group playing a *within-group* dynamic beauty contest (the spending-income multiplier running within the demand block and the strategic complementarity in price-setting running within the supply block), and provide new insights on how the PE vs. GE mechanism is dampened through higher-order beliefs in the beyond FIRE framework.[6]

We extend the textbook NK framework in Galí (2008) in two dimensions: financial frictions at the household level and dispersed information. We focus on the amplification multiplier. As laid out by Bilbiie (2019a, 2008) and Galí et al. (2007), as well as richer models by Gornemann et al. (2016), Werning (2015), Auclert (2019) and Hagedorn et al. (2019), whether aggregate shocks have bigger or smaller effects on aggregate consumption, compared to the representative agent framework, is ambiguous. In a model that combines the tractability of TANK models with the most important elements of heterogeneneous agent

the step at which the GE effects kick in, the change in wages and their income effect, depends deeply on the FIRE assumption. It is at this step when the financially constrained agents magnify the aggregate response, since their high MPC interacts with the aggregate wage rate change, giving in turn the well-known amplification result.

6. A dynamic beauty contest is a class of games in which the optimal decision for an individual agent depends on the expectation of the current and future decisions of others.

models, Bilbiie (2019a) shows that the output response to shocks is amplified if the income elasticity of constrained agents with respect to aggregate income is larger than one. He refers to this case as cyclical income inequality; a channel which is strengthened if a larger fraction of agents is constrained.[7] Brinca et al. (2016) and Almgren et al. (2018) find empirical evidence for the amplification multiplier in the case of fiscal and monetary policy, respectively. Regarding the latter, we extend the model to include dispersed information following Lucas (1972) approach to noisy information. Morris and Shin (2002) and Woodford (2003) are the first to study the economy as a static beauty contest, and Allen et al. (2006), Bacchetta and Van Wincoop (2006), Morris et al. (2006), and Nimark (2008) extend the economy to a dynamic beauty contest. More recently, Angeletos and Lian (2018), Angeletos and Huo (2018), and Huo and Takayama (2018) show that dispersed information attenuates the GE effects associated with the Keynesian multiplier and the inflation-spending feedback in a RANK economy, causing the economy to respond to news about the future *as if* the agents were myopic. We extend the framework in Angeletos and Huo (2018) by merging the two building blocks, the Dynamic IS and NK Philips curves, and study the inflation-spending feedback and its implications for the amplification multiplier.

We also study forward guidance in the beyond FIRE framework. We are not the first to attempt to solve the FGP: Del Negro et al. (2012), Mckay (2015), Andrade et al. (2019), Hagedorn et al. (2019), and Angeletos and Lian (2018) have contributed to a growing literature that tries to find an explanation for the FGP from different angles, our approach combining that of Hagedorn et al. (2019) and Angeletos and Lian (2018). We find that, although there is compounding at the aggregate DIS curve arising from countercyclical income inequality, higher-order uncertainty induces enough anchoring to cure the FG puzzle, a failure of the standard NK framework. That is, our model is not subject to what Bilbiie (2019a) denotes *Catch-22*.

The amplification multiplier magnitude is dampened in the dispersed information framework, in which partial equilibirum (PE) effects dominate general equilibrium (GE) effects initially, compared to the FIRE case. In this private and dispersed information economy, agents need to forecast the exogenous fundamental and aggregate inflation and output. While the information friction environment complicates the forecast of the fundamental, it does not give rise to higher-order beliefs since the realisation does not depend on other's actions and an agent does not need to predict others' beliefs on the fundamental. However,

7. In models that focus on the cyclicality of income risk , e.g., Werning (2015), amplification of aggregate shocks is caused by an increase in the probability of becoming constraint for the unconstrained, which leads the latter to save more and consume less.

forecasting aggregate output and inflation has the additional complication of having to deal with higher-order beliefs. In the standard framework, first-order and higher-order beliefs coincide, whereas in our case higher-order beliefs differ from first-order beliefs, and move less than lower-order beliefs since they are more anchored to the prior. As a product of this, the expectations of (endogenous) aggregate variables adjust less to news, and the GE effect is attenuated. The main consequence of the different PE vs. GE role is that aggregate dynamics will be entirely driven by PE effects initially, as estimated in Holm et al. (2021). After some periods and a sequence of signals, agents will learn that a monetary policy shock has occurred, and the aggregate dynamics will rely more and more on GE effects, until the PE share converges to the full information benchmark.[8] We find that (*i*) the peak response of output is about 1/3 of that in the FIRE case; (*ii*) impulse responses are hump-shaped, which the standard FIRE framework can only produce if there is habit formation, price indexation and lumpy investment;[9] and (*iii*) when income inequality is countercyclical (the case studied in Bilbiie (2008)), the response of output after a monetary policy shock is amplified by around 8%, compared to 10% in the benchmark model. That is, dispersed information effectively reduces the amplification multiplier and the overall effect of monetary policy.

Dispersed information adds aggregate anchoring and myopia. Under noisy information, individual forecasts are anchored to agents' own priors. Because expectations play a key role in the determination of aggregate variables in modern macroeconomics, anchoring in expectations translates into aggregate anchoring in endogenous aggregate variables and myopia towards the future. These two results, taken together, enlarge the determinacy region of interest rate rules and solve the FGP. In the NK framework, the determinacy region is ultimately linked to the forward-looking behaviour of the model equations. The Taylor rule provides an essential stabilisation role, and an excessively dovish monetary authority ends up creating explosive dynamics in the model equations. Adding information frictions produces aggregate myopia and widens the determinacy region, a result consistent with the behavioral NK framework in Gabaix (2016). Similarly, the FGP is solved by dispersed information via the introduction of aggregate myopia.

8. Formally, imperfect information reduces the degree of complementarity of actions across agents, and partially mutes the amplification multiplier mechanism that critically relies on them.

9. Havranek et al. (2017) present a meta-analysis of the different estimates of habits in the macro literature and the available micro-estimates. In general, macro models take values around 0.75, whereas micro-estimates suggest a value around 0.4. Groth and Khan (2010) conduct a similar analysis for the investment adjustment frictions case, finding that the microeconomic estimates an order of magnitude below the ones used in the empirical macro literature, in which they are estimated to minimize the distance between model dynamics and empirical IRFs. Finally, the price-indexation model suggests that every price is changed every period, which is inconsistent with micro-data estimates provided by Nakamura and Steinsson (2008).

Our last contribution is to study expectation shocks. We consider the case of public information, and we show that although the non-fundamental shock is only transitory, its effects are persistent, which aligns with the findings in Lorenzoni (2009). Because agents cannot fully disentangle whether the shock to the signal that they have observed comes from the fundamental monetary policy rule or the non-fundamental noise part, the "animal spirits" shock partially inherits the properties of the pure monetary shock, which in turn explains its persistent consequences. In a second extension we consider both public and private information. We find that monetary policy is more effective and closer to the FIRE benchmark, and the effect of belief shocks is lessened, as a result of effectively reducing the degree of information frictions by including an additional signal.

Roadmap The paper proceeds as follows. In section 2 we describe our theoretical framework, focusing on both household financial heterogeneity and dispersed information. Section 3 derives the equilibrium dynamics. In section 4 we discuss the different implications and insights that our HANK model beyond FIRE provides: amplification multiplier, the role of the PE vs. GE share, equilibrium determinacy, forward guidance and "animal spirits" shocks. Section 5 concludes.

2 The Analytical HANK model

The HANK framework described in this section is a reduced-form version of the standard incomplete markets (SIM) model, based on Bilbiie (2019a). Households face an idiosyncratic risk of not being able to access asset markets and firm profits, instead of risk in their labor income. This simplifying assumption allows us to solve the model in paper and pencil, and still provides the desired precautionary savings motive that two-agents New Keynesian (TANK) models lack. On the firm side, the model is kept close to the standard NK framework.

On top of household heterogeneity with respect to their market participation, agents (households and firms) face uncertainty about the state of nature. They receive idiosyncratic signals about the true state, which endogenously generates heterogeneous information sets. Since agents rely on different information, their beliefs and forecasts will differ. This aspect will be crucial for forecasts of endogenous aggregate variables like output or inflation. This gives rise to higher-order beliefs: in order to forecast these endogenous outcomes, an agent needs to forecast the action of other agents, other agents need to forecast the action of others, *ad infinitum*.

2.1 Households

Households save in one-period (liquid) bonds and consume. They have access to financial income, labor income, firm profits and government transfers.

Financial frictions Financial frictions are exogenous to individual behavior, in contrast with the SIM model. In every period, a household is either financially constrained or not. If the household is financially constrained, he is unable to save and loses access to the firm profits, but keeps access to previous-period savings. We denote constrained households as Hand-to-Mouth (HtM). In contrast, unconstrained households benefit from having access to asset markets and firm profits. To insure against the risk of becoming constrained, which entails losing access to part of their resources (firm profits) and the ability to borrow, unconstrained households save in bonds. That is, precautionary savings take the form of liquid bonds. As is standard in simple NK models, we assume that assets are in zero net supply.[10]

In every period there is a realized idiosyncratic shock. The household then knows if he will be financially constrained or not in that period. The exogenous shock takes the form of a Markov chain. Denote s the probability to stay unconstrained, denote h the probability to stay constrained, and denote $1-s$ and $1-h$ the transition probabilities. We assume for simplicity that the markov process induces a stationary distribution. Formally,

$$\begin{pmatrix}\lambda & 1-\lambda\end{pmatrix}\begin{pmatrix}h & 1-h \\ 1-s & s\end{pmatrix} = \begin{pmatrix}\lambda & 1-\lambda\end{pmatrix} \implies \lambda = \frac{1-s}{2-s-h}$$

Notice that this analytical HANK framework nests the standard TANK model when $s = h = 1$ (i.e., in the first period the state of each household is revealed and will never change). We show below that another convenient aspect of this model is that it also nests the RANK framework, and makes comparison between the three settings conveniently easy.

Information frictions On top of the financial heterogeneity dimension, which has a similar structure as Bilbiie (2019a), there is an additional source of heterogeneity. Households are able to observe their own current private variables (the salary they are paid, the consumption and saving decision they make, the transfers they receive) but not all of the current aggregate variables. For instance, they observe all goods prices and are thus able to see the (current) aggregate price index, but they do not observe output, inflation or the nominal

10. There is no institution providing liquidity and there is no capital accumulation.

interest rate.[11] In particular, they are not able to observe perfectly the aggregate (monetary policy) shock, the only state variable. Instead, households observe an imperfectly correlated signal. The signal and information structure will be introduced in section 3. For now it is only important to keep in mind that the aggregate expectation operator does not satisfy the Law of Iterated Expectations (LIE). This information structure produces heterogeneous information sets across households, since each of them has observed a different reality over time. It also hinders household's decision, since it is harder for them to predict what other households' actions will be.

2.1.1 Household problem

There is a measure-1 continuum of ex-ante identical consumers in the economy, indexed by $i \in \mathcal{I}_c = [0,1]$. Household i maximizes an infinite stream of his expected utility over consumption and dis-utility over labor supply,

$$\sum_{t=0}^{\infty} \beta^t \mathbb{E}_{it} u(C_{it}, N_{it})$$

where C_{it} denotes household i's consumption decision at time t, and N_{it} denotes his labor supply choice. Notice that, differently from standard FIRE models, there is an i subscript in the expectation operator, as a result of the heterogeneity in information sets and forecasts.

Unconstrained households A share $(1-\lambda)$ of unconstrained households have access to financial income B_{it}; they also have access to labor income $W_t N_{it}$, where W_t is the aggregate wage rate. Finally, they receive the untaxed share of firm profits $\frac{1-\tau}{1-\lambda} E_t$, where τ is the profit tax rate and E_t. With these resources, an unconstrained household can either consume or save in bonds B_{it} for tomorrow. The solution to their problem, derived in Appendix A, is given by an individual Euler condition,

$$C_{it}^{-\sigma} \geq \beta \mathbb{E}_{it} \left[R_t C_{it+1}^{-\sigma} \right] \tag{2.1}$$

where we have assumed that utility takes a CRRA form, with σ denoting the intertemporal elasticity of substitution and φ the inverse Frisch elasticity. Opening up the expectation

11. Vives and Yang (2016) motivate this through bounded rationality and inattention, while Angeletos and Huo (2018) argue that inflation contains little statistical information about real variables. Huo and Takayama (2018) allow for endogenous information, but such a choice complicates the dynamics and the concept of persistence becomes less clear.

operator, depending on which state the household can potentially go to (Markov structure), the condition can be written as

$$(C_{it}^S)^{-\sigma} = \beta \mathbb{E}_{it} \left\{ R_t \left[s(C_{it+1}^S)^{-\sigma} + (1-s)(C_{it+1}^H)^{-\sigma} \right] \right\} \tag{2.2}$$

Notice that this setting preserves the standard *individual* Euler condition. However, at the *aggregate* level, there will be a discounted Euler condition.

The intratemporal optimality condition of the household $i \in S$ problem is

$$\mathbb{E}_{it} W_t = (C_{it}^S)^{\sigma} (N_{it}^S)^{\varphi} \tag{2.3}$$

which is the optimal labor supply decision.

Constrained households In contrast, a share λ of households is financially constrained. They are banned from asset markets and do not have access to firm dividends, but they still have an intratemporal decision on how much labor to supply, and receive the taxed share of firm profits as government transfers, $\frac{\tau}{\lambda} E_t$. Formally, household $i \in H$ only faces an intratemporal labor decision,

$$\mathbb{E}_{it} W_t = (C_{it}^H)^{-\sigma} (N_{it}^H)^{\varphi} \tag{2.4}$$

which is the optimal labor supply decision.

2.1.2 Aggregate Consumption Function

The following proposition summarizes the aggregate consumption function for each household type.

Proposition 1. *The log-linearized aggregate consumption functions for households of type* S *and* H *at time* t *are*

$$\begin{aligned} c_t^S &= -\frac{\beta}{\sigma} \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^c r_{t+k} - (1-s) \frac{\varphi}{\varphi+\sigma} \left(\frac{1-\tau}{1-\lambda} - \frac{\tau}{\lambda} \right) \sum_{k=1}^{\infty} \beta^k \overline{\mathbb{E}}_t^c e_{t+k} \\ &\quad + (1-\beta) \sum_{k=0}^{\infty} \beta^k \mathbb{E}_{it} \left[\frac{1+\varphi}{\varphi+\sigma} \overline{\mathbb{E}}_t^c w_{t+k} + \frac{\varphi}{\varphi+\sigma} \frac{1-\tau}{1-\lambda} \overline{\mathbb{E}}_t^c e_{t+k} \right] \end{aligned} \tag{2.5}$$

$$c_t^H = \frac{1+\varphi}{\varphi+\sigma} \overline{\mathbb{E}}_t^c w_t^r + \frac{\varphi}{\varphi+\sigma} \frac{\tau}{\lambda} \overline{\mathbb{E}}_t^c e_t \tag{2.6}$$

where $\overline{\mathbb{E}}_t^c(\cdot) = \int_0^1 \mathbb{E}_{it}(\cdot)\, di$ *is the cross-sectional average forecast across households.*

Proof. See Appendix A. □

That is, we can write current aggregate consumption of the S type as a function of future streams of the real interest rate and future aggregate income of the S and H type. On the other hand, the consumption function of the H type depends on the current aggregate wage rate and the current share of transfers they receive.

Condition (2.5) has been derived without assuming a particular information structure, we have simply not applied the LIE at the aggregate level. Therefore, it should be interpreted as a *general* aggregate consumption function. Notice also that we have replaced the standard FIRE expectation operator by $\overline{\mathbb{E}}_t^c(\cdot)$, the average expectation operator for households.

2.2 Firms and the Phillips Curve

Households consume an aggregate basket of goods $j \in \mathcal{I}_\pi = [1, 2]$, which takes the form of the standard CES aggregator

$$C_t = \left(\int_1^2 C_{jt}^{\frac{\varepsilon-1}{\varepsilon}} dj \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

where $\varepsilon > 1$ is the elasticity of substitution between different varieties. Cost minimization from the final good firm implies that the demand from each good is $C_{jt+k} = \left(\frac{P_{jt}}{P_{t+k}}\right)^{-\varepsilon} C_{t+k}$, where $\frac{P_{jt}}{P_t}$ is good j's price in relative terms to the aggregate price index,

$$P_t = \left(\int_1^2 P_{jt}^{1-\varepsilon} dj \right)^{\frac{1}{1-\varepsilon}}$$

Each good is produced by an intermediate monopolistic firm that uses technology linear in labor $Y_{jt} = N_{jt}$.

Aggregate Price Dynamics As in the benchmark NK model, price rigidities take the form of Calvo-lottery friction. At every period, each firm is able to reset their price with probability $(1-\theta)$, independent of the time of the last price change. That is, only a measure $(1-\theta)$ of firms is able to reset their prices in a given period, and the average duration of a price is given by $1/(1-\theta)$. Such environment implies that aggregate price dynamics are given (in log-linear terms) by

$$\pi_t = \int_{\mathcal{I}_f} \pi_{jt} \, dj = (1-\theta) \left[\int_{\mathcal{I}_f} p_{jt}^* \, dj - p_{t-1} \right] = (1-\theta) \left(p_t^* - p_{t-1} \right) \tag{2.7}$$

Optimal Price Setting A firm re-optimizing in period t will choose the price P_{jt}^* that maximizes the current market value of the profits generated while the price remains effective. Formally,

$$P_{jt}^* = \arg\max_{P_{jt}} \sum_{k=0}^{\infty} \theta^k \mathbb{E}_{jt} \left\{ \Lambda_{t,t+k} \frac{1}{P_{t+k}} \left[P_{jt} Y_{j,t+k|t} - \mathcal{C}_{t+k}(Y_{j,t+j|t}) \right] \right\}$$

subject to the sequence of demand schedules

$$Y_{j,t+k|t} = \left(\frac{P_{jt}}{P_{t+k}} \right)^{-\varepsilon} Y_{t+k}$$

where $\Lambda_{t,t+k} \equiv \beta^k \left(\frac{C_{t+k}}{C_t} \right)^{-\sigma}$ is the stochastic discount factor, $\mathcal{C}_t(\cdot)$ is the (nominal) cost function, and $Y_{j,t+k|t}$ denotes output in period $t+k$ for a firm j that last reset its price in period t.

Note that, under flexible prices ($\theta = 0$), $P_{jt}^* = \frac{\varepsilon}{\varepsilon-1} W_t$. Aggregating over firms we obtain the standard result that the aggregate price level is greater than the aggregate marginal cost, due to the markup of monopolistic firms: $P_t = \frac{\varepsilon}{\varepsilon-1} W_t$. Aggregating the optimal labor supply condition (2.3) over households we obtain $N_t^{\varphi} = W_t C_t^{-\sigma}$. Combining the last two conditions, we can write

$$\frac{N_t^{\varphi}}{C_t^{-\sigma}} = W_t = \frac{\varepsilon - 1}{\varepsilon} P_t < P_t^{sp} = W_t$$

where P_t^{sp} is the price set by a hypothetical social planner. That is, the inequality implies that output and employment are below their efficient levels, which comes as a result of monopolistic competition. To solve this suboptimality, the government implements the standard optimal subsidy that induces marginal cost pricing, so that the model is efficient in equilibrium: with desired markup defined by $P_{jt}^* = \frac{\varepsilon}{\varepsilon-1} \frac{1}{1-\tau^s} W_t$, the optimal subsidy is $\tau^s = \frac{1}{\varepsilon-1}$. The profit function is

$$D_{jt} = (1 + \tau^s) P_{jt} Y_{jt} - W_t N_{jt} - T_t^f$$

The subsidy is financed by taxing firms $T_t^f = \tau^s Y_t$, which gives total profits $D_t = P_t Y_t - W_t N_t$.

Proposition 2. *The firm-level Phillips curve is given by*

$$\pi_{jt} = \kappa \theta \mathbb{E}_{jt} y_t + (1 - \theta) \mathbb{E}_{jt} \pi_t + \beta \theta \mathbb{E}_{jt} \pi_{i,t+1} \tag{2.8}$$

where $\pi_{jt} = (1-\theta)\left(p^*_{jt} - p_{t-1}\right)$, $\kappa = \frac{(1-\theta)(1-\beta\theta)}{\theta}(\sigma+\varphi)$, *and the aggregate Phillips curve can be written as*

$$\pi_t = \kappa\theta \sum_{k=0}^{\infty} (\beta\theta)^k \overline{\mathbb{E}}^f_t y_{t+k} + (1-\theta) \sum_{k=0}^{\infty} (\beta\theta)^k \overline{\mathbb{E}}^f_t \pi_{t+k} \tag{2.9}$$

where $\overline{\mathbb{E}}^f_t(\cdot) = \int_0^1 \mathbb{E}_{jt}(\cdot)\, dj$ *is the cross-sectional average forecast across firms.*

Proof. See Appendix A. □

Just as in the household's case, conditions (2.8)-(2.9) are derived under a general information structure, in which we relax the assumption that the aggregate firm expectation operator satisfies the LIE. Notice that this approach gives rise to an *individual* Phillips curve (2.8), each firm j's policy function. Each firm decision can be described as a beauty contest in which they need to forecast current output and inflation, which in turn depend on each household's and firm's actions, and their own future optimal action.

2.3 Fiscal and Monetary Policy

We assume that the government and the monetary authority do not face information frictions and know the current state of nature. The government conducts fiscal and monetary policy. In fiscal terms, on top of the aforementioned optimal production subsidy, it conducts a redistribution scheme: it taxes profits from unconstrained households and rebates the proceedings to the constrained. In log-linear terms

$$e^S_t = \frac{1-\tau}{1-\lambda} e_t$$
$$e^H_t = \frac{\tau}{\lambda} e_t$$

Monetary policy is conducted following a Taylor rule of the form

$$i_t = \phi_\pi \pi_t + \phi_y y_t + v_t \tag{2.10}$$
$$v_t = \rho v_{t-1} + \eta_t \tag{2.11}$$

where the error term follows an AR(1) process to match the empirically observed inertia in the interest rate.

2.4 The Dynamic IS Curve

As in the textbook NK, the model can be summarized in a system of two equations, representing the demand and supply side. Unlike the textbook NK, the system cannot be collapsed to two first-order expectational difference equations. The hierarchy of beliefs prevents the LIE from holding, and the system representation is given by the following proposition.

Proposition 3. *The average-household-level DIS curve is given by*

$$c_{it} = -\frac{\beta}{\sigma}(1-\lambda)\mathbb{E}_{it}r_t + [1-\beta(1-\lambda\chi)]\mathbb{E}_{it}y_t + \beta[\delta(1-\lambda\chi)-1]\mathbb{E}_{it}c_{t+1} + \beta\mathbb{E}_{it}c_{i,t+1} \quad (2.12)$$

and the aggregate DIS curve can be written as

$$y_t = -\frac{\beta}{\sigma}(1-\lambda)\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t r_{t+k} + [1-\beta(1-\lambda\chi)]\overline{\mathbb{E}}_t y_t + (\delta-\beta)(1-\lambda\chi)\sum_{k=1}^{\infty}\beta^k\overline{\mathbb{E}}_t y_{t+k} \quad (2.13)$$

where $\chi = 1+\varphi\left(1-\frac{\tau}{\lambda}\right)$ measures the degree of amplification with respect to RANK (if $\chi > 1$ there is amplification and if $\chi < 1$ there is lessening), and $\delta = 1 + \frac{(\chi-1)(1-s)}{1-\lambda\chi}$ measures the degree of compounding at the consumer's Euler condition (if $\delta > 1$ there is compounding and if $\delta < 1$ there is discounting).

Proof. See Appendix A. □

Again, conditions (2.12)-(2.13) are derived under a general information structure, in which we relax the assumption that the aggregate household expectation operator satisfies the LIE and where agents do not observe aggregate variables. Notice that this approach gives rise to an *individual* DIS curve (2.12), each household i's policy function. Each household decision can be described as a beauty contest in which they need to forecast current real interest rates and future output, which in turn depend on each household's and firm's actions, and their own future optimal action.

Note that, given that the inverse of the Frisch elasticity is strictly positive ($\varphi > 0$), $\chi > 1$ if $\tau < \lambda$. As we will show below, there is amplification of the effects of monetary policy if $\chi > 1$, and dampening otherwise, or if income inequality is countercyclical ($\tau < \lambda$). If instead there were too much redistribution, poor households would be better off at a recession, making income inequality procyclical and switching the result to lessening.[12] Almgren et

12. Werning (2015) argues that there is amplification as long as income inequality is countercyclical and liquidity is not countercyclical. In our model with zero liquidity we are in the acyclical liquidity case.

al. (2020) find empirical evidence for the amplification effects of monetary policy, and we therefore focus in the case $\chi > 1$.

A further remark helps in understanding the dynamics. Consider our benchmark framework with amplification ($\chi > 1$), which in turn implies $\delta > 1$. In an economy without financial frictions, Bilbiie (2019a) shows that $\delta > 1$ (coming from the precautionary savings motive) induces compounding in the aggregate DIS curve. To understand the mechanism, consider our DIS curve beyond FIRE (2.13). Absent information frictions, first-order beliefs coincide with higher-order beliefs and one can simplify the above expression by making use of the LIE and obtain

$$y_t = -\frac{1}{\varsigma}\mathbb{E}_t r_t + \delta \mathbb{E}_t y_{t+1} \tag{2.14}$$

where $\varsigma = \sigma \frac{1-\lambda\chi}{1-\lambda}$. One can see that $\delta > 1$ induces compounding at the aggregate DIS curve. A counterfactual consequence of compounding is that the FGP is exacerbated. That is, in the full information benchmark, one cannot have amplification of monetary policy and cure the FGP. This is a situation that Bilbiie (2019a) denominates *Catch-22*.

An additional benefit of our framework beyond FIRE is that it solves the *Catch-22*. In our case, even if $\delta > 1$, there will be discounting in the aggregate DIS curve (even if the individual Euler conditions preserve compounding due to precautionary savings). Aggregate discounting is, however, hard to see from the beyond FIRE DIS curve (2.13). In fact, it is hidden inside the cross-sectional average expectations. In the beyond FIRE economy, individuals update optimally their expectations using the Wiener filter, a close cousin of the Kalman filter. As a result, current expectations are partially anchored to lagged expectations and move only sluggishly. Since, as is common in the forward-looking NK framework, aggregate outcomes depend crucially on expectations, this anchoring in expectations translates to both anchoring in outcomes and myopia about the future. This myopia, consequence of information frictions, is sufficiently large to outweight the compounding induced by the precautionary savings motive, hereby parameterized in reduced form by $\delta > 1$.

As noted before, a convenient feature of this model is that it nests the more commonly known RANK and TANK settings. If $s = 1$ then $\delta = 1$ and we are in TANK; if further $\lambda = 0$ and $\tau = 0$, then $\chi = 1$ and we are in RANK. We now show in section 3 that the equilibrium solution to the above system (2.13), (2.9), (2.10) and (2.11) can be reduced to a system of 2 first-order difference equations.

3 Information Structure and Equilibrium Dynamics

Let us now describe the information structure. We take a deviation from FIRE that is standard and well-known in the literature: dispersed and noisy information. Similar to Lucas (1972), we assume that agents do not observe the fundamental shock and are therefore uncertain about the state of nature. Every period, each agent receives a dose of private information on the aggregate fundamental. Formally, there is a collection of private Gaussian signals, one per agent and per period. In particular, the period–t signal received by agent k in group g is given by

$$x_{kgt} = v_t + u_{kgt}, \qquad u_{kgt} \sim \mathcal{N}(0, \sigma_g^2) \tag{3.1}$$

where $g = \{\text{household}, \text{firm}\}$, $\sigma_g \geq 0$ parameterizes the noise in group g. Notice that, by allowing σ_g to differ by g, we accommodate rich information heterogeneity (for example, firms could be more informed than households on average).

Suppose that an agent wants to forecast an unobserved fundamental v_t that follows the AR(1) process (2.11) where $\eta_t \sim \mathcal{N}(0, \sigma_\eta^2)$. In such environment, an agent optimal expectation (Kalman filter) of an exogenous AR(1) process takes the following form

$$\mathbb{E}_{kgt} v_t = (1 - G_g)\mathbb{E}_{kg,t-1} v_t + G_g x_{kgt} \tag{3.2}$$

where G_g is the Kalman gain, the weight that agents (optimally) assign on new information x_{kgt} relative to the previous forecast, which depends on the variance of the exogenous marginal cost process σ_η^2, on the variance of the signal noise variance σ_u^2 and on v_t persistence ρ. In order to test the null of the FIRE assumption, or the role of dispersed information, Coibion and Gorodnichenko (2015) suggest the following: regress the ex-ante average forecast error, computed as the difference between the realized variable at $t+1$ and the expectation at time t of that variable at $t+1$ (that is, we compute the average mistake), on the average forecast revision. We define the average forecast revision as the change in the forecast of a variable at time $t+1$ formed at time t minus the forecast of that same variable formed at time $t-1$. Therefore, the forecast revision measures the rigidity of agents' expectations.

$$v_{t+1} - \overline{\mathbb{E}}_{gt} v_{t+1} = \gamma_{gv}(\overline{\mathbb{E}}_{gt} v_{t+1} - \overline{\mathbb{E}}_{g,t-1} v_{t+1}) + u_t \tag{3.3}$$

where $\gamma_v = \frac{1-G_g}{G_g}$ under noisy information. Notice that, under the full information rational expectations assumption, γ_{gv} should be zero. Under full information, each agent individual forecast of a future outcome is identical to each other agents' forecast. As a result, the average expectation operator in (3.3) could be interpreted as a representative agent forecast. Therefore, (3.3) would be effectively regressing the forecast error of the representative agent on its forecast revision. In that case, the forecast revision (dated at time t) should not consistently predict the forecast error. Otherwise, a rational representative agent would incorporate this information, dated at time t, into his information set. Therefore, the above regression suggests that the FIRE assumption is violated in data if $\gamma_{gv} \neq 0$, but is uninformative on whether the full information or the rational expectations (or both) are violated. In our model we will maintain the rational expectations assumption, and assume that agents face information frictions, generating heterogenous beliefs (information sets) across households.[13]

As we show below, the nowcast of inflation will take a similar functional form as (3.2), with the difference that the gain will not only depend on the information structure G_g but also on the other model parameters and on each agent policy function as a result of higher-order beliefs. Therefore, one can regress

$$\pi_{t+1} - \overline{\mathbb{E}}_{gt}\pi_{t+1} = \gamma_{g\pi}(\overline{\mathbb{E}}_{gt}\pi_{t+1} - \overline{\mathbb{E}}_{g,t-1}\pi_{t+1}) + u_t^{\pi} \tag{3.4}$$

for which Coibion and Gorodnichenko (2015) find $\gamma_{g\pi} > 0$ in the data.

Equilibrium Dynamics The equilibrium dynamics must therefore satisfy the individual-level optimal policy functions (2.12) and (2.8), and rational expectation formation should be consistent with the Taylor rule (2.10), the exogenous monetary shock process (2.11) and the signal process (3.1).

In this class of global games in which there is a signal about the stochastic fundamental, the literature has extensively used the Kalman filter to solve for optimal expectation updating, a form of bayesian learning. A caveat from the Kalman method is that it requires the knowledge of the dynamics of the forecasted variables. Recent work by Huo and Takayama (2018) shows that an equivalent version of the Kalman filter, namely the Wiener-Hopf fil-

13. Bordalo et al. (2018) and Broer and Kohlhas (2019) find evidence for a violation of the rational expectations assumption by regressing (3.3) at the individual level, finding evidence for agent over-confidence when forecasting inflation. Notice that, even if we assume information frictions, the above regression at the individual level should report a β estimate of zero, because at the individual level the forecast revision should not consistently predict the forecast error. We do not assume a departure from rational expectations because, as Angeletos and Huo (2018) show, over-confidence would have no effect on aggregate dynamics and would therefore not affect inflation persistence.

ter, can be used to solve for the optimal updating solution in closed-form without knowing the equilibrium dynamics of the forecasted variable. We derive such Wiener-Hopf filter in Appendix B.

In summary, although the system (2.13), (2.9), (2.10) and (2.11) has been solved only numerically in the literature, we take Huo and Takayama (2018)'s approach and show in Proposition 4 that the solutions to the fixed points are simply two ARMA(2,1) processes, which can be written jointly as a VAR(1).

Proposition 4. *In equilibrium the aggregate outcome obeys the following law of motion*

$$\boldsymbol{a}_t = A(\vartheta_1, \vartheta_2)\boldsymbol{a}_{t-1} + B(\vartheta_1, \vartheta_2)v_t \tag{3.5}$$

where $\boldsymbol{a}_t = \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}$ *is a vector containing output and inflation,* $A(\vartheta_1, \vartheta_2)$ *is a* 2×2 *matrix and* $B(\vartheta_1, \vartheta_2)$ *is a* 2×1 *vector*

$$A = \frac{1}{\psi_{11}\psi_{22} - \psi_{12}\psi_{21}} \begin{bmatrix} \psi_{11}\psi_{22}\vartheta_1 - \psi_{12}\psi_{21}\vartheta_2 & -\psi_{11}\psi_{12}(\vartheta_1 - \vartheta_2) \\ \psi_{21}\psi_{22}(\vartheta_1 - \vartheta_2) & -(\psi_{12}\psi_{21}\vartheta_1 - \psi_{11}\psi_{22}\vartheta_2) \end{bmatrix}$$
$$B = \begin{bmatrix} \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right) + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right) \\ \psi_{21}\left(1 - \frac{\vartheta_1}{\rho}\right) + \psi_{22}\left(1 - \frac{\vartheta_2}{\rho}\right) \end{bmatrix}$$

where $(\psi_{11}, \psi_{12}, \psi_{21}, \psi_{22})$ *are fixed scalars that depend on deep parameters of the model, satisfying the following conditions,*

$$\psi_{11} + \psi_{12} = -\frac{1 - \rho\beta}{(1 - \beta\rho)[\varsigma(1 - \delta\rho) + \phi_y] + \kappa(\phi_\pi - \rho)} \tag{3.6}$$
$$\psi_{21} + \psi_{22} = -\frac{\kappa}{(1 - \beta\rho)[\varsigma(1 - \delta\rho) + \phi_y] + \kappa(\phi_\pi - \rho)} \tag{3.7}$$

and $(\vartheta_1, \vartheta_2)$ *are two scalars that are given by the reciprocal of the two largest roots of the characteristic polynomial of the following matrix*

$$\boldsymbol{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix}$$

where

$$C_{11}(z) = \lambda_1 \left\{ (\beta - z)\left(z - \frac{1}{\rho}\right)(z-\rho) + \frac{\sigma_\eta^2}{\rho\sigma_1^2}\beta z \left[z\left(1 - \lambda\chi + \frac{\phi_y(1-\lambda)}{\sigma}\right) - \delta(1-\lambda\chi)\right]\right\}$$
$$C_{12}(z) = -\lambda_1 z \frac{\sigma_\eta^2}{\rho\sigma_1^2}\frac{\beta}{\sigma}(1-\lambda)(1-z\phi_\pi)$$
$$C_{21}(z) = -\lambda_2 z^2 \frac{\sigma_\eta^2}{\rho\sigma_2^2}\kappa\theta$$
$$C_{22}(z) = \lambda_2 \left[(\beta\theta - z)\left(z - \frac{1}{\rho}\right)(z-\rho) + \frac{\sigma_\eta^2}{\rho\sigma_2^2}\theta z (z-\beta)\right]$$

where λ_g*,* $g \in \{1, 2\}$ *is the inside root of the following polynomial*

$$\boldsymbol{D}(z) \equiv z^2 - \left(\frac{1}{\rho} + \rho + \frac{\sigma_\eta^2}{\rho\sigma_g^2}\right) z + 1$$

Proof. See Appendix A. □

The first aspect to notice is that the equilibrium dynamics follow a VAR(1) process. This result is consistent with the empirical macro literature, and easy to interpret. The square coefficient matrix $A(\vartheta_1, \vartheta_2)$ is endogenous to ϑ_1 and ϑ_2 (in fact, ϑ_1 and ϑ_2 are its roots). In our framework, ϑ_1 and ϑ_2 are two parameters that govern information frictions. When the signal noise is high enough such that signals are completely uninformative, ϑ_1 and ϑ_2 reach their maximum value of ρ. On the other hand, when signals are perfectly informative, $\vartheta_1 = \vartheta_2 = 0$. Because they are the roots of $A(\vartheta_1, \vartheta_2)$, $A(0, 0) = \mathbf{0}$. In that case, which is simply the standard NK model with full information, the model dynamics are instead $\boldsymbol{a}_t = B(0, 0)v_t$.

Two aspects are worth discussing. First, the beyond FIRE model produces anchoring, in the sense that $A(\vartheta_1, \vartheta_2) \neq \mathbf{0}$. Importantly, it does so wouthout assuming habits, adjustment costs or backward-looking firms. As shown by Havranek et al. (2017) and Groth and Khan (2010), the microeconomic estimates of those channels are well below the ones used in the empirical macro literature, in which they are estimated to minimize the distance between model dynamics and empirical IRFs. In summary, our model fixes the well-known failure of standard models to produce hump-shaped IRFs by introducing information frictions, and not by assumming ad-hoc sluggishness. Second, the equilibrium dynamics are less sensitive to monetary policy changes. Using our framework, we can compare the impact effect of monetary policy shocks on output and inflation. This is easily verified by comparing $B(\vartheta_1, \vartheta_2)$

Parameter	Description	Value	Source
β	Discount factor	0.99	Bilbiie (2019a)
θ	Calvo probability	0.75	Bilbiie (2019a)
σ	Intertemporal elasticity of substitution	1	Bilbiie (2019a)
φ	Inverse Frisch elasticity	1	Bilbiie (2019a)
ϕ_π	Inflation response in Taylor rule	1.5	Galí (2008)
ϕ_y	Output response in Taylor rule	0.125	Galí (2008)
ρ	Autocorrelation of monetary shock	0.8	Christiano et al. (2005)
σ_η^2	Variance of monetary shock	1	Bilbiie (2019a)
τ	Profit tax rate	0.19	Bilbiie (2019a)
λ	Share of HtM	0.37	Bilbiie (2019a)
s	Pr(unconstrained$_{t+1}$—unconstrained$_t$)	0.96	Bilbiie (2019a)
σ_1^2	Consumer signal innovation variance	2.18	Table II
σ_2^2	Firm signal innovation variance	2.18	Table II

Table I: Parameter values.

and $B(0,0)$: each element in $B(\vartheta_1, \vartheta_2)$ is smaller (in absolute terms) than each element in $B(0,0)$, given that $\vartheta_1, \vartheta_2 \in (0, \rho)$.

4 Applications and Additional Insights

In this section we study the different implications of our HANK beyond FIRE by conducting several policy experiments. We exploit the two main frictions, financial and informational, and explain their joint interaction and consequences. In particular, we show show that the Taylor Principle is satisfied in the economy beyond FIRE (with the determinacy region widened), we explain the key role of PE vs. GE effects and how these are affected by financial frictions, we show that the model solves the FGP, and we obtain the effect of an "animal spirits" shock.

Table I reports the parameters used in the different policy analyses. All these values are standard in the literature. The first block contains the standard RANK parameters. The discount factor β, Calvo inaction probability θ, the intertemporal rate of substitution σ, the inverse Frisch elasticity φ and the variance of the monetary policy shock σ_η^2 have standard values in the literature, taken from Bilbiie (2019a). We take $\rho = 0.8$ from Christiano et al. (2005) and Lindé (2005) to match the empirically observed inertia in the Taylor rule. We set the Taylor rule parameters ϕ_y and ϕ_π to the values used in Galí (2008).

The second block contains the parameters related to household financial heterogeneity. These are taken from Bilbiie (2019a), and include the probability of being financially restricted s, set to match the quarterly autocorrelation of the income process in Guvenen et al. (2014), the profit tax rate τ and the share of HtM λ, jointly set to match the aggregate MPC and the amplification magnitude in Kaplan et al. (2018).

The third block contains the parameters related to imperfect information. The informational friction in our HANK beyond FIRE setting and its dynamics depend critically on how precise signals that consumers and firms receive are. Coibion and Gorodnichenko (2015) were the first to point out the overall failure of the FIRE assumption. They suggest regressing (3.4) on survey data on expectations. Under FIRE, aggregate forecast errors (the left-hand-side term) should be uncorrelated with aggregate forecast revisions (the right-hand-side term). Because the economy can be reduced to a single representative agent, any additional piece on his information set should not produce systematical bias in forecast errors, for he would adjust his optimal action. Coibion and Gorodnichenko (2015) find $\gamma_{g\pi}$ consistently positive across variables and agents, suggesting under-reaction in aggregate forecasts. This, on its own, does not point out to the correct model: it provides evidence to reject FIRE as a whole, but does not suggest if the failure is coming from FI (but maintaining rational expectations) or from RE (but maintaining full information). Very likely, it is a convex combination of the failure of both in the real world. The latter finding motivated a line of research that has produced great advances in the behavioral macro area. Gabaix (2016) proposes a behavioral NK model in which agents miss-perceive the persistence of the exogenous monetary policy shock. Farhi and Werning (2019) suggest a *level-k* theory in which common knowledge is broken at some level k, such that individual best responses are only iterated k times. However, these models are restrictive, in the sense that they do not allow for learning since agents are cognitively bounded. Instead, the FI deviation that we take maintains RE and allows for learning.[14]

In order to find the values for σ_1 and σ_2 that are consistent with the available empirical evidence, we follow Coibion and Gorodnichenko (2015) and regress (3.4). Although our framework is flexible to accommodate heterogeneous signals precision, the literature has extensively focused on inflation forecasts when estimating (3.4). We therefore restrict atten-

14. In a previous version of this paper we allowed for bounded rationality, in the sense that agents misperceive signals' precision (over- and under-confidence) and mis-perceive the exogenous shock persistence (over-extrapolation). The former helps to match the individual-level version of (3.4), but is irrelevant for the aggregate dynamics presented in the IRFs. The latter has a small effect on the aggregate dynamics, driven by the fact that the data suggests a tiny degree of over-extrapolation.

	FE Inflation
FR Inflation	0.643***
	(0.120)
Constant	0.00964
	(0.0314)
Observations	203

HAC robust standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table II: Regression table

tion to inflation, which gives us a single moment to match, and calibrate the informational parameters to match empirical evidence.[15]

We report the estimates our estimates in Table II. We see that the forecast revision coefficient is positive and statistically significant. That is, in our HANK beyond FIRE we calibrate the private information precisions ($\sigma_1 = \sigma_2$) to match the empirical evidence on forecast revisions, $\widehat{\gamma}_{2\pi} = 0.643$. In order to do so, we first need to obtain the model-implied coefficient in our HANK beyond FIRE, $\gamma_{2\pi}^{\mathcal{M}}$. The following proposition serves that purpose.

Proposition 5. *In our beyond FIRE framework the regression coefficient $\mathcal{K}_g$ is given by*

$$\gamma_\pi^{\mathcal{M}} = \frac{\lambda_1^2}{\rho - \lambda_1}\left\{\frac{\psi_{21}(\rho-\vartheta_1)(1-\vartheta_2\lambda_1)[\rho(1-\vartheta_1\lambda_1)+\vartheta_1(1-\rho\lambda_1)-\rho\vartheta_1^2(1-\lambda_1^2)]}{(1-\vartheta_1\lambda_1)[\psi_{21}(\rho-\vartheta_1)(1-\vartheta_2\lambda_1)+\psi_{22}(\rho-\vartheta_2)(1-\vartheta_1\lambda_1)]}\right.$$
$$\left. + \frac{\psi_{22}(\rho-\vartheta_2)(1-\vartheta_1\lambda_1)[\rho(1-\vartheta_2\lambda_1)+\vartheta_2(1-\rho\lambda_1)-\rho\vartheta_2^2(1-\lambda_1^2)]}{(1-\vartheta_2\lambda_1)[\psi_{21}(\rho-\vartheta_1)(1-\vartheta_2\lambda_1)+\psi_{22}(\rho-\vartheta_2)(1-\vartheta_1\lambda_1)]}\right\} \tag{4.1}$$

Proof. See Appendix A ☐

Note that the set $(\lambda_1, \vartheta_1, \vartheta_2, \psi_{21}, \psi_{22})$ is endogenous to the signals' precisions σ_1 and σ_2. In order to find the noise levels that are consistent with our empirical findings, we need to solve the fixed point described by (4.1). We calibrate the pair (τ_1, τ_2) by minimizing the square distance between the model-implied coefficients $\gamma_\pi^{\mathcal{M}}$ and the empirically observed coefficients γ_π. As already reported in Table I, these values imply that $(\sigma_1^2, \sigma_2^2) = (2.18, 2.18)$.

15. We use data from the Survey of Professional Forecasters (SPF). This survey is interesting for us for many aspects. The most important one is that these professionals are asked to give a forecast for each future quarter, which allows us to obtain the forecast revision at any point in time in the data. Second, professional forecasters are an ideal match to firms in our model.

4.1 The Taylor Principle beyond FIRE

As in the standard NK model, the Taylor Principle boils down to studying the determinacy of the system (2.13), (2.9), (2.10) and (2.11). In the standard model, equilibrium is determinate whenever the system is not explosive. It turns out that in these forward-looking models, equilibrium is indeterminate when current outcomes are excessively affected by expectations of the future. One should therefore expect, as discussed in Gabaix (2016), that introducing myopia should widen the determinacy region, making system (2.13), (2.9), (2.10) and (2.11) stable for a larger set of (ϕ_π, ϕ_y) combinations.

Let us start by discussing the full-information rational-expectations benchmark. We are interested in isolating the role of financial frictions. As discussed earrlier, these are modelled in reduced-form by $\frac{1-\lambda}{1-\lambda\chi}$ and δ. In the empirically factual case of amplification, both terms are greater than unity. As we discussed in (2.14), $\delta > 1$ generates compounding in the DIS curve. As a result, the model becomes *more* forward-looking, and the stability region is reduced. To see this formally, we conduct the standard Blanchard and Kahn (1980) analysis in the FIRE case, summarized by the following proposition.

Proposition 6. *The FIRE equilibrium is determinate if*

$$(1 - \beta\delta) + \frac{1}{\varsigma}(\kappa\phi_\pi + \phi_y) > 0 \tag{4.2}$$

$$(1 - \beta)(1 - \delta) + \frac{1}{\varsigma}[\kappa(\phi_\pi - 1) + (1 - \beta)\phi_y] > 0 \tag{4.3}$$

$$(1 + \beta)(1 + \delta) + \frac{1}{\varsigma}[\kappa(\phi_\pi + 1) + (1 + \beta)\phi_y] > 0 \tag{4.4}$$

Proof. See Appendix A. □

In order to isolate the role of each of the financial frictions terms, we first compare a TANK model (in which δ is restricted to 1, for there is no precautionary savings motive) with the benchmark RANK. In these cases, (4.4) is always satisfied for strictly positive Taylor rule coefficients, and conditions (4.2)-(4.3) are reduced to

$$1 - \beta + \frac{1}{\varsigma}(\kappa\phi_\pi + \phi_y) > 0 \tag{4.5}$$

$$\kappa(\phi_\pi - 1) + (1 - \beta)\phi_y > 0 \tag{4.6}$$

Notice that condition (4.6) implies that (4.5) will always hold.[16] As one can see from (4.6), the term ς is completely inocuous when we study determinacy (it will be of interest when we study the sensitivity to aggregate shocks). As a result, the determinacy region in RANK and TANK is identical. This makes transparently clear that it is ultimately δ, which is the companion of the forward-looking element in (2.14), what will drive the restrictions on the Taylor Principle. If we now compare the HANK model (with $\delta > 1$) with the benchmark RANK, we verify that the determinacy region is reduced. In that case, (4.4) is always satisfied so we only need to consider the other two. As one can see from (4.2)-(4.3), $\delta > 1$ is affecting the leftmost term in both equations, by making it negative. As a result, the rightmost element on the left hand side in both conditions needs to be *sufficiently* larger. Precautionary savings are therefore reducing the determinacy region, which we see visually in Figure 1 panel 1a, since they generate compounding in the individual Euler condition.

We now turn to the more interesting beyond FIRE case. Under the parameter values reported in Table I, we conduct the beyond FIRE equivalent of Blanchard and Kahn (1980), which we summarize in Proposition 7.

Proposition 7. *Equilibrium exists and is unique if*

$$1 - \vartheta_1\vartheta_2 > 0 \tag{4.7}$$

$$(1 - \vartheta_1)(1 - \vartheta_2) > 0 \tag{4.8}$$

$$(1 + \vartheta_1)(1 + \vartheta_2) > 0 \tag{4.9}$$

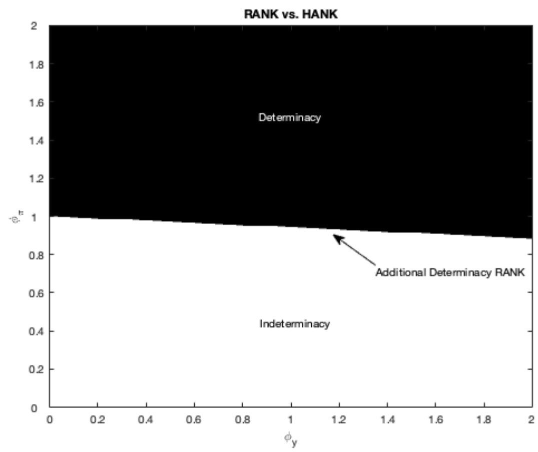
and ϑ_1 and ϑ_2 are the only two outside roots of polynomial $\boldsymbol{C}(z)$, defined in Proposition 4.

Proof. See Appendix A. ☐

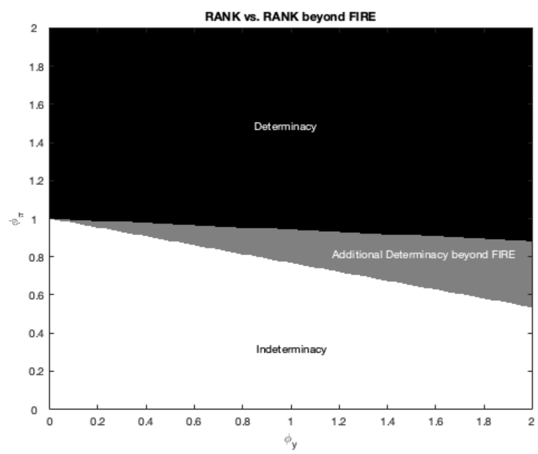
Condition (4.8) is usually the only one that we consider in the standard framework, since the FIRE equivalent of conditions (4.7) and (4.9) are trivially satisfied. In our beyond FIRE framework (4.7)-(4.9) are satisfied since $\vartheta_1 \in (0, \rho)$ and $\vartheta_2 \in (0, \rho)$, with $\rho < 1$. In fact, the most restrictive condition is that ϑ_1 and ϑ_2 are the *only* outside roots of polynomial $\boldsymbol{C}(z)$. Note that ϑ_1 and ϑ_2 are determined endogenously by the deep parameters in the model, so

16. Notice that we can rewrite (4.5) as

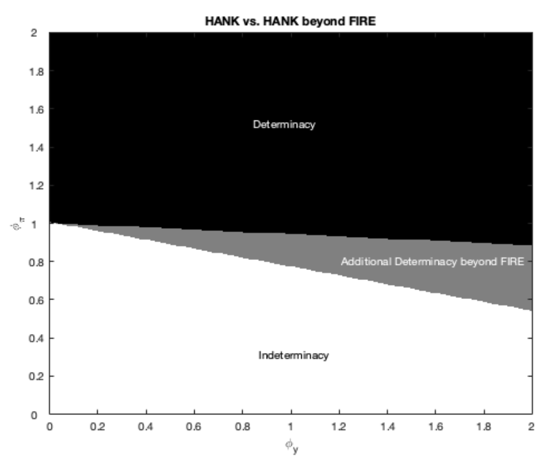
$$\frac{1-\beta}{\sigma}\frac{1-\lambda\chi}{1-\lambda} + \kappa + \beta\phi_y + \kappa(\phi_\pi - 1) + (1-\beta)\phi_y > 0$$



(a) RANK vs. HANK (standard)



(b) RANK vs. RANK beyond FIRE



(c) HANK vs. HANK beyond FIRE

Figure 1: Determinacy regions.

that some parameterizations can yield an indeterminacy even if conditions (4.7)-(4.9) are met but $C(z)$ contains more than two outside roots.

In order to compare how is equilibrium determinacy affected by the imperfect information structure, it is useful to plot the determinacy regions both beyond FIRE and under FIRE. Figure 1 plots the determinacy regions under both frameworks. As one can see, imperfect information widens the determinacy region. This is a result of aggregate myopia. As we discuss in section 4.3, the aggregate dynamics behavior is *as if* agents were myopic. This aggregate myopia is microfounded through (optimal) sluggishness updating of expectations. In the beyond FIRE economy, current expectations are partially anchored to lagged expectations (each individual's prior about the state of nature). Since, as is common in the forward-looking NK framework, aggregate outcomes depend crucially on expectations, this anchoring in expectations translates to both anchoring in outcomes and myopia about the future. Because the present aggregate actions are less sensitive to future actions (than in the standard NK), the determinacy region is widened.

4.2 Response after a Monetary Policy Shock

Our HANK beyond FIRE differs from the textbook NK in two dimensions: household heterogeneity, or *HA*, and information frictions. In order to isolate the effects of financial and information frictions, we will study them separately.

Amplification Let us first consider financial frictions. We are not the first to study the *HA* dimension. Galí et al. (2007) and **bilbiie2008** are two early examples of this literature. Their key result is that, under (plausible) parametric assumptions, adding rule-of-thumb households amplifies the response of aggregate variables to exogenous (monetary and fiscal) shocks. The proposed transmission mechanism works as follows. Unconstrained households change their consumption choice after a monetary policy shock (according to their individual Euler condition), which in turn affects aggregate demand. Because wages are fully flexible, they adjust to the new schedule. This is how the effects of monetary policy reach the HtM. Because they have a unity MPC, they will consume all the income change coming from wages and will magnify any change in aggregate demand.[17]

17. Almgren et al. (2020) test this mechanism in the data. Focusing on Euro Area economies, which are subject to the same monetary policy shock (for they share the Central Bank), they show that monetary policy has heterogeneous effects across countries and that the HtM channel drives these results: the larger the share of HtM (or rule-of-thumb) households in an economy, the larger are the effects of monetary policy.

We show this graphically in Figure 2. In panel 2a we plot the impulse response of output and inflation in the FIRE RANK framework

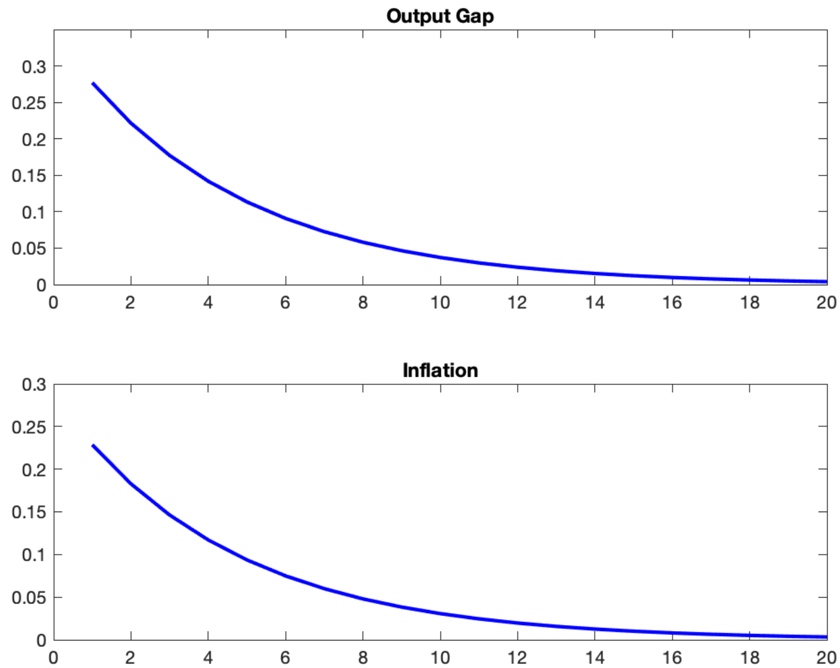
$$\boldsymbol{a}_t = B(0,0)v_t$$

after a monetary policy shock.[18] It is important to notice that in the standard framework without information frictions, the peak response occurs on impact. This, as argued before, comes from the lack of anchoring. Christiano et al. (2005) show that introducing consumption habits, investment adjustment costs and price indexation helps to produce hump-shaped IRFs. However, as argued in section 3, the microestimates for these frictions are an order of magnitude lower than those required in the macro literature, which simply calibrates them to match empirical IRFs. On top of this, empirical evidence suggests that *individual* consumption responses following an income shock have a monotonically decreasing pattern, which makes the consumption habits channel counterfactual (see Fagereng et al. 2019). A second remark is that the HtM transmission channel is present. In order to quantify the effects coming from the HA channel, consider a TANK and a HANK economy that are perturbed by the same monetary policy shock. We plot them in panel 2b, together with their RANK counterpart. We find that both output and inflation are more responsive in the HA economies, consistent with the empirical findings in Almgren et al. (2020), due to the HtM channel ($\chi > 1$). In particular, the TANK economy produces a peak value in the output IRF that is 190bp larger; and these effects are maximized in the HANK setting with precautionary savings where the peak value in the output IRF that is 280bp larger.

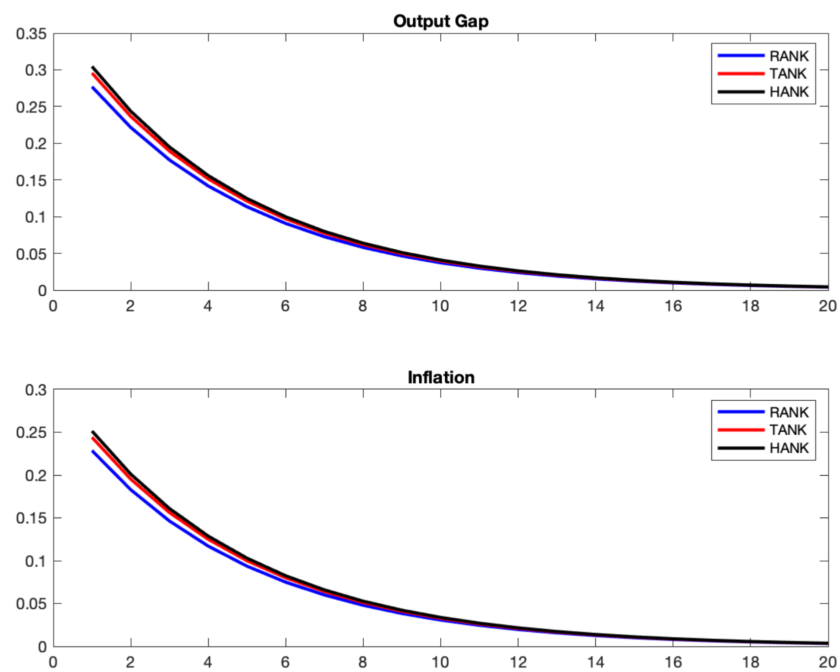
Two problems arise in the light of Figure 2. First, the finding that output increases by 0.3% after a 25bp monetary policy shock seems excessive. The empirical macro literature generally presents results in the range of $0.05\% - 0.15\%$ (see e.g. Ramey (2016) for a literature review.) Second, we find that IRFs' peak occurs on impact. Empirical evidence suggests that there is sluggishness both for output and inflation at the aggregate level, but the necessary micro adjustments made in theoretical frameworks are generally not micro-consistent. We propose a solution to these puzzles that takes the form of dispersed information, and we find that our framework beyond FIRE reconciles the micro- and macro-econometric evidence.

PE vs. GE In line with empirical evidence, the HtM channel proposed by Galí et al. (2007) and **bilbiie2008** is also present in our HANK beyond FIRE, yet partially muted. As long

18. A convenient feature of our model beyond FIRE, as already discussed, is that it nests the standard framework when we restrict $\vartheta_1 = \vartheta_2 = 0$, and makes thus easier the comparison.



(a) RANK



(b) RANK, TANK and HANK

Figure 2: Output gap and Inflation dynamics after a 25bp monetary policy shock

as there is not excessive fiscal redistribution, parameterized by $\tau < \lambda$, there is amplification of monetary policy. Our main finding is that the amplification magnitude is dampened. In order to interpret the role of information frictions in the model, it is convenient to decompose the total response in the DIS curve (2.13) into partial (direct) and general (indirect) effect components:

$$y_t = \underbrace{-\frac{\beta}{\sigma}(1-\lambda)\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t r_{t+k}}_{\text{PE effect}} \underbrace{+[1-\beta(1-\lambda\chi)]\overline{\mathbb{E}}_t y_t + (\delta-\beta)(1-\lambda\chi)\sum_{k=1}^{\infty}\beta^k\overline{\mathbb{E}}_t y_{t+k}}_{\text{GE effect}} \quad (4.10)$$

In IRF-terms, we can write the IRF at time $\tau \in \{t, t+1, t+2, ...\}$ in terms of the two PE and GE components,

$$\text{IRF}_{t,\tau} = \frac{\partial \text{PE}_\tau}{\partial \eta_t} + \frac{\partial \text{GE}_\tau}{\partial \eta_t}; \quad \tau \geq t$$

where $\text{IRF}_{t,\tau} = \frac{\partial y_\tau}{\partial \eta_t}$, and PE_τ and GE_τ are the direct (or partial equilibrium) effect and the general equilibrium effect. Let us now define the PE share μ_τ as

$$\mu_\tau = \frac{\text{PE}_\tau}{\text{PE}_\tau + \text{GE}_\tau}$$

The following proposition provides the PE share μ_τ beyond FIRE

Proposition 8. *Beyond FIRE, the time-varying PE share μ_τ is given by*

$$\mu_\tau = -\frac{\beta}{\sigma}(1-\lambda)\rho\frac{\delta_1\rho^\tau + \delta_2\lambda_1^\tau + \delta_3\vartheta_1^\tau + \delta_4\vartheta_2^\tau}{\psi_{11}(\rho^{\tau+1} - \vartheta_1^{\tau+1}) + \psi_{12}(\rho^{\tau+1} - \vartheta_2^{\tau+1})}$$

where

$$\delta_1 = \frac{1 + \phi_y(\psi_{11} + \psi_{12}) + (\phi_\pi - \rho)(\psi_{21} + \psi_{22})}{1 - \rho\beta}$$

$$\delta_2 = \frac{\lambda_1}{\rho^2(1-\rho\beta)}\Bigg\{ -\rho + \phi_y\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\lambda_1 - \rho\vartheta_j[\beta + \lambda_1(1-\beta(\rho+\vartheta_j-\lambda_1))]]\psi_{1j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}$$

$$+ \phi_\pi\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\lambda_1 - \rho\vartheta_j[\beta + \lambda_1(1-\beta(\rho+\vartheta_j-\lambda_1))]]\psi_{2j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}$$

$$-\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\rho\lambda_1(1+\rho\beta\vartheta_j^2) - \vartheta_j(\rho - \lambda_1(1-\rho(\beta+\lambda_1)))]\psi_{2j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}$$

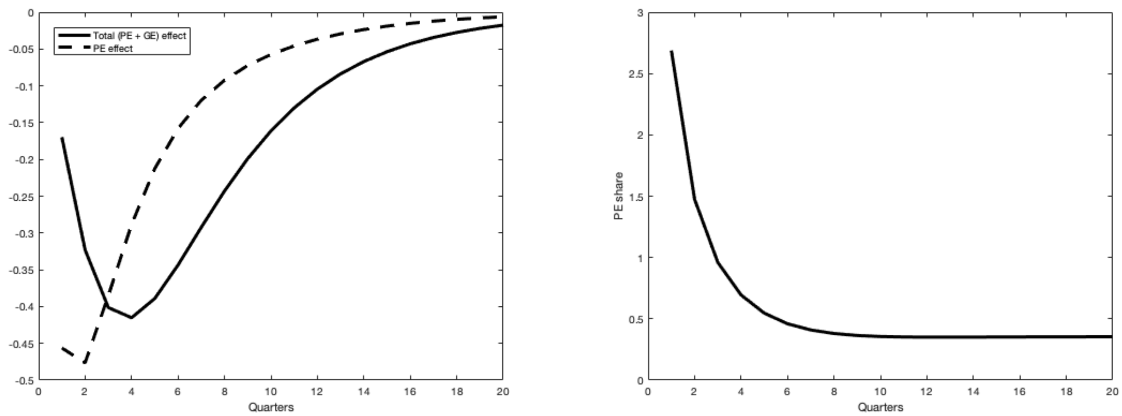
$$\delta_3 = -\frac{\vartheta_1^2(\rho-\lambda_1)(1-\rho\lambda_1)[\phi_y\psi_{11}+(\phi_\pi-\vartheta_1)\psi_{21}]}{\rho^2(1-\beta\vartheta_1)(\vartheta_1-\lambda_1)(1-\vartheta_1\lambda_1)}$$
$$\delta_4 = -\frac{\vartheta_2^2(\rho-\lambda_1)(1-\rho\lambda_1)[\phi_y\psi_{12}+(\phi_\pi-\vartheta_2)\psi_{22}]}{\rho^2(1-\beta\vartheta_2)(\vartheta_2-\lambda_1)(1-\vartheta_2\lambda_1)}$$

Proof. See Appendix A □

We plot the total response, the PE response and the PE share over time μ_τ in Figure 3. We find that, while the total effect is muted initially, PE effects arise immediately. This result, consistent with the empirical findings in Holm et al. (2021), comes from the information frictions dimension. GE effects depend on the hierarchy of beliefs, with each higher-order belief creating more anchoring. On the contrary, PE effects do not depend on higher-order beliefs, since the outcome realization does not depend on agents' beliefs.[19] The beyond FIRE PE share μ_τ is initially high (even above 1, consistent again with the empirical evidence in Holm et al. 2021) and converges over time to the FIRE PE share.

In order to interpret the differences with the FIRE case, it is key to understand that the transmission mechanism proposed by Galí et al. (2007) and **bilbiie2008** relies heavily on GE effects. The key impact of constrained households' MPC relies on them being perfectly aware of the state of nature *and* of others' actions for the nowcast of real wages. Notice that in this framework agents need to forecast not only the exogenous fundamental (the monetary policy shock) but aggregate inflation and output. While the information friction environment complicates the forecast of the fundamental, it does not give rise to higher-order beliefs since the realization does not depend on other's actions. On the other hand, forecasting aggregate output and inflation has the additional complication of having to deal with higher-order beliefs: agents need to infer what others believe, agents need to infer what others think they belief, *ad infinitum*. This rise of higher-order beliefs, which are more anchored to the prior at each increasing order, increases the anchoring in the GE dimension. As a result, aggregate dynamics will be entirely driven by PE effects initially. After some periods, agents will learn that a (persistent) monetary policy shock has occurred, and the aggregate dynamics will rely more and more on GE effects, until the PE vs. GE share converges to the full information benchmark. As one can see in Figure 3b, the GE multiplier is arrested in the first periods. As time goes by and agents have received enough signals, their aggregate action converges to the FIRE one (as we will see in Figure 4) and the PE share μ_τ converges to the standard

19. A caveat to this result is that PE effects will depend on higher-order beliefs in our case, since the real interest rate is endogenous to the stabilization role of the Taylor rule. Quantitatively, we find that PE effects are less anchored than GE effects.



(a) Total vs. PE effect.

(b) PE share μ_τ over time.

Figure 3: Total, Direct and Indirect Effects.

FIRE value.

To summarize the PE vs. GE discussion, imperfect information reduces the degree of complementarity of actions across agents, although it is important to remark that the amplification mechanism is still present in the model. Higher-order uncertainty, or beliefs, effectively arrests and slows down the GE effect.

Impulse Response Functions Now that we understand how is the HtM channel modified by dispersed information and the role of PE vs. GE effects, let us study the equilibrium dynamics. Suppose that the monetary authority shocks the economy (3.5) with a 25bp monetary policy shock. Figure 4 plots the impulse response of output and inflation.

Two aspects are worth mentioning. First, the HtM channel is still present. A larger degree of financial frictions leads to a more responsive economy to aggregate shocks (see table III). This result, as argued before, comes from countercyclical income inequality (here parameterized by $\chi > 1$). Our main finding is that the role of such mechanism is partially muted by dispersed information. The TANK economy produces a peak value in the output IRF that is 51bp larger; and these effects are maximized in the HANK setting with precautionary savings where the peak value in the output (inflation) IRF that is 77bp. Table III gives numbers for the amplification *magnitude* lessening in the dispersed information framework. Second, the peak effect in output is around 1/3 of that of the standard framework, around 0.10% and in line with the findings in Ramey (2016), and the IRFs have the hump-shape that we observe in the data (see e.g., Christiano et al. (2005) and Ramey (2016))

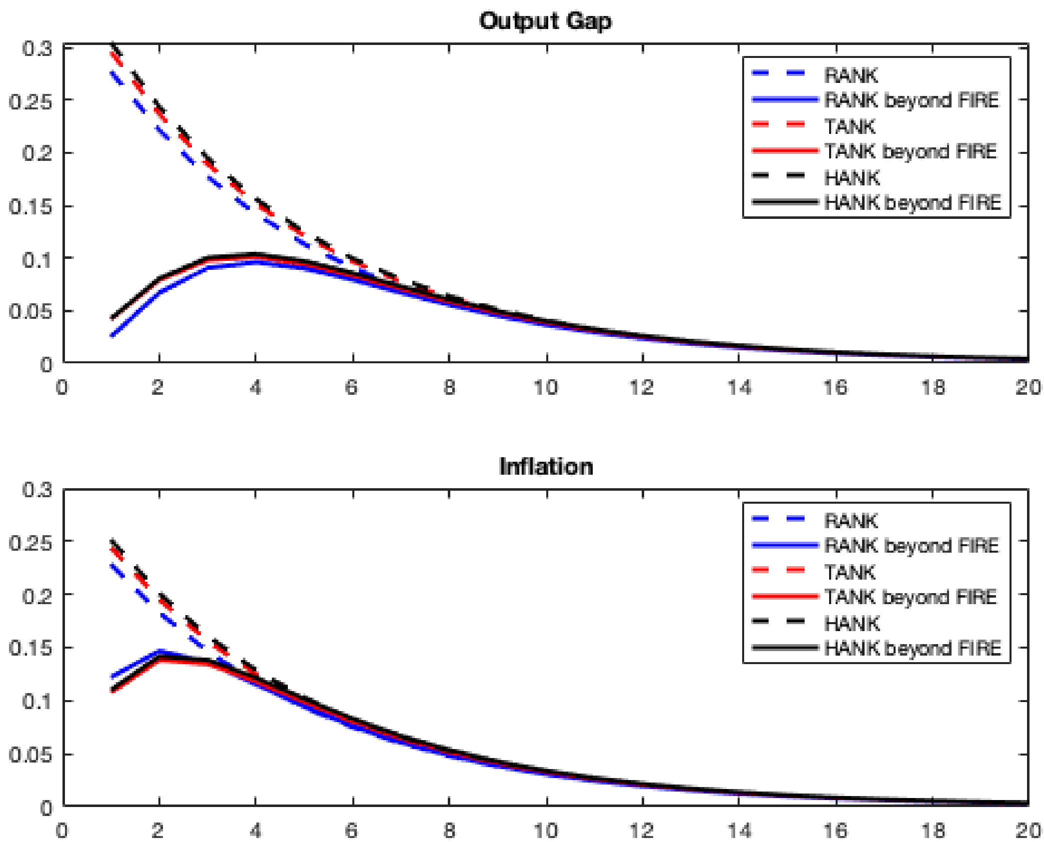


Figure 4: Impulse responses after a monetary policy shock.

Note: FIRE dynamics in dashed lines, beyond FIRE dynamics in straight lines. RANK dynamics in blue, TANK dynamics in red and HANK dynamics in black.

Framework	Comparison (vs. RANK)	Amplification (peak)
Standard	*TANK*	6.76%
	HANK	9.97%
Beyond FIRE	*TANK*	5.33%
	HANK	8.05%

Table III: Amplification magnitude from Output IRFs

without compromising the individual (monotonically decreasing) responses to income shocks documented in Fagereng et al. (2019).

4.3 Forward Guidance

A documented failure of the standard NK model is the *Forward Guidance Puzzle*. Forward guidance is an unconventional monetary policy tool that can be used by Central Banks in a situation in which the nominal interest rate (their main policy tool) is stuck at zero, so that further expansionary conventional policy is unfeasible. The Central Bank commits to keeping nominal interest rates low (relative to what their Taylor rule would mandate), in the hope of unanchoring inflation expectations and output. In the recent financial crisis, several Central Banks made use of it (see e.g. Angeletos and Sastry (2020) for a more comprehensive treatment.) The excessively forward-looking standard NK model predicts that a forward guidance $\tau-$shock (i.e., a promise at time t to shock the economy in period $\tau \geq t$ by using the real interest rate) has *the same (or more) effect* the more in the future it is promised. To see this, we iterate forward the FIRE DIS curve (2.14)

$$y_t = -\frac{1}{\varsigma}\sum_{j=0}^{\infty}\delta^j \mathbb{E}_t r_{t+j}$$

where we have assumed that expectations of long-run output gap are equal to the steady-state value, $\lim_{T\to\infty}\delta^T \mathbb{E}_t y_{T+1} = 0$. Recall that in the standard NK $\varsigma = \sigma$ and $\delta = 1$. In that case, $y_t = -1/\sigma\sum_{j=0}^{\infty}\mathbb{E}_t r_{t+j}$, and any future shock on the nominal interest rate (a forward guidance shock) has an *identical* impact on today's output, irrespective of the period in which it is realized. This is aggravated in the case in which there are financial constraints, since the precautionary savings motive and amplification induce compounding ($\delta > 1$) at the aggregate level, making the process explosive: the further in the future the shock takes place, the larger the increase in the output gap today. This is the situation that Bilbiie (2019a) denominates Catch-22: a realistic amplification of monetary policy effects aggravates the FGP. It is, however, wishful-thinking that this policy tool, perfectly valid in zero lower bound (ZLB) periods, is *so* effective. Del Negro et al. (2012) were the first to study this empirically, and find that forward guidance is indeed less effective than what the theoretical model suggests.

We argue in Proposition 9 that the information frictions at the individual level induce anchoring and myopia at the aggregate level, as discussed in Angeletos and Huo (2018). This result will be sufficient in order to cure the FGP, while still maintain the amplification result. In order to analyze the effects of forward guidance in our HANK beyond FIRE framework, consider a situation in which the economy is stuck in a liquidity trap. Suppose that the zero lower bound (ZLB) for nominal interest rates is binding between periods t and τ, such that $\tau \geq t$. The following proposition rewrites the DIS curve beyond FIRE in FIRE terms, and proves that there is FGP anymore.

Proposition 9. *The ad-hoc equilibrium dynamics*

$$\boldsymbol{a}_t = \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1} + \overline{\boldsymbol{\varphi}} v_t \tag{4.11}$$

produce identical dynamics to the supply-side dispersed information model for certain pair of 2×2 *matrices* $(\boldsymbol{\omega}_b, \boldsymbol{\omega}_f)$. *The DIS curve and the Phillips curve can be written in FIRE terms as*

$$\begin{aligned} y_t &= \omega_{by} y_{t-1} + \omega_{b\pi} \pi_{t-1} - \frac{1}{\varsigma}(i_t - \mathbb{E}_t \pi_{t+1}) + \omega_{fy} \mathbb{E}_t y_{t+1} + \left(\omega_{f\pi} - \frac{1}{\varsigma}\right) \mathbb{E}_t \pi_{t+1} \\ \pi_t &= \delta_{by} y_{t-1} + \delta_{b\pi} \pi_{t-1} + \kappa y_t + \delta_{fy} \mathbb{E}_t y_{t+1} + \delta_{f\pi} \mathbb{E}_t \pi_{t+1} \end{aligned} \tag{4.12}$$

where $\{\omega_{by}, \omega_{b\pi}, \omega_{fy}, \omega_{f\pi}, \delta_{by}, \delta_{b\pi}, \delta_{fy}, \delta_{f\pi}\}$ *are scalars defined in Appendix A. Dispersed information cures the Forward Guidance Puzzle if one of the roots of the polynomial* $\mathcal{P}(x) \equiv \omega_{fy} x^2 - x + \omega_{by}$ *lies outside the unit circle, and the other root lies inside the unit circle. Furthermore, the effect of forward guidance at period* τ *on consumption at period* t *is given by*

$$FG_{t,t+\tau} = \frac{\partial \tilde{y}_t}{\partial \mathbb{E}_t r_{t+\tau}} = -\left(\frac{\omega_{f\pi}}{\omega_{fy}\zeta} + \frac{\omega_{b\pi}}{\omega_{fy}\zeta^3}\right)\frac{1}{\zeta^\tau} \tag{4.13}$$

where $|\zeta| > 1$ *is the outside root of the polynomial* $\mathcal{P}(x)$.

Proof. See Appendix A □

A caveat of the above proposition is that scalars $\{\omega_{by}, \omega_{b\pi}, \omega_{fy}, \omega_{f\pi}, \delta_{by}, \delta_{b\pi}, \delta_{fy}, \delta_{f\pi}\}$ are not unique, although dynamics are unique. That is, different weights are consistent with the

	(1) DIS	(2) Phillips Curve
y_{t-1}	0.418***	−0.196**
	(0.0393)	(0.0992)
π_{t-1}	−0.108*	0.348***
	(0.0563)	(0.0767)
y_t		0.502***
		(0.194)
π_t	0.169*	
	(0.100)	
y_{t+1}	0.609***	−0.331***
	(0.0475)	(0.126)
π_{t+1}	−0.0499	0.638***
	(0.0837)	(0.0771)
Observations	203	203

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IV: Regression table

equilibrium dynamics described by (3.5). Intuitively, agents' actions can be anchored/myopic with respect to aggregate output or inflation, or a combination of both. Hence, in order to study the dynamics in the Phillips curve and the FGP, the theorist is left with two degrees of freedom. We therefore restrict ω_{by} and $\delta_{b\pi}$ to equate its empirical counterparts, reported in Table IV, and the myopia parameters ω_{fy} and $\omega_{f\pi}$ will be then determined endogenously.[20] Secondly, in the benchmark NK model with no information frictions we have $\omega_{by} = \omega_{b\pi} = \delta_{by} = \delta_{b\pi} = \delta_{fy} = 0$, $\omega_{fy} = \delta$, $\omega_{f\pi} = 1/\varsigma$ and $\delta_{f\pi} = \beta$ and the DIS and Phillips curves are reduced to (2.14) and (A.49), respectively.

Proposition 9 derives the general DIS curve in FIRE terms (4.12). In order to analyze the effects of forward guidance in our HANK beyond FIRE framework, consider a situation in which the economy is stuck at the ZLB in which nominal interest rates are binding at the zero constraint, $i_k = 0$ for $k \in (t, \tau)$. As a result, the ex-ante real interest rate is the (log)

20. To estimate the DIS and Phillips curves we rely on GMM methods, using four lags of the Effective Fed Funds rate, GDP Deflator, CBO Output Gap, Commodity Price Inflation, Real M2 Growth and the spread between the long-term bond rate and the three-month Treasury Bill rate as instruments.

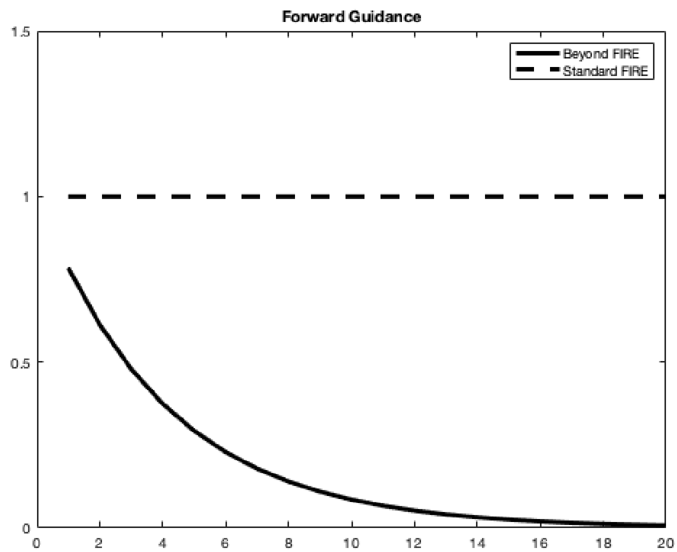


Figure 5: The effect of Forward Guidance on current output, RANK.

Note: Results shown for the RANK framework. FGP is also cured in the TANK and HANK cases ($|\zeta| > 1$), but exhibit an oscillatory pattern driven by $\zeta < -1$.

inverse of expected inflation, $\mathbb{E}_t r_k = -\mathbb{E}_t \pi_{k+1}$. In this case, the DIS curve (4.12) becomes

$$y_t = \omega_{by} y_{t-1} + \omega_{b\pi} \pi_{t-1} - \omega_{f\pi} \mathbb{E}_t r_t + \omega_{fy} \mathbb{E}_t y_{t+1}$$

Notice how dispersed information adds anchoring and myopia in the DIS curve: anchoring is added both via output and inflation by introducing two additional lagged terms. On the other hand, myopia is introduced by introducing a term $\omega_{fy} < 1 \leq \delta$. Because during the ZLB $\mathbb{E}_t r_k = -\mathbb{E}_t \pi_{k+1}$, and there is also myopia with respect to future inflation, the contemporaneous effect of a real interest rate shock is also diminished, $\omega_{f\pi} < 1/\varsigma$. In Figure 5 we plot the impact of a forward guidance shock at period τ on today's output for each τ. The FGP is cured, so that the further in time the forward guidance is implemented, the lesser the effect.

4.4 Beliefs Shock

4.4.1 Public Information

In this section we replace private information by public information, and obtain the model dynamics after a shock to the common signal. The benchmark model does not allow for this exercise, since a shock to an individual signal (whether household or firm) does not have any effect on aggregate variables, since agents are atomistic. We keep the rest of the model unchanged, except for the information structure. Instead of the individual signal $x_{igt} = v_t + u_{igt}$, all agents receive a common and public noisy signal informing them on the monetary policy shock v_t. Formally, there is a collection of public Gaussian signals, one per period and common across agents. In particular, the period–t signal received by all agents, regardless of their group g, is given by

$$z_t = v_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

where $\sigma_\epsilon \geq 0$ parameterizes the noise in the common signal. The following proposition summarizes the equilibrium dynamics under public information.[21]

Proposition 10. *In equilibrium the aggregate outcome obeys the following law of motion*

$$\boldsymbol{a}_t = A(\vartheta_1, \vartheta_2)\boldsymbol{a}_{t-1} + B(\vartheta_1, \vartheta_2)v_t + B(\vartheta_1, \vartheta_2)\epsilon_t \tag{4.14}$$

where $\boldsymbol{a}_t$ is a vector containing output and inflation, $A(\vartheta_1, \vartheta_2)$ is a 2×2 matrix and $B(\vartheta_1, \vartheta_2)$ is a 2×1 vector, both already presented in Proposition 4, where $(\psi_{11}, \psi_{12}, \psi_{21}, \psi_{22})$ are fixed scalars already presented in Proposition 4 and $(\vartheta_1, \vartheta_2)$ are two scalars that are given by the reciprocal of the two largest roots of the characteristic polynomial of the following matrix

$$\boldsymbol{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z) \end{bmatrix}$$

where

$$C_{11}(z) = \hat{\lambda}\left\{(\beta - z)\left(z - \frac{1}{\rho}\right)(z - \rho) + \frac{\sigma_\eta^2}{\rho\sigma_\epsilon^2} z\left[z\left(1 + \frac{\phi_y}{\varsigma}\right) - \delta\right]\right\}$$

$$C_{12}(z) = -\hat{\lambda} z \frac{\sigma_\eta^2}{\rho\sigma_\epsilon^2}\frac{\beta}{\varsigma}(1 - z\phi_\pi)$$

$$C_{21}(z) = -\hat{\lambda} z^2 \frac{\sigma_\eta^2}{\rho\sigma_\epsilon^2}\kappa\theta$$

$$C_{22}(z) = \hat{\lambda}\left[(\beta\theta - z)\left(z - \frac{1}{\rho}\right)(z - \rho) + \frac{\sigma_\eta^2}{\rho\sigma_\epsilon^2}\theta z\,(z - \beta)\right]$$

21. Although an extension in which the common signal is only common within each group is perfectly feasible, we find that the extension complicates significantly the (simple) representation of the model dynamics in (4.14).

where $\hat{\lambda}$ *is the inside root of the following polynomial*

$$\boldsymbol{D}(z) \equiv z^2 - \left(\frac{1}{\rho} + \rho + \frac{\sigma_\eta^2}{\rho\sigma_\epsilon^2}\right) z + 1$$

Proof. See Appendix A. □

The first aspect to notice is that the equilibrium dynamics still follow a VAR(1) process, with an additional contemporaneous exogenous shock ϵ_t. This term can be interpreted as a belief or "animal spirits" shock. Notice that both shocks have identical effects on impact on aggregate variables, given that agents cannot completely disentangle the noise and the fundamental shock from the signal. However, since the belief shock ϵ_t is transitory and not autocorrelated, it has less long-lasting effects than the monetary policy shock. We plot the impulse responses of output, inflation and the policy rate in Figure 6.[22]

Although the belief shock is purely transitory, it produces persistent and hump-shaped dynamics of output over time. This is the result of having imperfectly informed agents, which cannot immediately differentiate between a belief shock and a true monetary policy shock. Notice also the different response of the policy rate: after the expansionary monetary policy shock the policy rate moves down. On the other hand, after the non-fundamental belief shock, the Central Bank raises interest rates to cool down the economy, which reduces the general equilibrium effect of the belief shock.

4.4.2 Private and Public Information

What if instead of replacing private by public signals, we allow agents to observe two signals, one private and one public? In this section we extend the model to include public information, and obtain the model dynamics after a shock to the common signal. On top of the individual signal $x_{igt} = v_t + u_{igt}$, all agents receive a common and public noisy signal informing them on the monetary policy shock v_t. Formally, there is a collection of public Gaussian signals, one per period and common across agents. In particular, the period–t signal received by all agents, regardless of their group g, is given by

$$z_t = v_t + \epsilon_t, \qquad \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$$

where $\sigma_\epsilon \geq 0$ parameterizes the noise in the common signal. The following proposition summarizes the equilibrium dynamics under public information.

22. We set the public signal noise to the value reported in Table I, $\sigma_\epsilon^2 = 2.22$.

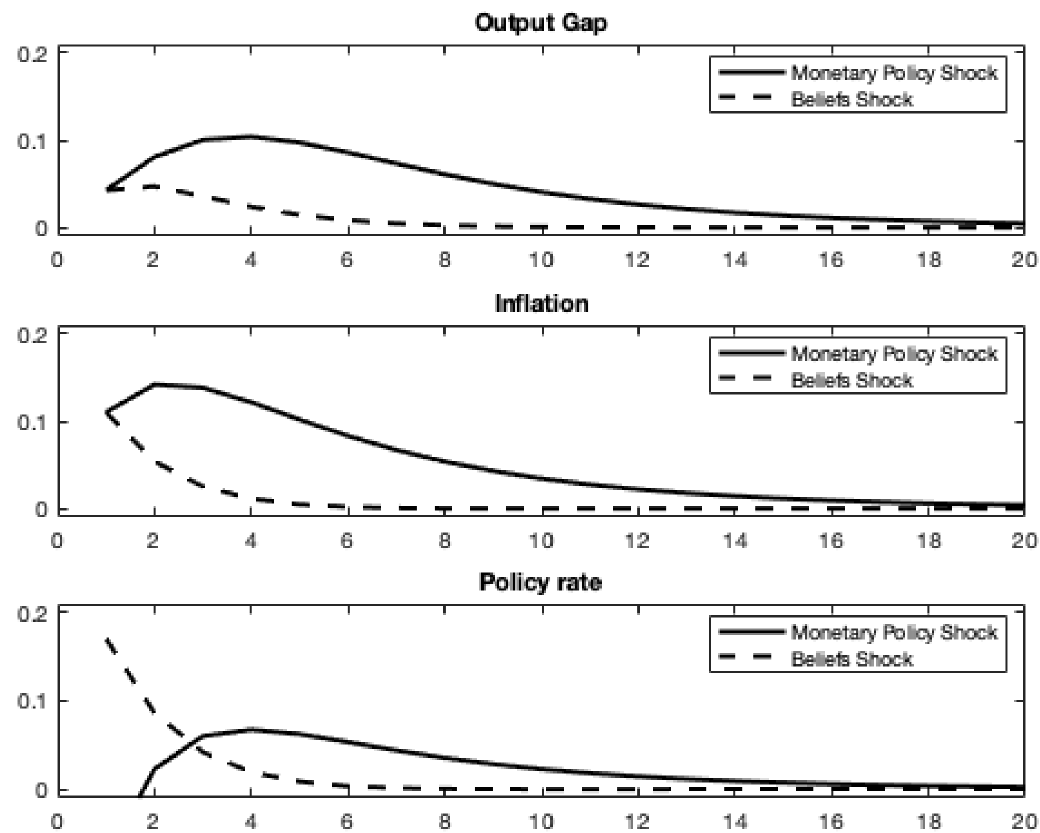


Figure 6: Impulse Responses of output, inflation and the policy rate after a monetary policy shock (dark line) and a belief shock (dotted line) in the HANK economy.

Proposition 11. *In equilibrium the aggregate outcome obeys the following law of motion*

$$\boldsymbol{a}_t = Q_v \sum_{k=0}^{\infty} \Lambda^k \Gamma v_{t-k} + Q_\epsilon \sum_{k=0}^{\infty} \Lambda^k \Gamma \epsilon_{t-k} \tag{4.15}$$

where $\boldsymbol{a}_t = \begin{bmatrix} y_t \\ \pi_t \end{bmatrix}$ *is a vector containing output and inflation, and*

$$Q_v = \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix}, \qquad Q_\epsilon = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$$
$$\Lambda = \begin{bmatrix} \vartheta_1 & 0 \\ 0 & \vartheta_2 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 1 - \vartheta_1/\rho \\ 1 - \vartheta_2/\rho \end{bmatrix}$$

where $(\psi_{11}, \psi_{12}, \psi_{21}, \psi_{22}, \phi_{11}, \phi_{12}, \phi_{21}, \phi_{22})$ *are fixed scalars that depend on deep parameters of the model, and* $(\vartheta_1, \vartheta_2)$ *are two scalars that are given by the reciprocal of the two largest roots of the characteristic polynomial of the following matrix*

$$\boldsymbol{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) & C_{13}(z) & C_{14}(z) \\ C_{21}(z) & C_{22}(z) & C_{23}(z) & C_{24}(z) \\ C_{31}(z) & C_{32}(z) & C_{33}(z) & C_{34}(z) \\ C_{41}(z) & C_{42}(z) & C_{43}(z) & C_{44}(z) \end{bmatrix}$$

where

$$C_{11}(z) = \beta \left[(1 - \lambda\chi) \left(1 - \frac{\delta \sigma_\eta^2}{z} \right) + \frac{\phi_y (1 - \lambda)}{\sigma} \right]$$
$$C_{12}(z) = -\frac{\lambda_1 \sigma_\eta^4 \left\{ \beta[\delta(1 - \lambda\chi) - 1] + z \left[1 - \beta \left(1 - \lambda\chi + \frac{\phi_y(1-\lambda)}{\sigma} \right) \right] \right\}}{(z - \lambda_1)(1 - \lambda_1 z) \rho \sigma_\epsilon^2}$$
$$C_{13}(z) = \frac{\beta(1 - \lambda)\sigma_\eta^2}{\sigma} \left(\phi_\pi - \frac{1}{z} \right)$$
$$C_{14}(z) = -\frac{\lambda_1 \sigma_\eta^4 \beta (1 - \lambda) (1 - \phi_\pi z)}{\sigma (z - \lambda_1)(1 - \lambda_1 z) \rho \sigma_\epsilon^2}$$
$$C_{21}(z) = 0$$
$$C_{22}(z) = 1 - \frac{\beta \sigma_\eta^2}{z} + C_{12}(z) \frac{\sigma_\epsilon^2}{\sigma_1^2}$$
$$C_{23}(z) = 0$$

$$C_{24}(z) = C_{14}(z)\frac{\sigma_\epsilon^2}{\sigma_1^2}$$

$$C_{31}(z) = -\sigma_\eta^2\kappa\theta$$

$$C_{32}(z) = -\frac{\lambda_2\sigma_\eta^4\kappa\theta z}{(z-\lambda_2)(1-\lambda_2 z)\rho\sigma_\epsilon^2}$$

$$C_{33}(z) = 1-\sigma_\eta^2\left[1-\theta\left(1-\frac{\beta}{z}\right)\right]$$

$$C_{34}(z) = -\frac{\lambda_2\sigma_\eta^4(1-\theta)z}{(z-\lambda_2)(1-\lambda_2 z)\rho\sigma_\epsilon^2}$$

$$C_{41}(z) = 0$$

$$C_{42}(z) = -\frac{\lambda_2\sigma_\eta^4\kappa\theta z}{(z-\lambda_2)(1-\lambda_2 z)\rho\sigma_2^2}$$

$$C_{43}(z) = 0$$

$$C_{44}(z) = 1-\sigma_\eta^2\left[\frac{\beta\theta}{z}+\frac{\lambda_2\sigma_\eta^2(1-\theta)z}{(z-\lambda_2)(1-\lambda_2 z)\rho\sigma_2^2}\right]$$

where λ_g, $g \in \{1,2\}$ *is the inside root of the following polynomial*

$$\boldsymbol{D}(z) \equiv z^2 - \left[\frac{1}{\rho}+\rho+\frac{(\sigma_g^2+\sigma_\epsilon^2)\sigma_\eta^2}{\rho\sigma_g^2\sigma_\epsilon^2}\right]z+1$$

Proof. See Appendix A. □

The first aspect to notice is that the equilibrium dynamics do not follow a VAR(1) process anymore, unless $Q_v = Q_\epsilon$ which is not generally satisfied. In this case the two exogenous shocks do not share the impact effect anymore, since agents can partly disentangle them through the two signals. Notice that by introducing an additional signal we are effectively reducing the degree of information frictions that agents face. Even if there is an exogenous shock to the common signal, private signals will be unaffected. As a result, agents will not fully react to the "animal spirits" shock. In fact, we find that the effect of the belief shock is smaller than before, and the monetary policy shock is more powerful and the produced dynamics are closer to the standard FIRE dynamics, as plotted in Figure 7.

To summarize, adding public information reduces information frictions, which in turn dampens the effect of any belief shock and enlarges the effect of monetary policy shocks.

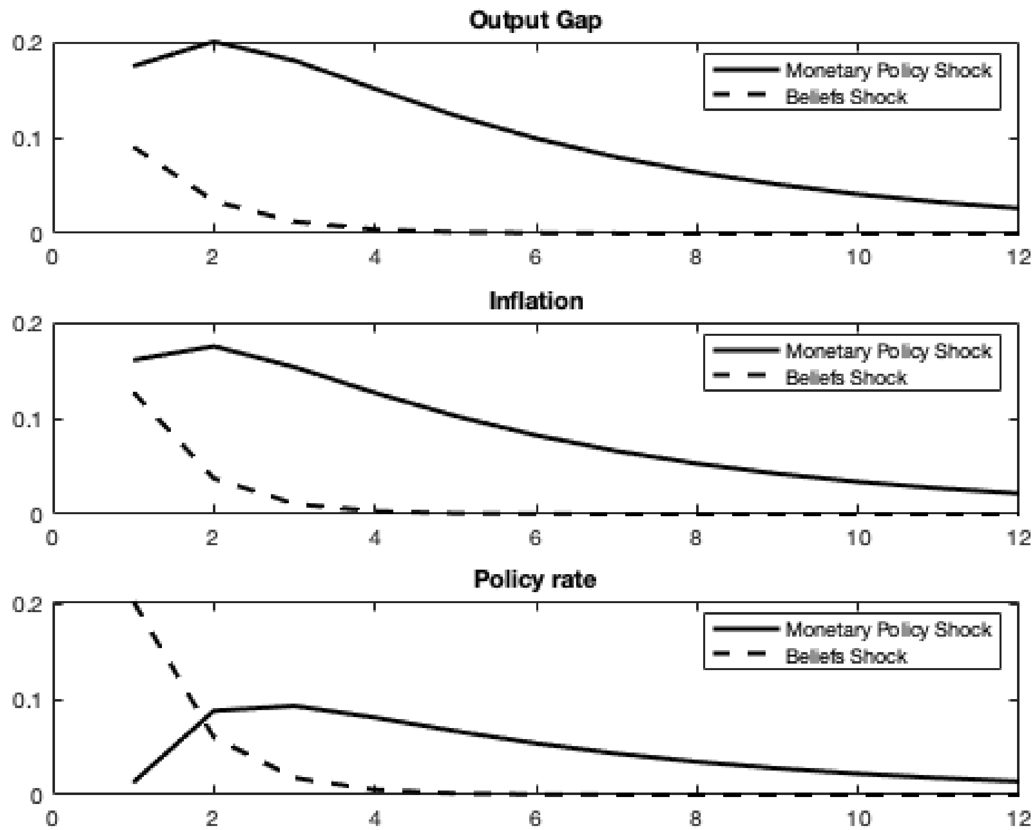


Figure 7: Impulse Responses of output and inflation after a monetary policy shock (dark line) and a shock to the common signal (dotted line).

5 Conclusion

The amplification result in the FIRE benchmark relies on financially constrained households being immediately affected after a monetary policy shock through the GE effects. We provide a new theory to explain the transmission channel of monetary policy in HANK economies. By introducing dispersed information, the GE considerations are dampened in the initial periods, reducing the magnitude of the multiplier.

We use our theory to shed light on other questions of first-order importance. We find that our framework produces hump-shaped IRFs without resorting to ad-hoc micro-inconsistent adjustment costs in habits, pricing or investment decisions. Instead, we microfound aggregate sluggishness using dispersed information and expectation formation sluggishness, for which we provide empirical evidence. This results in a different PE vs. GE role than in standard FIRE models, and is consistent with recent empirical evidence. We also show that dispersed information produces *as if* myopia, which extends the equilibrium determinacy region, and is crucial for the solution of the forward guidance puzzle.

Finally, we find that purely transitory "animal spirits" shocks can generate large and persistent effects in output and inflation.

References

Allen, Franklin, Stephen Morris, and Hyun Song Shin. 2006. "Beauty contests and iterated expectations in asset markets." *Review of Financial Studies.* ISSN: 08939454.

Almgren, Mattias, Jose Elias Gallegos, John Kramer, and Ricardo Lima. 2018. "A Proxy GVAR with High-Frequency Identification for Monetary Policy in the Euro Area."

Almgren, Mattias, José-Elías Gallegos, John Kramer, and Ricardo Lima. 2020. "Monetary Policy and Liquidity Constraints: Evidence from the Euro Area."

Andrade, Philippe, Gaetano Gaballo, Eric Mengus, and Benoît Mojon. 2019. "Forward guidance and heterogeneous beliefs." *American Economic Journal: Macroeconomics.* ISSN: 19457715.

Angeletos, George-Marios, and Zhen Huo. 2018. "Myopia and Anchoring." *NBER Working Paper Series:* 54.

Angeletos, George Marios, and Chen Lian. 2018. *Forward guidance without common knowledge.*

Angeletos, George-Marios, and Karthik A. Sastry. 2020. "Managing Expectations: Instruments vs. Targets."

Angeletos, George-marios, Zhen; Huo, and Karthik A. Sastry. 2020. "Imperfect Macroeconomic Expectations: Theory and Evidence." *Working Paper.*

Auclert, Adrien. 2019. "Monetary policy and the redistribution channel." *American Economic Review* 109 (6): 2333–2367. ISSN: 19447981.

Bacchetta, Philippe, and Eric Van Wincoop. 2006. "Can information heterogeneity explain the exchange rate determination puzzle?" *American Economic Review.* ISSN: 00028282.

Bilbiie, Florin O. 2008. "Limited asset markets participation, monetary policy and (inverted) aggregate demand logic." *Journal of Economic Theory.* ISSN: 10957235.

———. 2019a. "Monetary Policy and Heterogeneity: An Analytical Framework." *Mimeo.*

———. 2019b. "The New Keynesian cross." *Journal of Monetary Economics.* ISSN: 03043932.

Blanchard, Olivier Jean, and Charles M. Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations." *Econometrica.* ISSN: 00129682.

Bordalo, Pedro, Nicola Gennaioli, Yueran Ma, and Andrei Shleifer. 2018. *Over-reaction in Macroeconomic Expectations.* Technical report.

Brinca, Pedro, Hans A. Holter, Per Krusell, and Laurence Malafry. 2016. "Fiscal multipliers in the 21st century." *Journal of Monetary Economics* 77:53–69. ISSN: 03043932.

Broer, Tobias, and Alexandre N Kohlhas. 2019. "Forecaster (Mis-)Behavior."

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1): 1–45. ISSN: 0022-3808. arXiv: arXiv:1011.1669v3.

Coibion, Olivier, and Yuriy Gorodnichenko. 2015. "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review.* ISSN: 00028282.

Del Negro, Marco, Marc P. Giannoni, and Christina Patterson. 2012. "The Forward Guidance Puzzle." *SSRN Electronic Journal.*

Fagereng, Andreas, Martin Blomhoff Holm, Benjamin Moll, and Gisle Natvik. 2019. "Saving Behavior Across the Wealth Distribution: The Importance of Capital Gains." *NBER Working Paper.* ISSN: 1098-6596. arXiv: arXiv:1011.1669v3.

Farhi, Emmanuel, and Iván Werning. 2019. "Monetary policy, bounded rationality, and incomplete markets." *American Economic Review.* ISSN: 19447981.

Gabaix, Xavier. 2016. "A Behavioral New Keynesian Model." *Working Paper:* 1–58.

Galí, Jordi. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework,* 203. 1. ISBN: 9780874216561. arXiv: arXiv:1011.1669v3.

Galí, Jordi, J David Lopez-Salido, and Javier Valles. 2007. "Understanding the Effects of Government Spending on Consumption." *Journal of the European Economic Association* 5 (1): 227–270. ISSN: 1542-4774. arXiv: arXiv:1011.1669v3.

Gornemann, Nils, Keith Kuester, and Makoto Nakajima. 2016. *Doves for the Rich, Hawks for the Poor? Distributional Consequences of Monetary Policy.*

Groth, Charlotta, and Hashmat Khan. 2010. "Investment adjustment costs: An empirical assessment." *Journal of Money, Credit and Banking.* ISSN: 00222879.

Guvenen, Fatih, Serdar Ozkan, and Jae Song. 2014. "The nature of countercyclical income risk." *Journal of Political Economy.* ISSN: 1537534X.

Hagedorn, Marcus, Jinfeng Luo, Iourii Manovskii, and Kurt Mitman. 2019. "Forward guidance." *Journal of Monetary Economics.* ISSN: 03043932.

Hamilton, James Douglas. 1994. *Time Series Analysis.* 1–799. Princeton University Press. ISBN: 0-691-04289-6.

Havranek, Tomas, Marek Rusnak, and Anna Sokolova. 2017. "Habit formation in consumption: A meta-analysis." *European Economic Review.* ISSN: 00142921.

Holm, Martin Blomhoff, Pascal Paul, and Andreas Tischbirek. 2021. "The Transmission of Monetary Policy under the Microscope." *Journal of Political Economy* 0 (ja): null. eprint: https://doi.org/10.1086/715416.

Huo, Zhen, and Naoki Takayama. 2018. "Rational Expectations Models with Higher Order Beliefs."

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante. 2018. *Monetary policy according to HANK.* arXiv: arXiv:1011.1669v3.

Kaplan, Greg, Giovanni L. Violante, and Justin Weidner. 2014. "The Wealthy Hand-to-Mouth." *Brookings Papers on Economic Activity* 2014 (1): 77–138. ISSN: 1533-4465.

Lindé, Jesper. 2005. "Estimating New-Keynesian Phillips curves: A full information maximum likelihood approach." *Journal of Monetary Economics* 52 (6): 1135–1149. ISSN: 03043932.

Lorenzoni, Guido. 2009. "A theory of demand shocks." *American Economic Review* 99 (5): 2050–2084. ISSN: 00028282.

Lucas, Robert E. 1972. "Expectations and the neutrality of money." *Journal of Economic Theory.* ISSN: 10957235.

Mckay, Alisdair. 2015. "The Power of Forward Guidance Revisited." 106 (10): 3133–3158. ISSN: 0002-8282.

Morris, Stephen, and Hyun Song Shin. 2002. "Social value of public information." *American Economic Review.* ISSN: 00028282.

Morris, Stephen, Hyun Song Shin, Nobuhiro Kiyotaki, Marios Angeletos, and Andrew Atkeson. 2006. "Inertia of forward-looking expectations." In *American Economic Review.*

Nakamura, Emi, and Jón Steinsson. 2008. "Five Facts about Prices: A Reevaluation of Menu Cost Models." *Quarterly Journal of Economics* 123 (4): 1415–1464. ISSN: 0033-5533.

Nimark, Kristoffer. 2008. "Dynamic pricing and imperfect common knowledge." *Journal of Monetary Economics.* ISSN: 03043932.

Ramey, V.A. 2016. "Chapter 2 - Macroeconomic Shocks and Their Propagation," edited by John B. Taylor and Harald Uhlig, 2:71–162. Handbook of Macroeconomics. Elsevier.

Vives, Xavier, and Liyan Yang. 2016. "A Model of Costly Interpretation of Asset Prices." *SSRN Electronic Journal* (January).

Werning, Iván. 2015. "Incomplete Markets and Aggregate Demand." *National Bureau of Economic Research Working Paper Series.* ISSN: 0898-2937.

Woodford, M. 2003. "Imperfect Common Knowledge and the Effects of Monetary Policy." *Information and Expectations in Modern Macroeconomics.*

Appendix

A Proofs of Propositions in Main Text

Proof of Proposition 1. An unconstrained agent $i \in S$ chooses consumption, asset holdings and leisure solving the standard intertemporal problem: $\max E_{i0} \sum_{t=0}^{\infty} \beta^t U\left(C_{it}^S, N_{it}^S\right)$, subject to the sequence of constraints:

$$B_{it} + \Omega_{i,t+1} V_t \leq Z_{it} + \Omega_{it}\left(V_t + P_t D_t\right) + W_t N_{it}^S - P_t C_{it}^S \tag{A.1}$$

where C_{it}^S, N_{it}^S are consumption and hours worked, B_{it} is the nominal value at end of period t of a portfolio of all state-contingent assets held, except for shares in firms. Likewise for Z_{it}, beginning of period wealth. V_t is average market value at time t of shares, D_t their real dividend payoff and Ω_{it} are share holdings. Absence of arbitrage implies that there exists a stochastic discount factor $Q_{i,t,t+1}$ such that the price at t of a portfolio with uncertain payoff at $t+1$ is (for state-contingent assets and shares respectively, for an agent i who participates in those markets):

$$B_{it} = \mathbb{E}_{it}\left[Q_{i,t,t+1} Z_{i,t+1} \frac{P_t}{P_{t+1}}\right] \quad \text{and} \quad 1 = \mathbb{E}_{it}\left[Q_{i,t,t+1}\left(\frac{P_t}{P_{t+1}} \frac{V_{t+1}}{V_t} + \frac{P_t}{V_t} D_{t+1}\right)\right] \tag{A.2}$$

which iterated forward gives the fundamental pricing equation: $1 = \mathbb{E}_{it}\left[\frac{P_t}{V_t} \sum_{k=1}^{\infty} Q_{i,t,t+k} D_{t+k}\right]$. The riskless gross short-term real interest rate R_t is a solution to:

$$1 = \mathbb{E}_{it}\left(R_t Q_{i,t,t+1}\right)$$

Note that for nominal assets we have the nominal interest rate $1 = \mathbb{E}_{it}\left(\frac{P_t}{P_{t+1}} I_t Q_{i,t,t+1}\right)$. Substituting the no-arbitrage conditions (A.2) into the wealth dynamics equation (A.1) gives the flow budget constraint. Together with the usual no-borrowing limit for each state, and anticipating that in equilibrium all agents will hold a constant fraction of the shares (there is no trade in shares) Ω_i, whose integral across agents is 1, this implies the usual intertemporal budget constraint:

$$\mathbb{E}_{it}\left[\frac{P_t}{P_{t+1}} Q_{i,t,t+1} X_{i,t+1}\right] \leq \mathbb{E}_{it}\left[X_{it} + W_t N_{it}^S - P_t C_{it}^S\right]$$

where

$$
\begin{aligned}
\mathbb{E}_{it} X_{it} &= \mathbb{E}_{it}\left[Z_{it} + \Omega_i \left(V_t + P_t D_t\right)\right] \\
&= \mathbb{E}_{it}\left[Z_{it} + \Omega_i \left(\sum_{k=0}^{\infty} P_t Q_{i,t,t+k} D_{t+k}\right)\right]
\end{aligned}
$$

and

$$
\begin{aligned}
\mathbb{E}_{it} \sum_{k=0}^{\infty} Q_{i,t,t+k} C^S_{i,t+k} &\le \mathbb{E}_{it}\left[\frac{X_{it}}{P_t} + \sum_{k=0}^{\infty} Q_{i,t,t+k} \frac{W_{t+k}}{P_{t+k}} N^S_{i,t+k}\right] \\
&= \mathbb{E}_{it} \sum_{k=0}^{\infty} Q_{i,t,t+k} Y^S_{i,t+k} \qquad \text{(A.3)}
\end{aligned}
$$

where

$$
Y^S_{i,t+k} = \Omega_i D_{t+k} + \frac{W_{t+k}}{P_{t+k}} N^S_{i,t+k}
$$

is income of agent i. Maximizing utility subject to this constraint gives the first-order necessary and sufficient conditions at each date and in each state:

$$
\beta \frac{U_C\left(C_{i,t+1}\right)}{U_C\left(C_t\right)} = Q_{i,t,t+1}
$$

along with (A.3) holding with equality (or alternatively flow budget constraint holding with equality and transversality conditions ruling out Ponzi games be satisfied: $\lim_{k\to\infty} \mathbb{E}_{it}\left[Q_{i,t,t+k} Z_{i,t+k}\right] = \lim_{k\to\infty} \mathbb{E}_{it}\left[Q_{i,t,t+k} V_{t+k}\right] = 0$). Using (A.3) and the functional form of the utility function the short-term nominal interest rate must obey:

$$
1 = \beta \mathbb{E}_{it}\left[R_t \frac{U_C\left(C^S_{i,t+1}\right)}{U_C\left(C^S_{it}\right)}\right]
$$

Denote by small letter log deviations from steady-state, except for rates of return (where they denote absolute deviations). Notice that

$$
Q_{t,t+k} = \beta^k \frac{U_C\left(C^S_{i,t+k}\right)}{U_C\left(C^S_{it}\right)}
$$

and in steady state: $Q_k = \beta^k$. Thus we have

$$q_{i,t,t+k} = \ln \frac{Q^S_{i,t,t+k}}{Q^S_{ik}} = \ln \frac{U_C\left(C^S_{i,t+k}\right)}{U_C\left(C^S_{it}\right)} = -\sigma\left(c^S_{i,t+k} - c^S_{it}\right)$$

where

$$q_{i,t,t+k} = q_{i,t,t+1} + q_{i,t+1,t+2} + \ldots + q_{i,t+k-1,t+k}$$

Using the stochastic discount factor notation, we can write the unconstrained Euler condition as

$$\frac{1}{\sigma} q^S_{t,t+1} = c^S_{it} - s\mathbb{E}_{it} c^S_{i,t+1} - (1-s)\mathbb{E}_{it} c^H_{i,t+1}$$

Iterating forward the above condition,

$$c^S_{it} = s^k \mathbb{E}_{it} c^S_{t+k} - \sum_{j=0}^{k-1}\left[\frac{1}{\sigma}\mathbb{E}_{it} q^S_{t,t+1} + (1-s)\mathbb{E}_{it} c^H_{i,t+j}\right] \tag{A.4}$$

Using the definition of the stochastic discount factor, we can write

$$\frac{1}{\sigma} q^S_{t,t+k} = c^S_{it} - s\mathbb{E}_{it} c^S_{i,t+1} - (1-s)\mathbb{E}_{it} c^H_{i,t+1} + c^S_{i,t+1} - s\mathbb{E}_{it} c^S_{i,t+2} - (1-s)\mathbb{E}_{it} c^H_{i,t+2} + \ldots$$
$$+ c^S_{i,t+k-1} - s\mathbb{E}_{it} c^S_{i,t+k} - (1-s)\mathbb{E}_{it} c^H_{i,t+k}$$

and we can thus write

$$\frac{1}{\sigma}\mathbb{E}_{it} q^S_{t,t+k} = c^S_{it} + (1-s)\mathbb{E}_{it}\sum_{j=1}^{k}(c^S_{i,t+k} - c^H_{i,t+k})$$

Log-linearizing the intertemporal budget constraint around a steady-state with no shocks nor information frictions, zero profits and no inequality, $C^S = C^H$

$$\sum_{k=0}^{\infty} \beta^k c^S_{it+k} = \sum_{k=0}^{\infty} \beta^k y^S_{i,t+k} \tag{A.5}$$

Adding $\sigma^{-1}\mathbb{E}_{it}q^S_{t,t+k}$ on each side

$$\sum_{k=0}^{\infty}\beta^k\mathbb{E}\left[\frac{1}{\sigma}q^S_{t,t+k}+c^S_{it+k}\right]=\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}\left[\frac{1}{\sigma}q^S_{t,t+k}+y^S_{it+k}\right] \tag{A.6}$$

Using the iterated Euler condition (A.4), the LHS is reduced to

$$\frac{1}{1-\beta}c^S_{it}+\frac{1-s}{1-\beta}\sum_{k=1}^{\infty}\beta^k\mathbb{E}_{it}(c^S_{i,t+k}-c^H_{i,t+k})=\frac{1}{\sigma}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}q^S_{t,t+k}+\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}y^S_{it+k} \tag{A.7}$$

We can also write

$$\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}q^S_{t,t+k}=-\sum_{k=1}^{\infty}\beta^k\sum_{j=0}^{k-1}\mathbb{E}_{it}r_{t+k}=-\frac{\beta}{1-\beta}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}r_{t+k}$$

Hence, we can write the consumption policy function as

$$c^S_{it}=-(1-s)\sum_{k=1}^{\infty}\beta^k\mathbb{E}_{it}(c^S_{i,t+k}-c^H_{i,t+k})-\frac{\beta}{\sigma}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}r_{t+k}+(1-\beta)\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}y^S_{it+k} \tag{A.8}$$

We assume that the government implements an optimal steady-state subsidy such that there are zero profits and perfect consumption insurance in steady state, $\tau^S=(\epsilon-1)^{-1}$, and that the government implements a redistribution scheme by taxing profits, τ. Log-linearizing the budget constraints

$$c^S_{it}=w_t+n^S_{it}+\frac{1-\tau}{1-\lambda}e_t=y^S_{it} \tag{A.9}$$

$$c^H_{it}=w_t+n^H_{it}+\frac{\tau}{\lambda}e_t=y^H_{it} \tag{A.10}$$

Using the intratemporal labor supply conditions

$$\mathbb{E}_{it}w^r_t=\sigma c^S_{it}+\varphi n^S_{it} \tag{A.11}$$

$$\mathbb{E}_{it}w^r_t=\sigma c^H_{it}+\varphi n^H_{it} \tag{A.12}$$

Combining (A.9)-(A.12), we can write

$$c^S_{it}=\frac{1+\varphi}{\varphi+\sigma}\mathbb{E}_{it}w_t+\frac{\varphi}{\varphi+\sigma}\frac{1-\tau}{1-\lambda}\mathbb{E}_{it}e_t \tag{A.13}$$

$$c_{it}^H = \frac{1+\varphi}{\varphi+\sigma}\mathbb{E}_{it} w_t + \frac{\varphi}{\varphi+\sigma}\frac{\tau}{\lambda}\mathbb{E}_{it} e_t \tag{A.14}$$

Hence, we can rewrite the consumption function (A.8) as

$$\begin{aligned} c_{it}^S = &-(1-s)\sum_{k=1}^{\infty}\beta^k\left[\frac{\varphi}{\varphi+\sigma}\left(\frac{1-\tau}{1-\lambda}-\frac{\tau}{\lambda}\right)\mathbb{E}_{it} e_{t+k}\right] - \frac{\beta}{\sigma}\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it} r_{t+k} \\ &+ (1-\beta)\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}\left[\frac{1+\varphi}{\varphi+\sigma}\mathbb{E}_{it} w_{t+k} + \frac{\varphi}{\varphi+\sigma}\frac{1-\tau}{1-\lambda}\mathbb{E}_{it} e_{t+k}\right] \end{aligned} \tag{A.15}$$

Aggregating across $i \in S$ agents, we can write

$$\begin{aligned} c_{t}^S = &-(1-s)\sum_{k=1}^{\infty}\beta^k\left[\frac{\varphi}{\varphi+\sigma}\left(\frac{1-\tau}{1-\lambda}-\frac{\tau}{\lambda}\right)\overline{\mathbb{E}}_{t} e_{t+k}\right] - \frac{\beta}{\sigma}\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_{t} r_{t+k} \\ &+ (1-\beta)\sum_{k=0}^{\infty}\beta^k\mathbb{E}_{it}\left[\frac{1+\varphi}{\varphi+\sigma}\overline{\mathbb{E}}_{t} w_{t+k} + \frac{\varphi}{\varphi+\sigma}\frac{1-\tau}{1-\lambda}\overline{\mathbb{E}}_{t} e_{t+k}\right] \end{aligned} \tag{A.16}$$

□

Proof of Proposition 2. The First-Order Condition is

$$\sum_{k=0}^{\infty}\theta^k\mathbb{E}_{jt}\left[\Lambda_{t,t+k}Y_{j,t+k|t}\frac{1}{P_{t+k}}\left(P_{jt}^* - \mathcal{M}\Psi_{j,t+k|t}\right)\right] = 0$$

where $\Psi_{j,t+k|t} \equiv \mathcal{C}'_{t+k}(Y_{j,t+j|t})$ denotes the (nominal) marginal cost for firm j, and $\mathcal{M} = \frac{\epsilon}{\epsilon-1}$. Log-linearizing around the zero inflation steady-state, we obtain the familiar price-setting rule

$$p_{jt}^* = (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k\mathbb{E}_{jt}\left(\psi_{j,t+k|t}+\mu\right) \tag{A.17}$$

where $\psi_{j,t+k|t} = \log\Psi_{j,t+k|t}$ and $\mu = \log\mathcal{M}$.

Market clearing in the goods market implies that $Y_{jt} = C_{jt} = \int_{\mathcal{I}_h} C_{ijt}\, di$ for each j good/firm. Aggregating across firms, we obtain the aggregate market clearing condition: since assets are in zero net supply and there is no capital, investment, government consumption nor net exports, production equals consumption:

$$\int_{\mathcal{I}_f} Y_{jt}\, dj = \int_{\mathcal{I}_h}\int_{\mathcal{I}_f} C_{ijt}\, dj\, di \implies Y_t = C_t$$

Aggregate employment is given by the sum of employment across firms, and must meet aggregate labor supply

$$N_t = \int_{\mathcal{I}_h} N_{it}\, di = \int_{\mathcal{I}_f} N_{jt}\, dj$$

Using the production function and consumption demand, together with goods market clearing

$$N_t = \int_{\mathcal{I}_f} Y_{jt}\, dj = Y_t \int_{\mathcal{I}_f} \left(\frac{P_{jt}}{P_t}\right)^{-\epsilon} dj$$

Log-linearizing the above expression yields to

$$n_t = y_t \tag{A.18}$$

The (log) marginal cost for firm j at time $t+k|t$ is

$$\begin{aligned}\psi_{j,t+k|t} &= w_{t+k} - mpn_{j,t+k|t} \\ &= w_{t+k}\end{aligned}$$

where $mpn_{j,t+k|t}$ and $n_{j,t+k|t}$ denote (log) marginal product of labor and (log) employment in period $t+k$ for a firm that last reset its price at time t, respectively.

Let $\psi_t \equiv \int_{\mathcal{I}_f} \psi_{jt}$ denote the (log) average marginal cost. We can then write

$$\psi_t = w_t$$

Thus, the following relation holds

$$\psi_{j,t+k|t} = \psi_{t+k} \tag{A.19}$$

Introducing (A.19) into (A.17), we can rewrite the firm price-setting condition as

$$p^*_{jt} = (1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\left(p_{t+k} - \hat{\mu}_{t+k}\right)$$

where $\hat{\mu} = \mu_t - \mu$ is the deviation between the average and desired markups, where $\mu_t = -(\psi_t - p_t)$.

Suppose that firms observe the aggregate prices up to period $t-1$, p^{t-1}, then we can restate the above condition as

$$p_{jt}^* - p_{t-1} = -(1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\hat{\mu}_{t+k} + \sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\pi_{t+k}$$

Define the firm-specific inflation rate as $\pi_{jt} = (1-\theta)(p_{jt}^* - p_{t-1})$. Then we can write the above expression as

$$\begin{aligned}
\pi_{jt} &= -(1-\theta)(1-\beta\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\hat{\mu}_{t+k} + (1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{jt}\pi_{t+k} \\
&= (1-\theta)\mathbb{E}_{jt}[\pi_t - (1-\beta\theta)\hat{\mu}_t] + \beta\theta\mathbb{E}_{jt}\left\{(1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k[\pi_{t+1+k} - (1-\beta\theta)\hat{\mu}_{t+1+k}]\right\} \\
&= (1-\theta)\mathbb{E}_{jt}[\pi_t - (1-\beta\theta)\hat{\mu}_t] + \beta\theta\mathbb{E}_{jt}\left\{(1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k \mathbb{E}_{j,t+1}[\pi_{t+1+k} - (1-\beta\theta)\hat{\mu}_{t+1+k}]\right\} \\
&= -(1-\theta)(1-\beta\theta)\mathbb{E}_{jt}\hat{\mu}_t + (1-\theta)\mathbb{E}_{jt}\pi_t + \beta\theta\mathbb{E}_{jt}\pi_{j,t+1} \qquad \text{(A.20)}
\end{aligned}$$

where $\pi_t = \int_{\mathcal{I}_f} \pi_{jt}\, dj$.

Note that we can write the deviation between average and desired markups as

$$\begin{aligned}
\mu_t &= p_t - \psi_t \\
&= p_t - w_t \\
&= -(\sigma y_t + \varphi n_t) \\
&= -\left(\sigma + \varphi\right) y_t
\end{aligned}$$

As in the benchmark model, under flexible prices ($\theta = 0$) the average markup is constant and equal to the desired μ. Consider the natural level of output, y_t^n as the equilibrium level under flexible prices and full-information rational expectations. Rewriting the above condition under the natural equilibrium,

$$\mu = -\left(\sigma + \varphi\right) y_t^n$$

which we can write as

$$y_t^n = \psi_y$$

where $\psi_y = -\frac{\mu}{\sigma+\varphi}$. Therefore, we can write

$$\hat{\mu}_t = -(\sigma+\varphi)\,\tilde{y}_t$$

where $\tilde{y}_t = y_t - y_t^n$ is defined as the output gap. Finally, we can write the individual Phillips curve as

$$\begin{aligned}\pi_{jt} &= (1-\theta)(1-\beta\theta)\,(\sigma+\varphi)\,\mathbb{E}_{jt}\tilde{y}_t + (1-\theta)\mathbb{E}_{jt}\pi_t + \beta\theta\mathbb{E}_{jt}\pi_{i,t+1} \\ &= \kappa\theta\mathbb{E}_{jt}\tilde{y}_t + (1-\theta)\mathbb{E}_{jt}\pi_t + \beta\theta\mathbb{E}_{jt}\pi_{i,t+1} \end{aligned} \tag{A.21}$$

where $\kappa = \frac{(1-\theta)(1-\beta\theta)}{\theta}\,(\sigma+\varphi)$, and the aggregate Phillips curve can be written as

$$\pi_t = \kappa\theta\sum_{k=0}^{\infty}(\beta\theta)^k\overline{\mathbb{E}}_t^f\tilde{y}_{t+k} + (1-\theta)\sum_{k=0}^{\infty}(\beta\theta)^k\overline{\mathbb{E}}_t^f\pi_{t+k} \tag{A.22}$$

□

Proof of Proposition 3. Denote aggregate consumption and aggregate labor supply for the unconstrained household as

$$C_t^S = \int C_{it}^S\,di, \quad N_t^S = \int N_{it}^S\,di$$

and aggregate consumption and aggregate labor supply for the constrained household given by

$$C_t^H = \int C_{it}^H\,di, \quad N_t^H = \int N_{it}^H\,di$$

Equilibrium in the goods market requires that consumption of unconstrained and constrained households equals total consumption

$$C_t = \lambda C_t^H + (1-\lambda)C_t^S$$

Since we are in a closed economy without investment and government spending, the resource constraint is $Y_t = C_t$. Equilibrium in the labor market requires

$$N_t = \lambda N_t^H + (1-\lambda) N_t^S$$

With uniform stady-state hours by normalization ($N^S = N^H = N$), and the fiscal policy inducing $C^S = C^H = C$, the above log-linearized market clearing conditions yields

$$y_t = c_t = \lambda c_t^H + (1-\lambda) c_t^S \tag{A.23}$$

$$n_t = \lambda n_t^H + (1-\lambda) n_t^S \tag{A.24}$$

Finally, because the final good sector is competitive and observes all relevant prices p_{jt}, we have

$$p_t = \int p_{jt}\, dj$$

$$y_t = \int y_{jt}\, dj = \int n_{jt}\, dj$$

$$y_t = n_t = \int n_{it}\, di \tag{A.25}$$

$$y_t = c_t = \int c_{it}\, di \tag{A.26}$$

Combining the (expectation augmented) optimal labor supply condition of unconstrained households (A.11) and that of constrained households (A.12), and the labor and goods market clearing conditions (A.23)-(A.24), we can write

$$\begin{aligned} \overline{\mathbb{E}}_t^c w_t &= \sigma \overline{\mathbb{E}}_t^c c_t + \varphi \overline{\mathbb{E}}_t^c n_t \\ &= (\varphi + \sigma) \overline{\mathbb{E}}_t^c y_t \end{aligned} \tag{A.27}$$

where we have used the aggregate market clearing condition in goods and labor sectors. As is common in NK models without nominal wage rigidities, profits are countercyclical. This results in dividends (and transfers received by firms) being countercyclical. Using the fact that $e_t = -w_t$, we can write (A.14) as

$$\begin{aligned} c_t^H &= \frac{1}{\varphi + \sigma} \left[1 + \varphi \left(1 - \frac{\tau^D}{\lambda} \right) \right] \overline{\mathbb{E}}_t^c w_t \\ &= \left[1 + \varphi \left(1 - \frac{\tau^D}{\lambda} \right) \right] \overline{\mathbb{E}}_t^c y_t \end{aligned}$$

$$= \chi \overline{\mathbb{E}}_t^c y_t$$

Hence, we can finally write the aggregate consumption function as

$$\begin{aligned} c_t &= (1-\lambda)c_t^S + \lambda c_t^H \\ &= -\frac{\beta}{\sigma}(1-\lambda)\sum_{k=0}^{\infty}\beta^k \overline{\mathbb{E}}_t^c r_{t+k} + [1-\beta(1-\lambda\chi)]\overline{\mathbb{E}}_t^c y_t + (\delta-\beta)(1-\lambda\chi)\sum_{k=1}^{\infty}\beta^k \overline{\mathbb{E}}_t^c c_{t+k} \end{aligned} \tag{A.28}$$

where $\delta = 1 + \frac{(\chi-1)(1-s)}{1-\lambda\chi}$ and $\varsigma = \sigma\frac{1-\lambda\chi}{1-\lambda}$. Finally, notice that this is implied by the following beauty-contest game for a representative household i,

$$c_{it} = -\frac{\beta}{\sigma}(1-\lambda)\mathbb{E}_{it}r_t + [1-\beta(1-\lambda\chi)]\mathbb{E}_{it}y_t + \beta[\delta(1-\lambda\chi)-1]\mathbb{E}_t c_{t+1} + \beta\mathbb{E}_{it}c_{i,t+1}$$

is equivalent to (A.28) provided that $\lim_{T\to\infty}\beta^T\mathbb{E}_{it}c_{i,t+T}$, which is broadly assumed in the literature given $\beta < 1$.

Replacing the real interest rate by the Taylor rule process in the DIS curve (A.28), $r_t = i_t - \pi_{t+1} = \phi_\pi\pi_t + \phi_y y_t + v_t - \pi_{t+1}$, we obtain

$$\begin{aligned} c_{it} &= \left[1-\beta\left(1-\lambda\chi+\frac{\phi_y}{\sigma}(1-\lambda)\right)\right]\mathbb{E}_t y_t - \frac{\beta\phi_\pi}{\sigma}(1-\lambda)\mathbb{E}_{it}\pi_t - \frac{\beta}{\varsigma}(1-\lambda)\mathbb{E}_{it}v_t \\ &\quad + \beta[\delta(1-\lambda\chi)-1]\mathbb{E}_{it}c_{t+1} + \frac{\beta}{\varsigma}(1-\lambda)\mathbb{E}_{it}\pi_{t+1} + \beta\mathbb{E}_{it}c_{i,t+1} \end{aligned}$$

□

Proof of Proposition 4. The final goal is to write the DIS curve and the NKPC (2.13)-(2.9) in a similar way to Huo and Takayama (2018). That is,

$$\begin{aligned} a_{i1t} &= \varphi_1\mathbb{E}_{i1t}\xi_t + \beta_1\mathbb{E}_{i1t}a_{i1t+1} + \gamma_{11}\mathbb{E}_{i1t}a_{1t} + \alpha_{11}\mathbb{E}_{i1t}a_{1t+1} + \gamma_{12}\mathbb{E}_{i1t}a_{2t} + \alpha_{12}\mathbb{E}_{i1t}a_{2t+1} \qquad \text{(A.29)} \\ a_{j2t} &= \varphi_2\mathbb{E}_{j2t}\xi_t + \beta_2\mathbb{E}_{j2t}a_{j2t+1} + \gamma_{21}\mathbb{E}_{j2t}a_{1t} + \alpha_{21}\mathbb{E}_{j2t}a_{1t+1} + \gamma_{22}\mathbb{E}_{j2t}a_{2t} + \alpha_{22}\mathbb{E}_{j2t}a_{2t+1} \end{aligned}$$

To show this, let me focus on (A.29). Iterating forward

$$a_{i1t} = \varphi_1\sum_{k=0}^{\infty}\beta_1^k\mathbb{E}_{i1t}\xi_{t+k} + \gamma_{11}\mathbb{E}_{i1t}a_{1t} + (\beta_1\gamma_{11}+\alpha_{11})\sum_{k=0}^{\infty}\beta_1^k\mathbb{E}_{i1t}a_{1t+k+1} + \gamma_{12}\mathbb{E}_{i1t}a_{2t} + (\beta_1\gamma_{12}+\alpha_{12})\sum_{k=0}^{\infty}\beta_1^k\mathbb{E}_{i1t}a_{2t+k+1}$$

And the aggregate action for households is

$$a_{1t} = \varphi_1 \sum_{k=0}^{\infty} \beta_1^k \overline{\mathbb{E}}_{1t}\xi_{t+k} + \gamma_{11}\overline{\mathbb{E}}_{1t}a_{1t} + (\beta_1\gamma_{11} + \alpha_{11}) \sum_{k=0}^{\infty} \beta_1^k \overline{\mathbb{E}}_{1t}a_{1t+k+1} + \gamma_{12}\overline{\mathbb{E}}_{1t}a_{2t} + (\beta_1\gamma_{12} + \alpha_{12}) \sum_{k=0}^{\infty} \beta_1^k \overline{\mathbb{E}}_{1t}a_{2t+k+1} \tag{A.30}$$

In a similar way, we can derive the aggregate action for firms

$$a_{2t} = \varphi_2 \sum_{k=0}^{\infty} \beta_2^k \overline{\mathbb{E}}_{2t}\xi_{t+k} + \gamma_{21}\overline{\mathbb{E}}_{2t}a_{1t} + (\beta_2\gamma_{21} + \alpha_{21}) \sum_{k=0}^{\infty} \beta_2^k \overline{\mathbb{E}}_{2t}a_{1t+k+1} + \gamma_{22}\overline{\mathbb{E}}_{2t}a_{2t} + (\beta_2\gamma_{22} + \alpha_{22})) \sum_{k=0}^{\infty} \beta_2^k \overline{\mathbb{E}}_{2t}a_{2t+k+1} \tag{A.31}$$

Notice that (A.30)–(A.31) are equivalent to (2.13) and (2.9), respectively, if $a_{1t} = y_t$, $a_{2t} = \pi_t$, $\xi_t = v_t$, $\overline{\mathbb{E}}_{1t}(\cdot) = \overline{\mathbb{E}}_{ct}(\cdot)$, $\overline{\mathbb{E}}_{2t}(\cdot) = \overline{\mathbb{E}}_{\pi t}(\cdot)$ and the following parametric restrictions are satisfied

$$\begin{aligned}
\varphi_1 &= -\frac{\beta(1-\lambda)}{\sigma} \\
\beta_1 &= \beta \\
\gamma_{11} &= 1 - \beta\left[1 - \lambda\chi + \frac{\phi_y}{\sigma}(1-\lambda)\right] \\
\gamma_{12} &= -\beta(1-\lambda)\frac{\phi_\pi}{\sigma} \\
\alpha_{11} &= \beta[\delta(1-\lambda\chi) - 1] \\
\alpha_{12} &= \frac{\beta}{\sigma}(1-\lambda)
\end{aligned}
\qquad
\begin{aligned}
\varphi_2 &= 0 \\
\beta_2 &= \beta\theta \\
\gamma_{21} &= \kappa\theta \\
\gamma_{22} &= 1-\theta \\
\alpha_{21} &= 0 \\
\alpha_{22} &= 0
\end{aligned}$$

The best response of agent i in group g is specified as follows

$$a_{igt} = \varphi_g \mathbb{E}_{igt}\xi_t + \beta_g \mathbb{E}_{igt}a_{igt+1} + \sum_{j=1}^{2} \gamma_{gj}\mathbb{E}_{igt}a_{jt} + \sum_{j=1}^{2} \alpha_{gj}\mathbb{E}_{igt}a_{jt+1} \tag{A.32}$$

where a_{-gt} is the aggregate action of the other group at time t. Parameters $\{\beta_g\}$, $\{\gamma_{gk}\}$, $\{\alpha_{gk}\}$ help parameterize PE and GE considerations. Notice that GE effects run not only within groups but also across groups (the interaction of the two blocks of the NK model). Parameters $\{\varphi_g\}$ capture the direct exposure of group g to the exogenous shock.[23]

23. The parameter φ_g is allowed to be zero for some, but not all, g. For instance, if ξ_t represents a monetary policy shock, it shows up in the Dynamic IS curve but not in the NKPC (and the converse would be true

Let $\boldsymbol{a}_t = (a_{gt})$ be a column vector collecting the aggregate actions of all groups (e.g., the vector of aggregate consumption and aggregate inflation)

$$\boldsymbol{a}_t = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Let $\boldsymbol{\varphi} = (\varphi_g)$ be a column vector containing the value of φ_g across groups

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix}$$

Let $\boldsymbol{\beta} = \text{diag}(\beta_g)$ be a 2×2 diagonal matrix of discount factors, with off-diagonal elements equal to 0.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}$$

Let $\boldsymbol{\gamma}$ be a 2×2 matrix collecting the (contemporaneous) interaction parameters γ_{gj}

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$$

Let $\boldsymbol{\alpha} = (\alpha_{gk})$ be a 2×2 matrix collecting the (future) interaction parameters α_{gj}

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$$

Finally, let $\boldsymbol{\delta} \equiv \boldsymbol{\beta} + \boldsymbol{\alpha}$,

$$\boldsymbol{\delta} = \begin{bmatrix} \beta_1 + \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \beta_2 + \alpha_{22} \end{bmatrix}$$

Let us now have a look at the fundamental representation of the signal process. We know that

$$\begin{aligned} \xi_t &= \rho \xi_{t-1} + \eta_t \\ &= \frac{1}{1-\rho L} \eta_t, & \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \\ x_{igt} &= \xi_t + u_{igt}, & u_{igt} \sim \mathcal{N}(0, \sigma_g^2) \end{aligned}$$

for a markup shock, for example, which we do not study in this paper).

Notice that the signal process admits the following state-space representation

$$Z_t = FZ_{t-1} + \Phi \widehat{\boldsymbol{s}}_{igt}$$
$$X_t = HZ_t + \Psi \widehat{\boldsymbol{s}}_{igt}$$

with $F = \rho$, $\Phi = \begin{bmatrix} \sigma_\eta & 0 \end{bmatrix}$, $Z_t = \xi_t$, $H = 1$, $\Psi = \begin{bmatrix} 0 & \sigma_g \end{bmatrix}$ and $X_t = x_{igt}$. Define $\tau_\eta \equiv \frac{1}{\sigma_\eta^2}$, $\tau_g \equiv \frac{1}{\sigma_g^2}$. The signal system can be written as

$$\begin{aligned} x_{igt} &= \frac{1}{1-\rho L}\eta_t + u_{igt} \\ &= \frac{\sigma_\eta}{1-\rho L}\widehat{\eta}_t + \sigma_g \widehat{u}_{igt} \\ &= \begin{bmatrix} \frac{\sigma_\eta}{1-\rho L} & \sigma_g \end{bmatrix} \begin{bmatrix} \widehat{\eta}_t \\ \widehat{u}_{igt} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L} & \tau_g^{-1/2} \end{bmatrix} \begin{bmatrix} \widehat{\eta}_t \\ \widehat{u}_{igt} \end{bmatrix} \\ &= \boldsymbol{M}_g(L)\widehat{\boldsymbol{s}}_{igt}, \qquad \widehat{\boldsymbol{s}}_{igt} \sim \mathcal{N}(0, I) \end{aligned}$$

Suppose that there exists $\{B_g(L), w_{gt}\}$ such that

$$x_{igt} = \boldsymbol{M}_g(L)\widehat{\boldsymbol{s}}_{igt} = B_g(L)w_{igt} \tag{A.33}$$

with $B_g(L)$ invertible, w_{igt} serially uncorrelated and $w_t \sim (0, V)$. Then $x_{igt} = B_g(L)w_{igt}$ is a fundamental representation of x_{igt}. Since $B_g(L)$ is invertible, x_{ig}^t and w_{ig}^t contain the same information. (A.33) implies that both processes share the same autocorrelation function

$$\begin{aligned} G_x^g(z) \equiv \rho_{xx}^g(z) &= \boldsymbol{M}_g(z)\boldsymbol{M}_g'(z^{-1}) \\ &= B_g(z)V_g B'(z^{-1}) \end{aligned}$$

By Propositions 13.1-13.4 in Hamilton (1994),

$$B_g(L) = I + H(I - FL)^{-1}FK_g$$
$$V_g = HP_gH' + \Psi_g\Psi_g'$$

We need to find

$$P_g = F[P_g - P_g H'(HP_g H' + \Psi_g \Psi_g')^{-1} HP_g]F + \Phi\Phi' \tag{A.34}$$

$$K_g = P_g H'(HP_g H' + \Psi_g \Psi_g')^{-1} \tag{A.35}$$

We can write (A.34) as

$$\begin{aligned} P_g =& \rho[P_g - P_g(P_g + \sigma_g^2)^{-1} P_g]\rho + \sigma_\eta^2 \\ &\implies P_g^2 + P_g[(1-\rho^2)\sigma_g^2 - \sigma_\eta^2] = \sigma_\eta^2 \sigma_g^2 \end{aligned} \tag{A.36}$$

Denote $\kappa_g = P_g^{-1}$. Then we can rewrite (A.36) as

$$\begin{aligned} \sigma_g^2 \sigma_\eta^2 \kappa_g^2 &= [(1-\rho^2)\sigma_g^2 - \sigma_\eta^2]\kappa_g + 1 \\ \implies \kappa_g &= \frac{\tau_\eta}{2}\left\{1 - \rho^2 - \frac{\tau_g}{\tau_\eta} \pm \sqrt{\left[\frac{\tau_g}{\tau_\eta} - (1-\rho^2)\right]^2 + 4\frac{\tau_g}{\tau_\eta}}\right\} \end{aligned}$$

We also need to find K_g. From (A.35)

$$K_g = P_g(P_g + \sigma_g^2)^{-1} = \frac{1}{1 + \kappa_g \sigma_g^2}$$

Using the Kalman filter, the forecast of the fundamental is given by

$$\begin{aligned} \mathbb{E}_{igt}\xi_t &= (I - K_g H)F\mathbb{E}_{igt-1}\xi_{t-1} + K_g x_{igt} \\ &= \lambda_g \mathbb{E}_{igt-1}\xi_{t-1} + K_g x_{igt} \end{aligned}$$

where

$$\begin{aligned} \lambda_g &= (I - K_g H)F \\ &= \frac{\kappa_g \sigma_g^2 \rho}{1 + \kappa_g \sigma_g^2} \\ &= \frac{1}{2}\left[\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\eta} \mp \sqrt{\left(\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\eta}\right)^2 - 4}\right] \end{aligned} \tag{A.37}$$

where we choose the $\lambda_{1,2}$ root that lies inside the unit circle (the one with the '$-$' sign). We can also write

$$
\begin{aligned}
V_g &= \kappa_g^{-1} + \sigma_g^2 \\
&= \frac{\rho}{\lambda_g \tau_g}
\end{aligned}
$$

where I have used the identity $\kappa_g = {}^{\lambda_g \tau_g}/_{(\rho - \lambda_g)}$. Finally, we can obtain $B_g(L)$

$$
\begin{aligned}
B_g(L) &= 1 + \frac{\rho L}{(1-\rho L)(1+\kappa_g \sigma_g^2)} \\
&= \frac{1 - \lambda_g L}{1 - \rho L}
\end{aligned}
$$

z and therefore one can verify that

$$
B_g(z) V_g B_g'(z^{-1}) = M_g(z) M_g'(z^{-1})
$$

with

$$
\begin{aligned}
B_g(z) V_g B_g'(z^{-1}) &= \frac{\rho}{\lambda_g \tau_g} \frac{(1-\lambda_g L)(L - \lambda_g)}{(1-\rho L)(L-\rho)} \\
M_g(z) M_g'(z^{-1}) &= \frac{\tau_\eta L}{(1-\rho L)(L-\rho)} + \tau_g
\end{aligned}
$$

Finally, we can write

$$
\boldsymbol{M}_g'(L^{-1}) B'(L^{-1})^{-1} = \frac{G(L)}{\prod_{\tau=1}^{u}(L - \lambda_\tau)}
$$

Let us now move to the forecasting part. The forecast of a random variable f_t

$$
f_t = A(L) \widehat{\boldsymbol{s}}_t
$$

can be obtained using the Wiener-Hopf prediction filter[24]

$$
\mathbb{E}_{it} f_t = \left[A(L) M'(L^{-1}) B(L^{-1})^{-1} \right]_+ B(L)^{-1} x_{it}
$$

24. To explain the $[\,\cdot\,]_+$ operator, let us give an example. Suppose $g(z)$ is a rational function of z that does not contain negative powers of z in expansion, and all its roots lie inside the unit circle ($|\xi_j| < 1\ \forall j$). Then

$$
\left[\frac{g(z)}{(z-\xi_1)(z-\xi_2)\ldots(z-\xi_l)} \right]_+ = \frac{g(z)}{(z-\xi_1)(z-\xi_2)\ldots(z-\xi_l)} - \sum_{k=1}^{l} \frac{g(\xi_k)}{(z-\xi_k)\prod_{\tau \neq k}^{l}(\xi_k - \xi_\tau)}
$$

Based on this result, we can solve the model. Denote agent i in group g policy function

$$a_{igt} = h_g(L)x_{igt}$$

(in this model, agents only observe signals. As a result, the policy function can only depend on current and past signals). The aggregate outcome in group g can then be expressed as follows

$$\begin{aligned}
a_{gt} &= \int a_{igt}\, di \\
&= \int h_g(L)x_{igt}\, di \\
&= \int h_g(L)\left(\frac{\sigma_\eta}{1-\rho L}\widehat{\eta}_t + \sigma_g \widehat{u}_{igt}\right)\, di \\
&= h_g(L)\frac{\sigma_\eta}{1-\rho L}\widehat{\eta}_t
\end{aligned}$$

Let us now obtain the forecast of the fundamental ξ_t. We can write the fundamental as

$$\begin{aligned}
\xi_t &= \frac{\sigma_\eta}{1-\rho L}\widehat{\eta}_t \\
&= \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \begin{bmatrix} \widehat{\eta}_t \\ \widehat{u}_{igt} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \widehat{\boldsymbol{s}}_{igt}
\end{aligned}$$

Consider now the forecast of the own and average contemporaneous and future actions. Using the guess that $a_{igt+1} = h_g(L)x_{igt+1}$ and $a_{gt+1} = h_g(L)\xi_{t+1}$, we have

$$\begin{aligned}
a_{gt+1} &= h_g(L)\xi_{t+1} \\
&= \frac{h_g(L)}{L}\xi_t \\
&= \frac{h_g(L)}{L}\begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \widehat{\boldsymbol{s}}_{igt} \\
&= \begin{bmatrix} \tau_\eta^{-1/2}\frac{h_g(L)}{L(1-\rho L)} & 0 \end{bmatrix} \widehat{\boldsymbol{s}}_{igt} \\
a_{gt} &= h_g(L)\xi_t \\
&= h_g(L)\begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L} & 0 \end{bmatrix} \widehat{\boldsymbol{s}}_{igt} \\
&= \begin{bmatrix} \tau_\eta^{-1/2}\frac{h_g(L)}{1-\rho L} & 0 \end{bmatrix} \widehat{\boldsymbol{s}}_{igt}
\end{aligned}$$

$$
\begin{aligned}
a_{igt+1} &= h_g(L)x_{igt+1} \\
&= \frac{h_g(L)}{L}x_{igt} \\
&= \frac{h_g(L)}{L}\boldsymbol{M}_g\widehat{\boldsymbol{s}}_{igt} \\
&= \frac{h_g(L)}{L}\begin{bmatrix}\frac{\tau_\eta^{-1/2}}{1-\rho L} & \tau_g^{-1/2}\end{bmatrix}\widehat{\boldsymbol{s}}_{igt} \\
&= \begin{bmatrix}\tau_\eta^{-1/2}\frac{h_g(L)}{L(1-\rho L)} & \tau_g^{-1/2}\frac{h_g(L)}{L}\end{bmatrix}\widehat{\boldsymbol{s}}_{igt} \\
a_{igt+1} - a_{gt+1} &= \begin{bmatrix}0 & \tau_g^{-1/2}\frac{h_g(L)}{L}\end{bmatrix}\widehat{\boldsymbol{s}}_{igt}
\end{aligned}
$$

Let us now obtain the forecasts. Recall that, for

$$
f_t = A(L)\widehat{\boldsymbol{s}}_t = \frac{a(L)}{\prod_{\tau=1}^{d}(L-\beta_\tau)}\widehat{\boldsymbol{s}}_t
$$

The optimal forecast is given by

$$
\begin{aligned}
\mathbb{E}_{it}f_t &= \left[A(L)M'(L^{-1})B'(L^{-1})^{-1}\right]_+ V^{-1}B(L)^{-1}x_t \\
&= \frac{a(L)}{\prod_{\tau=1}^{d}(L-\beta_\tau)}M'(L^{-1})\rho_{xx}(L)^{-1}x_t - \sum_{k=1}^{u}\frac{a(\lambda_k)G(\lambda_k)V^{-1}B(L)^{-1}}{(L-\lambda_k)\prod_{\tau\neq k}^{u}(\lambda_k-\lambda_\tau)\prod_{\tau=1}^{d}(\lambda_k-\beta_\tau)}x_t - \\
&\quad - \sum_{k=1}^{d}\frac{a(\beta_k)G(\beta_k)V^{-1}B(L)^{-1}}{(L-\beta_k)\prod_{\tau=1}^{k}(\beta_k-\lambda_\tau)\prod_{\tau\neq k}^{d}(\beta_k-\beta_\tau)}x_t
\end{aligned}
$$

Hence, applying this general example to our particular case

$$
\begin{aligned}
\mathbb{E}_{igt}\xi_t &= \begin{bmatrix}\frac{\tau_\eta^{-1/2}}{1-\rho L} & 0\end{bmatrix}\begin{bmatrix}\frac{\tau_\eta^{-1/2}L}{L-\rho} \\ \tau_g^{-1/2}\end{bmatrix}\frac{\tau_g\lambda_g}{\rho_g}\frac{(1-\rho L)(L-\rho)}{(1-\lambda_g L)(L-\lambda_g)}x_{it} - \\
&\quad - \begin{bmatrix}\frac{\tau_\eta^{-1/2}}{1-\rho\lambda_g} & 0\end{bmatrix}\begin{bmatrix}\frac{\tau_\eta^{-1/2}\lambda_g}{\lambda_g-\rho} \\ \tau_g^{-1/2}\end{bmatrix}\frac{(\lambda_g-\rho)(\lambda_g-\lambda_g)}{(\lambda_g-\lambda_g)(L-\lambda_g)}\frac{\tau_g\lambda_g}{\rho}\frac{(1-\rho L)}{(1-\lambda_g L)}x_{igt} \\
&= \frac{\tau_g\lambda_g}{\rho\tau_\eta(1-\lambda_g\rho)}\frac{1}{1-\lambda_g L}x_{igt} \\
&= G_{1g}(L)x_{igt}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}_{igt} a_{kt+1} &= \begin{bmatrix} \tau_\eta^{-1/2} \frac{h_k(L)}{L(1-\rho L)} & 0 \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L^{-1}} \\ \tau_g^{-1/2} \end{bmatrix} \frac{(L-\rho)(1-\rho L)}{(L-\lambda_g)(1-\lambda_g L)} \frac{\lambda_g \tau_g}{\rho} x_{igt} - \\
&\quad - \begin{bmatrix} \tau_\eta^{-1/2} \frac{h_k(\lambda_g)}{\lambda_g(1-\rho\lambda_g)} & 0 \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho \lambda_g^{-1}} \\ \tau_g^{-1/2} \end{bmatrix} \frac{(\lambda_g-\rho)(1-\rho L)}{(L-\lambda_g)(1-\lambda_g L)} \frac{\lambda_g \tau_g}{\rho} x_{igt} \\
&= \frac{\lambda_g \tau_g}{\rho \tau_\eta} \left[\frac{h_k(L)}{(L-\lambda_g)(1-\lambda_g L)} - \frac{(1-\rho L) h_k(\lambda_g)}{(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)} \right] x_{igt} \\
&= G_{2k}(L) x_{igt} \\
\mathbb{E}_{igt} a_{kt} &= \begin{bmatrix} \tau_\eta^{-1/2} \frac{h_k(L)}{1-\rho L} & 0 \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L^{-1}} \\ \tau_g^{-1/2} \end{bmatrix} \frac{(L-\rho)(1-\rho L)}{(L-\lambda_g)(1-\lambda_g L)} \frac{\lambda_g \tau_g}{\rho} x_{igt} - \\
&\quad - \begin{bmatrix} \tau_\eta^{-1/2} \frac{h_k(\lambda_g)}{1-\rho\lambda_g} & 0 \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho \lambda_g^{-1}} \\ \tau_g^{-1/2} \end{bmatrix} \frac{(\lambda_g-\rho)(1-\rho L)}{(L-\lambda_g)(1-\lambda_g L)} \frac{\lambda_g \tau_g}{\rho} x_{igt} \\
&= \frac{\lambda_g \tau_g}{\rho \tau_\eta} \left[\frac{L h_k(L)}{(L-\lambda_g)(1-\lambda_g L)} - \frac{(1-\rho L) \lambda_g h_k(\lambda_g)}{(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)} \right] x_{igt} \\
&= G_{3k}(L) x_{igt} \\
\mathbb{E}_{igt} \left(a_{igt+1} - a_{gt+1} \right) &= \begin{bmatrix} 0 & \tau_g^{-1/2} \frac{h_g(L)}{L} \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L^{-1}} \\ \tau_g^{-1/2} \end{bmatrix} \frac{(L-\rho)(1-\rho L)}{(L-\lambda_g)(1-\lambda_g L)} \frac{\lambda_g \tau_g}{\rho} x_{igt} - \\
&\quad - \begin{bmatrix} 0 & \tau_g^{-1/2} \frac{h_g(\lambda_g)}{\lambda_g} \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho \lambda_g^{-1}} \\ \tau_g^{-1/2} \end{bmatrix} \frac{(\lambda_g-\rho)(1-\rho L)}{(L-\lambda_g)(1-\lambda_g L)} \frac{\lambda_g \tau_g}{\rho} x_{igt} - \\
&\quad - \tau_g^{-1} \frac{h_g(0)}{L} \frac{(1-\rho L)\tau_g}{1-\lambda_g L} x_{igt} \\
&= \frac{\lambda_g}{\rho} \left[\frac{(L-\rho) h_g(L)}{L(L-\lambda_g)} - \frac{(\lambda_g-\rho) h_g(\lambda_g)}{\lambda_g(L-\lambda_g)} - \frac{\rho h_g(0)}{\lambda_g L} \right] \frac{1-\rho L}{1-\lambda_g L} x_{igt} \\
&= G_{4g}(L) x_{igt}
\end{aligned}
$$

Recall the best response for agent i in group g (A.32), which we rewrite for convenience

$$
a_{igt} = \varphi_g \mathbb{E}_{igt} \xi_t + \beta_g \mathbb{E}_{igt} a_{igt+1} + \sum_{j=1}^{2} \gamma_{gj} \mathbb{E}_{igt} a_{jt} + \sum_{j=1}^{2} \alpha_{gj} \mathbb{E}_{igt} a_{jt+1}
$$

The fixed point problem is

$$
h_g(L) x_{igt} = \varphi_g G_{1g}(L) x_{igt} + \beta_g G_{4g}(L) x_{igt} + \sum_{k=1}^{n} \gamma_{gk} G_{3k}(L) x_{igt} + \sum_{k=1}^{n} \alpha_{gk} G_{2k}(L) x_{igt} + \beta_g G_{2g}(L) x_{igt}
$$

$$
\begin{aligned}
&= \frac{\varphi_g \lambda_g \tau_g}{(1-\lambda_g L)(1-\lambda_g \rho)\rho\tau_\eta} x_{igt} + \frac{\beta_g \lambda_g}{\rho}\left[\frac{(L-\rho)h_g(L)}{L(L-\lambda_g)} - \frac{(\lambda_g-\rho)h_g(\lambda_g)}{\lambda_g(L-\lambda_g)} - \frac{\rho h_g(0)}{\lambda_g L}\right]\frac{1-\rho L}{1-\lambda_g L} x_{igt} + \\
&\quad + \sum_{k=1}^{n} \frac{\gamma_{gk}\lambda_g\tau_g}{\rho\tau_\eta}\left[\frac{L h_k(L)}{(L-\lambda_g)(1-\lambda_g L)} - \frac{(1-\rho L)\lambda_g h_k(\lambda_g)}{(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)}\right] x_{igt} + \\
&\quad + \sum_{k=1}^{n} \frac{\alpha_{gk}\lambda_g\tau_g}{\rho\tau_\eta}\left[\frac{h_k(L)}{(L-\lambda_g)(1-\lambda_g L)} - \frac{(1-\rho L) h_k(\lambda_g)}{(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)}\right] x_{igt} + \\
&\quad + \frac{\beta_g\lambda_g\tau_g}{\rho\tau_\eta}\left[\frac{h_g(L)}{(L-\lambda_g)(1-\lambda_g L)} - \frac{(1-\rho L) h_g(\lambda_g)}{(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)}\right] x_{igt}
\end{aligned}
$$

Rearranging terms on the LHS and RHS

$$
\begin{aligned}
&h_g(L)\left[1 - \frac{\beta_g\lambda_g(L-\rho)(1-\rho L)}{\rho L(L-\lambda_g)(1-\lambda_g L)} - \frac{\beta_g\lambda_g\tau_g}{\rho\tau_\eta(L-\lambda_g)(1-\lambda_g L)}\right] x_{igt} - \\
&\qquad - \sum_{k=1}^{n} h_k(L)\frac{\gamma_{gk}\lambda_g\tau_g L}{\rho\tau_\eta(L-\lambda_g)(1-\lambda_g L)} x_{igt} - \sum_{k=1}^{n} h_k(L)\frac{\alpha_{gk}\lambda_g\tau_g}{\rho\tau_\eta(L-\lambda_g)(1-\lambda_g L)} x_{igt} = \\
&= \frac{\varphi_g\lambda_g\tau_g}{(1-\lambda_g L)(1-\lambda_g\rho)\rho\tau_\eta} x_{igt} - h_g(\lambda_g)\left[\frac{\beta_g(\lambda_g-\rho)(1-\rho L)}{\rho(L-\lambda_g)(1-\lambda_g L)} + \frac{\beta_g\lambda_g\tau_g(1-\rho L)}{\rho\tau_\eta(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)}\right] x_{igt} - \\
&\qquad - h_g(0)\frac{\beta_g(1-\rho L)}{L(1-\lambda_g L)} x_{igt} - \sum_{k=1}^{n} h_k(\lambda_g)\frac{\gamma_{gk}\lambda_g^2\tau_g(1-\rho L)}{\rho\tau_\eta(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)} x_{igt} - \\
&\qquad - \sum_{k=1}^{n} h_k(\lambda_g)\frac{\alpha_{gk}\lambda_g\tau_g(1-\rho L)}{\rho\tau_\eta(1-\rho\lambda_g)(L-\lambda_g)(1-\lambda_g L)} x_{igt}
\end{aligned}
$$

Multiplying both sides by $L(L-\lambda_g)(1-\lambda_g L)$,

$$
\begin{aligned}
&h_g(L)\left\{L(L-\lambda_g)(1-\lambda_g L) - \frac{\beta_g\lambda_g}{\rho}\left[(L-\rho)(1-\rho L) + \frac{\tau_g}{\tau_\eta}L\right]\right\} x_{igt} - \\
&\quad - \frac{\lambda_g\tau_g}{\rho\tau_\eta} L \sum_{k=1}^{n} h_k(L)\left[\gamma_{gk}L + \alpha_{gk}\right] x_{igt} \\
&= \frac{\varphi\lambda_g\tau_g}{\rho\tau_\eta(1-\lambda_g\rho)} L(L-\lambda_g) x_{igt} - h_g(0)\beta_g(1-\rho L)(L-\lambda_g) x_{igt} - \\
&\quad - \left\{h_g(\lambda_g)\beta_g\left[\frac{\lambda_g-\rho}{\rho} + \frac{\lambda_g\tau_g}{\rho\tau_\eta(1-\lambda_g\rho)}\right] + \frac{\lambda_g\tau_g}{\rho\tau_\eta(1-\lambda_g\rho)}\sum_{k=1}^{n} h_k(\lambda_g)[\lambda_g\gamma_{gk} + \alpha_{gk}]\right\} L(1-\rho L) x_{igt}
\end{aligned}
\tag{A.38}
$$

Aggregating across agents, we can write the above system of equations in terms of $\boldsymbol{h}(L)$ in matrix form

$$\begin{aligned}\boldsymbol{C}(L)\boldsymbol{h}(L)x_t &= \boldsymbol{d}(L)x_t \\ \boldsymbol{C}(L)\boldsymbol{h}(L) &= \boldsymbol{d}(L) \end{aligned} \tag{A.39}$$

where

$$\begin{aligned}\boldsymbol{C}(L) = \text{diag}\left\{L(L-\lambda_g)(1-\lambda_g L)\right\} - \boldsymbol{\beta}\ \text{diag}\left\{\frac{\lambda_g}{\rho}\left[(L-\rho)(1-\rho L) + \frac{\tau_g}{\tau_\eta}L\right]\right\} - \\ - \text{diag}\left\{\frac{\lambda_g \tau_g}{\rho \tau_\eta}L^2\right\}\boldsymbol{\gamma} - \text{diag}\left\{\frac{\lambda_g \tau_g}{\rho \tau_\eta}L\right\}\boldsymbol{\alpha}\end{aligned} \tag{A.40}$$

One can verify from (A.37) the following identity

$$\lambda_g + \frac{1}{\lambda_g} = \rho + \frac{1}{\rho} + \frac{\tau_g}{\rho\tau_\eta}$$

and, to simplify algebra, one can rewrite (A.40) as

$$\begin{aligned}\boldsymbol{C}(L) &= \text{diag}\left\{-\lambda_g L\left[L^2 - L\left(\lambda_g + \frac{1}{\lambda_g}\right) + 1\right]\right\} - \boldsymbol{\beta}\ \text{diag}\left\{-\lambda_g\left[L^2 - L\left(\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\eta}\right) + 1\right]\right\} - \\ &\quad - \text{diag}\ \left\{\frac{\lambda_g \tau_g}{\rho \tau_\eta}L^2\right\}\boldsymbol{\gamma} - \text{diag}\ \left\{\frac{\lambda_g \tau_g}{\rho \tau_\eta}L\right\}\boldsymbol{\alpha} \\ &= \text{diag}\left\{\lambda_g\right\}\left[(\boldsymbol{\beta} - IL)\ \text{diag}\left\{L^2 - \left(\frac{1}{\rho} + \rho + \frac{\tau_g}{\rho\tau_\eta}\right)L + 1\right\} - L^2\ \text{diag}\left\{\frac{\tau_g}{\rho\tau_\eta}\right\}\boldsymbol{\gamma} - L\ \text{diag}\left\{\frac{\tau_g}{\rho\tau_\eta}\right\}\boldsymbol{\alpha}\right] \\ &= \text{diag}\left\{\lambda_g\right\}\left[(\boldsymbol{\beta} - IL)\ \text{diag}\left\{\left(L - \frac{1}{\rho}\right)(L-\rho)\right\} - (\boldsymbol{\beta} - IL)L\ \text{diag}\left\{\frac{\tau_g}{\rho\tau_\eta}\right\} - \right. \\ &\quad \left. - L^2\ \text{diag}\left\{\frac{\tau_g}{\rho\tau_\eta}\right\}\boldsymbol{\gamma} - L\ \text{diag}\left\{\frac{\tau_g}{\rho\tau_\eta}\right\}\boldsymbol{\alpha}\right]\end{aligned}$$

That is, we can write $\boldsymbol{C}(z) = \begin{bmatrix} C_{11}(z) & C_{12}(z) \\ C_{21}(z) & C_{22}(z)\end{bmatrix}$, where

$$\begin{aligned}C_{11}(z) &= \lambda_1\left[(\beta_1 - z)\left(z - \frac{1}{\rho}\right)(z-\rho) + \frac{\tau_1}{\rho\tau_\eta}z\left[z(1-\gamma_{11}) - \delta_{11}\right]\right] \\ C_{12}(z) &= -\lambda_1 z\frac{\tau_1}{\rho\tau_\eta}(z\gamma_{12} + \delta_{12})\end{aligned}$$

$$C_{21}(z) = -\lambda_2 z \frac{\tau_2}{\rho\tau_\eta}(z\gamma_{21} + \delta_{21})$$

$$C_{22}(z) = \lambda_2 \left[(\beta_2 - z)\left(z - \frac{1}{\rho}\right)(z-\rho) + \frac{\tau_2}{\rho\tau_\eta} z\left[z(1-\gamma_{22}) - \delta_{22}\right]\right]$$

We can also write the RHS of (A.38) as

$$d_g(L) = \frac{\varphi_g \lambda_g \tau_g}{\rho\tau_\eta(1-\lambda_g\rho)} L(L-\lambda_g) - \beta_g(1-\rho L)(L-\lambda_g)h_g(0) -$$
$$- \left\{ h_g(\lambda_g)\beta_g \left[\frac{\lambda_g - \rho}{\rho} + \frac{\lambda_g \tau_g}{\rho\tau_\eta(1-\lambda_g\rho)}\right] + \frac{\lambda_g\tau_g}{\rho\tau_\eta(1-\lambda_g\rho)} \sum_{k=1}^{n} h_k(\lambda_g)(\lambda_g\gamma_{gk} + \alpha_{gk}) \right\} L(1-\rho L)$$

Hence, we can write $\boldsymbol{d}(z) = \begin{bmatrix} d_1[z; h_1(\cdot)] \\ d_2[z; h_2(\cdot)] \end{bmatrix}$, where

$$d_1[z;h_1(\cdot)] = \frac{\varphi_1\lambda_1\tau_1}{\rho(1-\lambda_1\rho)\tau_\eta} z(z-\lambda_1) - \frac{1}{\rho}\left[\beta_1(\lambda_1-\rho) + (\delta_{11} + \lambda_1\gamma_{11})\frac{\lambda_1\tau_1}{(1-\lambda_1\rho)\tau_\eta}\right] z(1-\rho z)h_1(\lambda_1) -$$
$$- \beta_1(z-\lambda_1)(1-\rho z)h_1(0)$$

$$d_2[z;h_2(\cdot)] = \frac{\varphi_2\lambda_2\tau_2}{\rho(1-\lambda_2\rho)\tau_\eta} z(z-\lambda_2) - \frac{1}{\rho}\left[\beta_2(\lambda_2-\rho) + (\delta_{22} + \lambda_2\gamma_{22})\frac{\lambda_2\tau_2}{(1-\lambda_2\rho)\tau_\eta}\right] z(1-\rho z)h_2(\lambda_2) -$$
$$- \beta_2(z-\lambda_2)(1-\rho z)h_2(0)$$

From (A.39), the solution to the policy function is given by

$$\boldsymbol{h}(L) = \boldsymbol{C}(L)^{-1}\boldsymbol{d}(L) = \frac{\text{adj } \boldsymbol{C}(L)}{\det \boldsymbol{C}(L)} \boldsymbol{d}(L)$$

Hence, we need to obtain $\det \boldsymbol{C}(L)$. Note that the degree of $\det \boldsymbol{C}(L)$ is a polynomial of degree 6 on z. Denote the inside roots of $\det \boldsymbol{C}(L)$ as $\{\zeta_1, \zeta_2, \zeta_3, \zeta_4\}$, and the outside roots as $\{\vartheta_1^{-1}, \vartheta_2^{-1}\}$. Because agents cannot use future signals, the inside roots have to be removed. Note that the number of free constants in $\boldsymbol{d}(L)$ is 4:

$$\{h_1(0)\}, \quad \left\{ h_1(\lambda_1)\beta_1 \left[\frac{\lambda_1 - \rho}{\rho} + \frac{\lambda_1\tau_1}{\rho\tau_\eta(1-\lambda_1\rho)}\right] + \frac{\lambda_1\tau_1}{\rho\tau_\eta(1-\lambda_1\rho)} \sum_{k=1}^{2} h_k(\lambda_1)(\lambda_1\gamma_{1k} + \alpha_{1k}) \right\}$$

$$\{h_2(0)\}, \quad \left\{ h_2(\lambda_2)\beta_2 \left[\frac{\lambda_2 - \rho}{\rho} + \frac{\lambda_2\tau_2}{\rho\tau_\eta(1-\lambda_2\rho)}\right] + \frac{\lambda_2\tau_2}{\rho\tau_\eta(1-\lambda_2\rho)} \sum_{k=1}^{2} h_k(\lambda_2)(\lambda_2\gamma_{2k} + \alpha_{2k}) \right\}$$

With a unique solution, it has to be the case that the number of outside roots is 2. By

Cramer's rule, $h_g(L)$ is given by

$$h_1(z) = \frac{\det \begin{bmatrix} d_1(z) & C_{12}(z) \\ d_2(z) & C_{22}(z) \end{bmatrix}}{\det \boldsymbol{C}(z)}$$

$$h_2(z) = \frac{\det \begin{bmatrix} C_{11}(z) & d_1(z) \\ C_{21}(z) & d_2(z) \end{bmatrix}}{\det \boldsymbol{C}(z)}$$

Which are the policy function for groups 1 (consumers) and 2 (firms). The degree of the numerator is 5, as the highest degree of $d_g(L)$ is 1 degree less than that of $\boldsymbol{C}(L)$. By choosing the appropriate constants $\left\{h_1(0), \widetilde{h}(\lambda_1), h_2(0), \widetilde{h}(\lambda_2)\right\}$, the 4 inside roots will be removed. Therefore, the 4 constants are solutions to the following system of linear equations

$$\det \begin{bmatrix} d_1(\zeta_1) & C_{12}(\zeta_1) \\ d_2(\zeta_1) & C_{22}(\zeta_1) \end{bmatrix} = 0$$

$$\det \begin{bmatrix} d_1(\zeta_2) & C_{12}(\zeta_2) \\ d_2(\zeta_2) & C_{22}(\zeta_2) \end{bmatrix} = 0$$

$$\det \begin{bmatrix} C_{11}(\zeta_3) & d_1(\zeta_3) \\ C_{21}(\zeta_3) & d_2(\zeta_3) \end{bmatrix} = 0$$

$$\det \begin{bmatrix} C_{11}(\zeta_4) & d_1(\zeta_4) \\ C_{21}(\zeta_4) & d_2(\zeta_4) \end{bmatrix} = 0$$

After removing the inside roots in the denominator, the degree of the numerator is 1 and the degree of the denominator is 2. The above determinants can be written as a system of 4 equations and 4 unknowns (our free constants). Once we have set the appropriate free constants the policy functions will be

$$h_g(z) = \frac{\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z}{(1 - \vartheta_1 z)(1 - \vartheta_2 z)}$$

and hence we have

$$\begin{aligned} a_{gt} &= h_g(L)\xi_t \\ &= \frac{\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z}{(1 - \vartheta_1 z)(1 - \vartheta_2 z)} \xi_t \end{aligned}$$

$$
\begin{aligned}
&= \psi_{g1}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 L}\xi_t + \psi_{g2}\left(1-\frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta_2 L}\xi_t \\
&= \psi_{g1}\widetilde{\vartheta}_{1t} + \psi_{g2}\widetilde{\vartheta}_{2t}
\end{aligned}
$$

We can write

$$
\begin{aligned}
\boldsymbol{a}_t &= \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \\
&= Q\widetilde{\vartheta}_t \\
&= \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix} \begin{bmatrix} \widetilde{\vartheta}_{1t} \\ \widetilde{\vartheta}_{2t} \end{bmatrix} \\
&= \begin{bmatrix} \psi_{11}\widetilde{\vartheta}_{1t} + \psi_{12}\widetilde{\vartheta}_{2t} \\ \psi_{21}\widetilde{\vartheta}_{1t} + \psi_{22}\widetilde{\vartheta}_{2t} \end{bmatrix}
\end{aligned}
$$

Notice that we can write

$$
\begin{aligned}
\widetilde{\vartheta}_{1t}(1-\vartheta_1 L) = \left(1-\frac{\vartheta_1}{\rho}\right)\xi_t \implies \widetilde{\vartheta}_{1t} = \vartheta_1\widetilde{\vartheta}_{1t-1} + \left(1-\frac{\vartheta_1}{\rho}\right)\xi_t \\
\widetilde{\vartheta}_{2t}(1-\vartheta_2 L) = \left(1-\frac{\vartheta_2}{\rho}\right)\xi_t \implies \widetilde{\vartheta}_{2t} = \vartheta_2\widetilde{\vartheta}_{2t-1} + \left(1-\frac{\vartheta_2}{\rho}\right)\xi_t
\end{aligned}
$$

Which we can write as a system as

$$
\widetilde{\vartheta}_t = \Lambda\widetilde{\vartheta}_{t-1} + \Gamma\xi_t
$$

where

$$
\Lambda = \begin{bmatrix} \vartheta_1 & 0 \\ 0 & \vartheta_2 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 1-\frac{\vartheta_1}{\rho} \\ 1-\frac{\vartheta_2}{\rho} \end{bmatrix}
$$

Hence, we can write

$$
\begin{aligned}
\boldsymbol{a}_t &= Q\widetilde{\theta}_t \\
&= Q(\Lambda\widetilde{\theta}_{t-1} + \Gamma\xi_t) \\
&= Q\Lambda\widetilde{\theta}_{t-1} + Q\Gamma\xi_t \\
&= Q\Lambda Q^{-1}\boldsymbol{a}_{t-1} + Q\Gamma\xi_t
\end{aligned}
$$

$$= A\boldsymbol{a}_{t-1} + B\xi_t \tag{A.41}$$

Finally, we need to show that (3.6)-(3.7) hold. First, notice that in the standard FIRE framework there is no information friction, $\vartheta_1 = \vartheta_2 = 0$.[25] Therefore, the dynamics under equilibrium follow $\boldsymbol{a}_t = A_{\text{FIRE}}\boldsymbol{a}_{t-1} + B_{\text{FIRE}}\xi_t$ where

$$A_{\text{FIRE}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \qquad B_{\text{FIRE}} = \begin{bmatrix} \psi_{11} + \psi_{12} \\ \psi_{21} + \psi_{22} \end{bmatrix}$$

Under the standard FIRE case (A.51), which I rewrite for convenience

$$\boldsymbol{a}_t = (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\varphi}\xi_t + (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\delta}\mathbb{E}_t\boldsymbol{a}_{t+1}$$

In order to find the solution dynamics under FIRE, we proceed with a guess and verify approach. Assume that $\boldsymbol{a}_t = D\xi_t$. Using the method of undetermined coefficients

$$\begin{aligned} D\xi_t &= (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\varphi}\xi_t + (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\delta}\mathbb{E}_t D\xi_{t+1} \\ &= (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\varphi}\xi_t + (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\delta} D\rho\xi_t \\ D &= (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\varphi} + (I - \boldsymbol{\gamma})^{-1}\boldsymbol{\delta} D\rho \end{aligned}$$

hence, it must be that $D = [\boldsymbol{I} - (\boldsymbol{I} - \boldsymbol{\gamma})^{-1}\boldsymbol{\delta}\rho]^{-1}(\boldsymbol{I} - \boldsymbol{\gamma})^{-1}\boldsymbol{\varphi}$. Notice that, for consistency, $B_{\text{FIRE}} = D$. As a result, even if we cannot find the analytical form of the individual $(\psi_{11}, \psi_{12}, \psi_{21}, \psi_{22})$, we know that conditions (3.6)-(3.7) hold. □

Proof of Proposition 5. We know that

$$\begin{aligned} a_{gt+1} = h_g(L)\xi_{t+1} &= \left[\frac{\psi_{g1}(\rho - \vartheta_1)}{\rho(1 - \vartheta_1 L)} + \frac{\psi_{g2}(\rho - \vartheta_2)}{\rho(1 - \vartheta_2 L)}\right]\xi_{t+1} \\ \overline{\mathbb{E}}_{gt}a_{gt+1} &= \frac{(\rho - \lambda_g)(1 - \rho\lambda_g)}{\rho(L - \lambda_g)(1 - \lambda_g L)}\left[h_g(L) - \frac{1 - \rho L}{1 - \rho\lambda_g}h_g(\lambda_g)\right]\xi_t \\ \overline{\mathbb{E}}_{gt-1}a_{gt+1} &= \frac{(\rho - \lambda_g)(1 - \rho\lambda_g)}{\rho(L - \lambda_g)(1 - \lambda_g L)}\left[\frac{h_g(L)}{L} - \frac{1 - \rho L}{\lambda_g(1 - \rho\lambda_g)}h_g(\lambda_g)\right]\xi_{t-1} \end{aligned}$$

25. To verify this, solve the household problem as if there were no firms (only the consumer group), assuming that the nominal interest rate follows an AR(1) process. In that case one can obtain analytically ψ_1, and verify that we are back to the standard case holds when $\vartheta_1 = 0$ (see e.g., Angeletos and Huo (2018)). Alternatively, one can verify numerically that the dynamics implied by the beyond FIRE case when we restrict $\vartheta_1 = \vartheta_2 = 0$ coincide with the dynamics of the FIRE case $\boldsymbol{a}_t = D\xi_t$.

Let us first obtain the LHS of the regression

$$
\begin{aligned}
a_{gt+1} - \overline{\mathbb{E}}_{gt} a_{gt+1} &= \frac{h_g(L)}{L}\xi_t - \frac{(\rho-\lambda_g)(1-\rho\lambda_g)}{\rho(L-\lambda_g)(1-\lambda_g L)}\left[h_g(L) - \frac{1-\rho L}{1-\rho\lambda_g}h_g(\lambda_g)\right]\xi_t \\
&= \left\{h_g(L)\left[\frac{1}{L} - \frac{(\rho-\lambda_g)(1-\rho\lambda_g)}{\rho(L-\lambda_g)(1-\lambda_g L)}\right] - h_g(\lambda_g)\frac{(\rho-\lambda_g)(1-\rho L)}{\rho(L-\lambda_g)(1-\lambda_g L)}\right\}\xi_t \\
&= \left\{h_g(L)\left[\frac{1}{L} - \frac{(\rho-\lambda_g)(1-\rho\lambda_g)}{\rho(L-\lambda_g)(1-\lambda_g L)}\right] - h_g(\lambda_g)\frac{(\rho-\lambda_g)(1-\rho L)}{\rho(L-\lambda_g)(1-\lambda_g L)}\right\}\frac{1}{1-\rho L}\eta_t \\
&= \left[h_g(L)\frac{\lambda_g(L-\rho)(1-\rho\lambda_g)}{L\rho(L-\lambda_g)(1-\lambda_g L)(1-\rho L)} - h_g(\lambda_g)\frac{(\rho-\lambda_g)}{\rho(L-\lambda_g)(1-\lambda_g L)}\right]\eta_t \\
&= \left\{\frac{\psi_{g1}(\rho-\vartheta_1)(1-\rho L)}{\rho(1-\vartheta_1\lambda_g)(1-\rho L)(1-\lambda_g L)} - \frac{\psi_{g1}(\rho-\vartheta_1)\vartheta_1\lambda_g(\rho-L)}{\rho^2(1-\vartheta_1\lambda_g)(1-\vartheta_1 L)(1-\lambda_g L)} + \right. \\
&\qquad \left. + \frac{\psi_{g2}(\rho-\vartheta_2)(1-\rho L)}{\rho(1-\vartheta_2\lambda_g)(1-\rho L)(1-\lambda_g L)} - \frac{\psi_{g2}(\rho-\vartheta_2)\vartheta_2\lambda_g(\rho-L)}{\rho^2(1-\vartheta_2\lambda_g)(1-\vartheta_2 L)(1-\lambda_g L)}\right\}\frac{1}{L}\eta_t \\
&= \left[\frac{\psi_{g1}(\rho-\vartheta_1)}{\rho(1-\vartheta_1\lambda_g)} + \frac{\psi_{g2}(\rho-\vartheta_2)}{\rho(1-\vartheta_2\lambda_g)}\right]\frac{1}{1-\lambda_g L}\eta_{t+1} \\
&\qquad - \frac{\psi_{g1}(\rho-\vartheta_1)\vartheta_1\lambda_g}{\rho^2(1-\vartheta_1\lambda_g)}\frac{\rho-L}{1-\vartheta_1 L}\frac{1}{1-\lambda_g L}\eta_{t+1} \\
&\qquad - \frac{\psi_{g2}(\rho-\vartheta_2)\vartheta_2\lambda_g}{\rho^2(1-\vartheta_2\lambda_g)}\frac{\rho-L}{1-\vartheta_2 L}\frac{1}{1-\lambda_g L}\eta_{t+1} \\
&= \underbrace{\left[\frac{\psi_{g1}(\rho-\vartheta_1)}{\rho(1-\vartheta_1\lambda_g)} + \frac{\psi_{g2}(\rho-\vartheta_2)}{\rho(1-\vartheta_2\lambda_g)}\right]}_{\Xi_{g,1}}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j} \\
&\qquad - \frac{\psi_{g1}(\rho-\vartheta_1)\vartheta_1\lambda_g}{\rho(1-\vartheta_1\lambda_g)}\frac{1-\rho^{-1}L}{(1-\vartheta_1 L)(1-\lambda_g L)}\eta_{t+1} \\
&\qquad - \frac{\psi_{g2}(\rho-\vartheta_2)\vartheta_2\lambda_g}{\rho(1-\vartheta_2\lambda_g)}\frac{1-\rho^{-1}L}{(1-\vartheta_2 L)(1-\lambda_g L)}\eta_{t+1} \\
&= \Xi_{g,1}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j} \\
&\qquad - \frac{\psi_{g1}(\rho-\vartheta_1)\vartheta_1\lambda_g}{\rho(1-\vartheta_1\lambda_g)}\left[\frac{\vartheta_1-\rho^{-1}}{\vartheta_1-\lambda_g}\frac{1}{1-\vartheta_1 L} + \frac{\rho^{-1}-\lambda_g}{\vartheta_1-\lambda_g}\frac{1}{1-\lambda_g L}\right]\eta_{t+1} \\
&\qquad - \frac{\psi_{g2}(\rho-\vartheta_2)\vartheta_2\lambda_g}{\rho(1-\vartheta_2\lambda_g)}\left[\frac{\vartheta_2-\rho^{-1}}{\vartheta_2-\lambda_g}\frac{1}{1-\vartheta_2 L} + \frac{\rho^{-1}-\lambda_g}{\vartheta_2-\lambda_g}\frac{1}{1-\lambda_g L}\right]\eta_{t+1} \\
&= \Xi_{g,1}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j}
\end{aligned}
$$

$$\underbrace{-\frac{\psi_{g1}(\rho-\vartheta_1)\vartheta_1\lambda_g(\vartheta_1-\rho^{-1})}{\rho(1-\vartheta_1\lambda_g)(\vartheta_1-\lambda_g)}}_{+\Xi_{g,2}}\sum_{j=0}^{\infty}\vartheta_1^j\eta_{t+1-j}$$

$$\underbrace{-\frac{\psi_{g1}(\rho-\vartheta_1)\vartheta_1\lambda_g(\rho-\lambda_g)}{\rho(1-\vartheta_1\lambda_g)(\vartheta_1-\lambda_g)}}_{+\Xi_{g,3}}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j}$$

$$\underbrace{-\frac{\psi_{g2}(\rho-\vartheta_2)\vartheta_2\lambda_g(\vartheta_2-\rho^{-1})}{\rho(1-\vartheta_2\lambda_g)(\vartheta_2-\lambda_g)}}_{+\Xi_{g,4}}\sum_{j=0}^{\infty}\vartheta_2^j\eta_{t+1-j}$$

$$\underbrace{-\frac{\psi_{g2}(\rho-\vartheta_2)\vartheta_2\lambda_g(\rho^{-1}-\lambda_g)}{\rho(1-\vartheta_2\lambda_g)(\vartheta_2-\lambda_g)}}_{+\Xi_{g,5}}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j}$$

$$=\underbrace{(\Xi_{g,1}+\Xi_{g,3}+\Xi_{g,5})}_{\chi_{g,1}}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j}$$

$$+\underbrace{\Xi_{g,2}}_{\chi_{g,2}}\sum_{j=0}^{\infty}\vartheta_1^j\eta_{t+1-j}+\underbrace{\Xi_{g,4}}_{\chi_{g,3}}\sum_{j=0}^{\infty}\vartheta_2^j\eta_{t+1-j}$$

which is the LHS of the regression. Let us now compute the RHS of the regression

$$\begin{aligned}\overline{\mathbb{E}}_{gt}a_{gt+1}-\overline{\mathbb{E}}_{gt-1}a_{gt+1}&=\frac{(\rho-\lambda_g)(1-\rho\lambda_g)}{\rho(L-\lambda_g)(1-\lambda_g L)}\left[\frac{h_g(L)}{L}-\frac{1-\rho L}{L(1-\rho\lambda_g)}h_g(\lambda_g)-\frac{h_g(L)}{L}+\frac{1-\rho L}{\lambda_g(1-\rho\lambda_g)}h_g(\lambda_g)\right]\xi_{t-}\\&=\frac{(\rho-\lambda_g)h_g(\lambda_g)}{\rho\lambda_g}\frac{1}{1-\lambda_g L}\eta_t\\&=\underbrace{\frac{(\rho-\lambda_g)h_g(\lambda_g)}{\rho\lambda_g}}_{\chi_{g,4}}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t-j}\end{aligned}$$

and we have the RHS. We can write the numerator of $\mathcal{K}_{g,CG}$ as

$$\mathbb{C}\left[a_{gt+1}-\overline{\mathbb{E}}_{gt}a_{gt+1},\overline{\mathbb{E}}_{gt}a_{gt+1}-\overline{\mathbb{E}}_{gt-1}a_{gt+1}\right]=$$

$$=\mathbb{C}\Bigg[\chi_{g,1}\sum_{j=0}^{\infty}\lambda_g^j\eta_{t+1-j}+\chi_{g,2}\sum_{j=0}^{\infty}\vartheta_1^j\eta_{t+1-j}+\chi_{g,3}\sum_{j=0}^{\infty}\vartheta_2^j\eta_{t+1-j},$$

$$+ \chi_{g,4} \sum_{j=0}^{\infty} \lambda_g^j \eta_{t-j} \Bigg] =$$

$$= \sigma_\eta^2 \left[\frac{\chi_{g,1}\chi_{g,4}\lambda_g}{1-\lambda_g^2} + \frac{\chi_{g,2}\chi_{g,4}\vartheta_1}{1-\lambda_g\vartheta_1} + \frac{\chi_{g,3}\chi_{g,4}\vartheta_2}{1-\lambda_g\vartheta_2} \right]$$

The denominator is

$$\mathbb{V}\left[\overline{\mathbb{E}}_{gt} a_{gt+1} - \overline{\mathbb{E}}_{gt-1} a_{gt+1}\right] = \mathbb{V}\left[\chi_{g,4} \sum_{j=0}^{\infty} \lambda_g^j \eta_{t-j}\right]$$
$$= \sigma_\eta^2 \frac{\chi_{g,4}^2}{1-\lambda_g^2}$$

Finally, the coefficient is

$$\mathcal{K}_{g,CG} = \frac{\chi_{g,1}}{\chi_{g,4}} \lambda_g + \frac{\chi_{g,2}}{\chi_{g,4}} \frac{\vartheta_1(1-\lambda_g^2)}{1-\lambda_g\vartheta_1} + \frac{\chi_{g,3}}{\chi_{g,4}} \frac{\vartheta_2(1-\lambda_g^2)}{1-\lambda_g\vartheta_2}$$

□

Proof of Proposition 6. Recall that equilibrium dynamics satisfy,

$$\boldsymbol{a}_t = \overline{\boldsymbol{\varphi}} v_t + \overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1} \tag{A.42}$$

We need to find the conditions under which the equilibrium process is stationary. This sums up to having all the eigenvalues in matrix $\overline{\boldsymbol{\delta}}^{-1}$ outside the unit circle. This restriction is satisfied if

$$\det \overline{\boldsymbol{\delta}}^{-1} > 1 \tag{A.43}$$
$$\det \overline{\boldsymbol{\delta}}^{-1} - \operatorname{tr} \overline{\boldsymbol{\delta}}^{-1} > -1 \tag{A.44}$$
$$\det \overline{\boldsymbol{\delta}}^{-1} + \operatorname{tr} \overline{\boldsymbol{\delta}}^{-1} > -1 \tag{A.45}$$

Introducing the respective values in (A.43)-(A.45), we obtain (4.2)-(4.4).

□

Proof of Proposition 7. Recall the equilibrium dynamics described by Proposition 4. We need to find the conditions under which the equilibrium process is stationary. This sums up to having all the eigenvalues in matrix A inside the unit circle. This restriction is satisfied if

$$\det A < 1 \tag{A.46}$$

$$\det A - \mathrm{tr}A > -1 \tag{A.47}$$

$$\det A + \mathrm{tr}A > -1 \tag{A.48}$$

Notice that $\mathrm{tr}A = \vartheta_1 + \vartheta_2$, where ϑ_1 and ϑ_2 are the two roots of the characteristic polynomial of A, and $\det\, A = \vartheta_1\vartheta_2$. Therefore, the above conditions can be translated to

$$\begin{aligned} \vartheta_1\vartheta_2 &< 1 \\ (\vartheta_1 - 1)(\vartheta_2 - 1) &> 0 \\ (\vartheta_1 + 1)(\vartheta_2 + 1) &> 0 \end{aligned}$$

Notice that such system can only be satisfied if both roots are inside the unit circle. Introducing the respective values in (A.46)-(A.48), we obtain (4.7)-(4.9). □

Proof of Proposition 8. The aggregate outcome is

$$\begin{aligned}
y_t &= \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 L}\xi_t + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta_2 L}\xi_t \\
&= \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right)\frac{1}{(1-\vartheta_1 L)(1-\rho L)}\eta_t + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right)\frac{1}{(1-\vartheta_2 L)(1-\rho L)}\eta_t \\
&= \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right)\left[\frac{1}{(\rho-\vartheta_1)L(1-\rho L)} - \frac{1}{(\rho-\vartheta_1)L(1-\vartheta_1 L)}\right]\eta_t + \\
&\qquad + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right)\left[\frac{1}{(\rho-\vartheta_2)L(1-\rho L)} - \frac{1}{(\rho-\vartheta_2)L(1-\vartheta_2 L)}\right]\eta_t \\
&= \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right)\frac{1}{\rho-\vartheta_1}\left[\frac{1}{1-\rho L} - \frac{1}{1-\vartheta_1 L}\right]\eta_{t+1} + \\
&\qquad + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right)\frac{1}{\rho-\vartheta_2}\left[\frac{1}{1-\rho L} - \frac{1}{1-\vartheta_2 L}\right]\eta_{t+1} \\
&= \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right)\frac{1}{\rho-\vartheta_1}\left[\sum_{j=0}^{\infty}(\rho L)^j\eta_{t+1} - \sum_{j=0}^{\infty}(\vartheta_1 L)^j\eta_{t+1}\right] + \\
&\qquad + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right)\frac{1}{\rho-\vartheta_2}\left[\sum_{j=0}^{\infty}(\rho L)^j\eta_{t+1} - \sum_{j=0}^{\infty}(\vartheta_2 L)^j\eta_{t+1}\right] \\
&= \psi_{11}\left(1 - \frac{\vartheta_1}{\rho}\right)\frac{1}{\rho-\vartheta_1}\sum_{\tau=0}^{\infty}(\rho^{\tau+1} - \vartheta_1^{\tau+1})\eta_{t-\tau} + \psi_{12}\left(1 - \frac{\vartheta_2}{\rho}\right)\frac{1}{\rho-\vartheta_2}\sum_{\tau=0}^{\infty}(\rho^{\tau+1} - \vartheta_2^{\tau+1})\eta_{t-\tau}
\end{aligned}$$

$$= \frac{\psi_{11}}{\rho} \sum_{\tau=0}^{\infty} (\rho^{\tau+1} - \vartheta_1^{\tau+1}) \eta_{t-\tau} + \frac{\psi_{12}}{\rho} \sum_{\tau=0}^{\infty} (\rho^{\tau+1} - \vartheta_2^{\tau+1}) \eta_{t-\tau}$$

The PE component is given by

$$\begin{aligned} \text{PE}_t &= -\frac{\beta}{\sigma}(1-\lambda) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t r_{t+k} \\ &= -\frac{\beta}{\sigma}(1-\lambda) \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^c \left[\phi_\pi \pi_{t+k} + \phi_y y_{t+k} + v_{t+k} - \pi_{t+k+1}\right] \end{aligned}$$

where

$$\begin{aligned} \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^c \pi_{t+k} &= \overline{\mathbb{E}}_t^c \frac{L\pi_t}{L-\beta} = \sum_{j=1}^{2} \psi_{2j} \left(1 - \frac{\vartheta_j}{\rho}\right) \overline{\mathbb{E}}_t^c \left[\frac{L}{(L-\beta)(1-\vartheta_j L)} v_t\right] \\ \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^c y_{t+k} &= \overline{\mathbb{E}}_t^c \frac{L y_t}{L-\beta} = \sum_{j=1}^{2} \psi_{1j} \left(1 - \frac{\vartheta_j}{\rho}\right) \overline{\mathbb{E}}_t^c \left[\frac{L}{(L-\beta)(1-\vartheta_j L)} v_t\right] \\ \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t v_{t+k} &= \overline{\mathbb{E}}_t^c \frac{L v_t}{L-\beta} \\ \sum_{k=0}^{\infty} \beta^k \overline{\mathbb{E}}_t^c \pi_{t+k+1} &= \overline{\mathbb{E}}_t^c \frac{\pi_t}{L-\beta} = \sum_{j=1}^{2} \psi_{2j} \left(1 - \frac{\vartheta_j}{\rho}\right) \overline{\mathbb{E}}_t^c \left[\frac{1}{(L-\beta)(1-\vartheta_j L)} v_t\right] \end{aligned}$$

with

$$\begin{aligned} \overline{\mathbb{E}}_t^c \left[\frac{L}{(L-\beta)(1-\vartheta_j L)} v_t\right] &= \left[\begin{bmatrix} \frac{\tau_\eta^{-1/2} L}{(L-\beta)(1-\vartheta_j L)(1-\rho L)} & 0 \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L^{-1}} \\ \tau_1^{-1/2} \end{bmatrix} \frac{1-\rho L^{-1}}{1-\lambda_1 L^{-1}}\right]_+ \frac{\lambda_1 \tau_1}{\rho} \frac{1}{1-\lambda_1 L} \eta_t \\ &= \frac{\rho}{(1-\rho\beta)(\rho-\vartheta_j)} \sum_{k=0}^{\infty} \rho^k \eta_{t-k} \\ &\quad - \frac{\vartheta_j^2 (\rho-\lambda_1)(1-\rho\lambda_1)}{\rho(\rho-\vartheta_j)(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)} \sum_{k=0}^{\infty} \vartheta_j^k \eta_{t-k} \\ &\quad + \frac{\lambda_1 \left\{\lambda_1 - \rho\vartheta_j \left[\beta + \lambda_1 \left(1 - \beta(\rho + \vartheta_j - \lambda_1)\right)\right]\right\}}{\rho(1-\rho\beta)(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)} \sum_{k=0}^{\infty} \lambda_1^k \eta_{t-k} \\ \overline{\mathbb{E}}_t^c \left[\frac{L}{L-\beta} v_t\right] &= \left[\begin{bmatrix} \frac{\tau_\eta^{-1/2} L}{(L-\beta)(1-\rho L)} & 0 \end{bmatrix} \begin{bmatrix} \frac{\tau_\eta^{-1/2}}{1-\rho L^{-1}} \\ \tau_1^{-1/2} \end{bmatrix} \frac{1-\rho L^{-1}}{1-\lambda_1 L^{-1}}\right]_+ \frac{\lambda_1 \tau_1}{\rho} \frac{1}{1-\lambda_1 L} \eta_t \end{aligned}$$

$$= \frac{1}{1-\rho\beta}\sum_{k=0}^{\infty}\rho^k\eta_{t-k} - \frac{\lambda_1}{\rho(1-\rho\beta)}\sum_{k=0}^{\infty}\lambda_1^k\eta_{t-k}$$

$$\overline{\mathbb{E}}_t^c\left[\frac{1}{(L-\beta)(1-\vartheta_j L)}v_t\right] = \left[\left[\tfrac{\tau_\eta^{-1/2}}{(L-\beta)(1-\vartheta_j L)(1-\rho L)} \quad 0\right]\begin{bmatrix}\frac{\tau_\eta^{-1/2}}{1-\rho L^{-1}}\\ \tau_1^{-1/2}\end{bmatrix}\frac{1-\rho L^{-1}}{1-\lambda_1 L^{-1}}\right]_+ \frac{\lambda_1\tau_1}{\rho}\frac{1}{1-\lambda_1 L}\eta_t$$

$$= \frac{\rho^2}{(1-\rho\beta)(\rho-\vartheta_j)}\sum_{k=0}^{\infty}\rho^k\eta_{t-k}$$

$$- \frac{\vartheta_j^3(\rho-\lambda_1)(1-\rho\lambda_1)}{\rho(\rho-\vartheta_j)(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\vartheta_j^k\eta_{t-k}$$

$$- \frac{\lambda_1\left\{\rho\vartheta_j(1+\lambda_1^2)-\lambda_1[\rho+\vartheta_j(1-\rho\beta(1-\rho\vartheta_j))]\right\}}{\rho(1-\rho\beta)(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\lambda_1^k\eta_{t-k}$$

Hence, we have

$$\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t^c\pi_{t+k} = \frac{\psi_{21}+\psi_{22}}{1-\rho\beta}\sum_{k=0}^{\infty}\rho^k\eta_{t-k}$$

$$+ \frac{\lambda_1}{\rho^2(1-\rho\beta)}\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\lambda_1-\rho\vartheta_j[\beta+\lambda_1(1-\beta(\rho+\vartheta_j-\lambda_1))]]\psi_{2j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\lambda_1^k\eta_{t-k}$$

$$- \sum_{j=1}^{2}\frac{\vartheta_j^2(\rho-\lambda_1)(1-\rho\lambda_1)\psi_{2j}}{\rho^2(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\vartheta_j^k\eta_{t-k}$$

$$\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t^c y_{t+k} = \frac{\psi_{11}+\psi_{12}}{1-\rho\beta}\sum_{k=0}^{\infty}\rho^k\eta_{t-k}$$

$$+ \frac{\lambda_1}{\rho^2(1-\rho\beta)}\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\lambda_1-\rho\vartheta_j[\beta+\lambda_1(1-\beta(\rho+\vartheta_j-\lambda_1))]]\psi_{1j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\lambda_1^k\eta_{t-k}$$

$$- \sum_{j=1}^{2}\frac{\vartheta_j^2(\rho-\lambda_1)(1-\rho\lambda_1)\psi_{1j}}{\rho^2(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\vartheta_j^k\eta_{t-k}$$

$$\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t v_{t+k} = \frac{1}{1-\rho\beta}\sum_{k=0}^{\infty}\rho^k\eta_{t-k} - \frac{\lambda_1}{\rho(1-\rho\beta)}\sum_{k=0}^{\infty}\lambda_1^k\eta_{t-k}$$

$$\sum_{k=0}^{\infty}\beta^k\overline{\mathbb{E}}_t^c\pi_{t+k+1} = \rho\frac{\psi_{21}+\psi_{22}}{1-\rho\beta}\sum_{k=0}^{\infty}\rho^k\eta_{t-k}$$

$$+ \frac{\lambda_1}{\rho^2(1-\rho\beta)}\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\rho\lambda_1(1+\rho\beta\vartheta_j^2)-\vartheta_j(\rho-\lambda_1(1-\rho(\beta+\lambda_1)))]\psi_{2j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\lambda_1^k\eta_{t-k}$$

$$-\sum_{j=1}^{2}\frac{\vartheta_j^3(\rho-\lambda_1)(1-\rho\lambda_1)\psi_{2j}}{\rho^2(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\sum_{k=0}^{\infty}\vartheta_j^k\eta_{t-k}$$

Therefore, the PE share μ_τ is given by

$$\begin{aligned}\mu_\tau &= \frac{\partial \mathrm{PE}_\tau/\partial\eta_t}{\partial y_\tau/\partial\eta_t}\\ &= -\frac{\beta}{\sigma}(1-\lambda)\rho\frac{\delta_1\rho^\tau+\delta_2\lambda_1^\tau+\delta_3\vartheta_1^\tau+\delta_4\vartheta_2^\tau}{\psi_{11}(\rho^{\tau+1}-\vartheta_1^{\tau+1})+\psi_{12}(\rho^{\tau+1}-\vartheta_2^{\tau+1})}\end{aligned}$$

where

$$\begin{aligned}\delta_1 &= \frac{1+\phi_y(\psi_{11}+\psi_{12})+(\phi_\pi-\rho)(\psi_{21}+\psi_{22})}{1-\rho\beta}\\ \delta_2 &= \frac{\lambda_1}{\rho^2(1-\rho\beta)}\Bigg\{-\rho+\phi_y\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\lambda_1-\rho\vartheta_j[\beta+\lambda_1(1-\beta(\rho+\vartheta_j-\lambda_1))]]\psi_{1j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\\ &\quad+\phi_\pi\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\lambda_1-\rho\vartheta_j[\beta+\lambda_1(1-\beta(\rho+\vartheta_j-\lambda_1))]]\psi_{2j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\\ &\quad-\sum_{j=1}^{2}\frac{(\rho-\vartheta_j)[\rho\lambda_1(1+\rho\beta\vartheta_j^2)-\vartheta_j(\rho-\lambda_1(1-\rho(\beta+\lambda_1)))]\psi_{2j}}{(1-\beta\vartheta_j)(\vartheta_j-\lambda_1)(1-\vartheta_j\lambda_1)}\\ \delta_3 &= -\frac{\vartheta_1^2(\rho-\lambda_1)(1-\rho\lambda_1)[\phi_y\psi_{11}+(\phi_\pi-\vartheta_1)\psi_{21}]}{\rho^2(1-\beta\vartheta_1)(\vartheta_1-\lambda_1)(1-\vartheta_1\lambda_1)}\\ \delta_4 &= -\frac{\vartheta_2^2(\rho-\lambda_1)(1-\rho\lambda_1)[\phi_y\psi_{12}+(\phi_\pi-\vartheta_2)\psi_{22}]}{\rho^2(1-\beta\vartheta_2)(\vartheta_2-\lambda_1)(1-\vartheta_2\lambda_1)}\end{aligned}$$

□

Proof of Proposition 9. In order to study forward guidance it is convenient to write the model *as if* there is full information. To see this, let us first recall the dynamics in the standard model. In the benchmark NK model the Phillips curve is given by

$$\pi_t = \kappa y_t + \beta\mathbb{E}_t\pi_{t+1} \tag{A.49}$$

the DIS curve is given by (2.14), and the Taylor rule is given by (2.10)-(2.11). Inserting the Taylor rule into the DIS curve, one can write the model as a system of two first-order stochastic difference equations

$$\tilde{A}\boldsymbol{a}_t = \tilde{B}\mathbb{E}_t\boldsymbol{a}_{t+1} + \tilde{C}v_t \tag{A.50}$$

where $\boldsymbol{a}_t = [y_t \quad \pi_t]'$ is a 2×1 vector containing output and inflation, $\tilde{A}$ is a 2×2 coefficient matrix, $\tilde{B}$ is a 2×2 coefficient matrix and $\tilde{C}$ is a 2×1 vector satisfying

$$\tilde{A} = \begin{bmatrix} \varsigma + \phi_y & \phi_\pi \\ -\kappa & 1 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \varsigma & \delta \\ 0 & \beta \end{bmatrix}, \quad \text{and} \quad \tilde{C} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

Premultiplying the system by $\tilde{A}^{-1}$ we obtain

$$\boldsymbol{a}_t = \overline{\boldsymbol{\varphi}} v_t + \overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1} \tag{A.51}$$

where $\overline{\boldsymbol{\delta}} = \tilde{A}^{-1}\tilde{B}$ and $\overline{\boldsymbol{\varphi}} = \tilde{A}^{-1}\tilde{C}$.

Notice that the DI DIS and Phillips curves, (2.13) and (2.9), involve higher-order beliefs. Angeletos and Huo (2018) show that the following system dynamics under FIRE mimic the dynamics of our DI model

$$\boldsymbol{a}_t = \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1} + \overline{\boldsymbol{\varphi}} v_t \tag{A.52}$$

Anchoring is obtained by including an ad-hoc term $\boldsymbol{\omega_b a}_{t-1}$, where $\boldsymbol{\omega_b}$ is a 2×2 matrix. This element will induce inertia in the model dynamics, as we see in the data. Myopia is obtained by including an ad-hoc term $\boldsymbol{\omega_f}$ interacting with expected future outcomes. Formally, we replace $\overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1}$ by $\boldsymbol{\omega_f} \overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1}$. Here $\boldsymbol{\omega_f}$ is a 2×2 matrix that induces myopia in the model dynamics as long as its spectral radius lies within the unit circle. Effectively, this element is reducing the forward-looking behavior in the model. As we will see in the next sections, this is key for our findings on the Taylor Principle or curing the FGP.

To show that the ad-hoc model presented above captures our HANK beyond FIRE under certain $(\boldsymbol{\omega}_f, \boldsymbol{\omega}_b)$, we rely on the Method for Undetermined Coefficients. The ad-hoc behavioral dynamics (A.52) and the HANK beyond FIRE dynamics (3.5) are observationally equivalent if

$$\begin{aligned}
A\boldsymbol{a}_{t-1} + B\xi_t &= \overline{\boldsymbol{\varphi}}\xi_t + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} \mathbb{E}_t \boldsymbol{a}_{t+1} + \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} \\
&= \overline{\boldsymbol{\varphi}}\xi_t + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} \mathbb{E}_t (A\boldsymbol{a}_t + B\xi_{t+1}) + \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} \\
&= \overline{\boldsymbol{\varphi}}\xi_t + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} (A\boldsymbol{a}_t + B\mathbb{E}_t \xi_{t+1}) + \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} \\
&= \overline{\boldsymbol{\varphi}}\xi_t + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} (A\boldsymbol{a}_t + B\rho\xi_t) + \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} \\
&= \overline{\boldsymbol{\varphi}}\xi_t + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} [A(A\boldsymbol{a}_{t-1} + B\xi_t) + B\rho\xi_t] + \boldsymbol{\omega}_b \boldsymbol{a}_{t-1} \\
&= \left[\overline{\boldsymbol{\varphi}} + \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} (A + \rho) B\right] \xi_t + \left[\boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} AA + \boldsymbol{\omega}_b\right] \boldsymbol{a}_{t-1}
\end{aligned}$$

They are thus equivalent if

$$\boldsymbol{\omega}_b = [I - \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} A] A$$
$$B - \overline{\boldsymbol{\varphi}} = \boldsymbol{\omega}_f \overline{\boldsymbol{\delta}} (A + \rho) B \tag{A.53}$$

Now that we have the system dynamics (A.52), we just need to multiply the system by $\tilde{A}$ to back out the DI DIS and DI Phillips curves, which we can write as

$$\tilde{y}_t = \frac{\omega_{yy}}{\varsigma} \tilde{y}_{t-1} + \frac{\omega_{y\pi}}{\varsigma} \pi_{t-1} - \frac{1}{\varsigma}(i_t - \mathbb{E}_t \pi_{t+1}) + \frac{\delta_{yy}}{\varsigma} \mathbb{E}_t \tilde{y}_{t+1} + \frac{\delta_{y\pi} - 1}{\varsigma} \mathbb{E}_t \pi_{t+1} \tag{A.54}$$
$$\pi_t = \omega_{\pi y} \tilde{y}_{t-1} + \omega_{\pi\pi} \pi_{t-1} + \kappa y_t + \delta_{\pi y} \mathbb{E}_t \tilde{y}_{t+1} + \delta_{\pi\pi} \mathbb{E}_t \pi_{t+1} \tag{A.55}$$

where

$$\omega_{yy} = (\varsigma + \phi_y)\omega_{b,11} + \phi_\pi \omega_{b,21}$$
$$\omega_{y\pi} = (\varsigma + \phi_y)\omega_{b,12} + \phi_\pi \omega_{b,22}$$
$$\omega_{\pi y} = \omega_{b,21} - \kappa \omega_{b,11}$$
$$\omega_{\pi\pi} = \omega_{b,22} - \kappa \omega_{b,12}$$
$$\delta_{yy} = \frac{\varsigma}{\varsigma + \phi_y + \kappa \phi_\pi} \left[(\sigma + \phi_y)(\omega_{f,11} + \kappa \omega_{f,12}) + \phi_\pi (\omega_{f,21} + \kappa \omega_{f,22}) \right]$$
$$\delta_{y\pi} = \frac{1}{\sigma + \phi_y + \kappa \phi_\pi} \left\{ (\delta - \beta \phi_\pi)[(\sigma + \phi_y)\omega_{f,11} + \phi_\pi \omega_{f,21}] + [\delta\kappa + \beta(\sigma + \phi_y)][(\sigma + \phi_y)\omega_{f,12} + \phi_\pi \omega_{f,22})] \right\}$$
$$\delta_{\pi y} = \frac{\varsigma}{\varsigma + \phi_y + \kappa \phi_\pi} \left[(\omega_{f,21} - \kappa \omega_{f,11}) + \kappa \left(\omega_{f,22} - \kappa \omega_{f,12} \right) \right]$$
$$\delta_{\pi\pi} = \frac{1}{\varsigma + \phi_y + \kappa \phi_\pi} \left\{ (\delta - \beta \phi_\pi)(\omega_{f,21} - \kappa \omega_{f,11}) + \left[\delta\kappa + \beta(\sigma + \phi_y) \right] \left(\omega_{f,22} - \kappa \omega_{f,12} \right) \right\}$$

In order to analyze the effects of forward guidance in our HANK beyond FIRE framework, consider a situation in which the economy is stuck in a liquidity trap. Suppose that the zero lower bound (ZLB) for nominal interest rates is binding between periods t and τ, such that $\tau \geq t$. During the ZLB period, nominal interest rates are against the constraint, $i_k = 0$ for $k \in (t, \tau)$, and thus the ex-ante real interest rate is the (log) inverse of expected inflation, $\mathbb{E}_t r_k = -\mathbb{E}_t \pi_{k+1}$. In this case, the DIS curve (A.54) becomes

$$\tilde{y}_t = \frac{\omega_{yy}}{\varsigma} \tilde{y}_{t-1} + \frac{\omega_{y\pi}}{\varsigma} \pi_{t-1} - \frac{\delta_{y\pi}}{\varsigma} \mathbb{E}_t r_t + \frac{\delta_{yy}}{\varsigma} \mathbb{E}_t \tilde{y}_{t+1} \tag{A.56}$$

Using the lag operator, we can factorize (A.56)

$$\begin{aligned}\mathbb{E}_t\left(\delta_{y\pi}r_t-\omega_{y\pi}\pi_{t-1}\right)&=\mathbb{E}_t\left[\left(\delta_{yy}L^{-2}-\varsigma L^{-1}+\omega_{yy}\right)\tilde{y}_{t-1}\right]\\&=\mathbb{E}_t\left[\delta_{yy}\left(L^{-1}-\gamma_1\right)\left(L^{-1}-\gamma_2\right)\tilde{y}_{t-1}\right]\end{aligned}$$

where γ_1 and γ_2 are the roots of the polynomial $\mathcal{P}(x)\equiv\delta_{yy}x^{-2}-\varsigma x^{-1}+\omega_{yy}$, with the two roots satisfying $|\gamma_1|<1$ and $|\gamma_2|>1$. Dividing both sides by $(L^{-1}-\gamma_1)$

$$\begin{aligned}\delta_{yy}\mathbb{E}_t[(L^{-1}-\gamma_2)\tilde{y}_{t-1}]&=\mathbb{E}_t\left(-\delta_{y\pi}\frac{1}{\gamma_2-L^{-1}}r_t+\omega_{y\pi}\frac{1}{\gamma_2-L^{-1}}\pi_{t-1}\right)\\&=\mathbb{E}_t\left(-\delta_{y\pi}\frac{\gamma_2^{-1}}{1-(\gamma_2L)^{-1}}r_t+\omega_{y\pi}\frac{\gamma_2^{-1}}{1-(\gamma_2L)^{-1}}\pi_{t-1}\right)\end{aligned}$$

Hence, we can write the dynamics as

$$\begin{aligned}\tilde{y}_t&=\gamma_1\tilde{y}_{t-1}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2}\pi_{t-1}-\frac{\delta_{y\pi}}{\delta_{yy}\gamma_2}\sum_{k=0}^{\infty}\left(\frac{1}{\gamma_2}\right)^k\mathbb{E}_tr_{t+k}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2^2}\sum_{k=0}^{\infty}\left(\frac{1}{\gamma_2}\right)^k\mathbb{E}_t\pi_{t+k-1}\\&=\gamma_1\tilde{y}_{t-1}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2}\pi_{t-1}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2}r_{t-2}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2^2}r_{t-1}-\left(\frac{\delta_{y\pi}}{\delta_{yy}\gamma_2}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2^3}\right)\sum_{k=0}^{\infty}\left(\frac{1}{\gamma_2}\right)^k\mathbb{E}_tr_{t+k}\end{aligned}\tag{A.57}$$

Therefore, the effect of a forward guidance shock promised at time t in period τ is

$$FG_{t,t+\tau}=\frac{\partial\tilde{y}_t}{\partial\mathbb{E}_tr_{t+\tau}}=-\left(\frac{\delta_{y\pi}}{\delta_{yy}\gamma_2}+\frac{\omega_{y\pi}}{\delta_{yy}\gamma_2^3}\right)\frac{1}{\gamma_2^{\tau}}\tag{A.58}$$

which is decreasing in τ provided that $|\gamma_2|>1$. Since γ_2^{-1} lies inside the unit circle, $\lim_{\tau\to\infty}FG_{t+\tau}=0$, and the forward guidance puzzle is solved.

□

Proof of Proposition 10. The proof is identical to the proof of Proposition 4, modulo the replacement of σ_g for σ_ϵ. In the public information case, the individual action is given by

$$\begin{aligned}a_{igt}&=h_g(L)z_t\\&=h_g(L)(\xi_t+\epsilon_t)\end{aligned}$$

The policy function of an agent in group g is given by

$$h_g(z) = \frac{\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z}{(1-\theta_1 z)(1-\theta_2 z)}$$

and hence we have

$$\begin{aligned}
a_{gt} &= h_g(L)(\xi_t + \epsilon_t) \\
&= \frac{\widetilde{\psi}_{g1} + \widetilde{\psi}_{g2} z}{(1-\theta_1 z)(1-\theta_2 z)}(\xi_t + \epsilon_t) \\
&= \psi_{g1}\left(1 - \frac{\theta_1}{\rho}\right)\frac{1}{1-\theta_1 L}(\xi_t + \epsilon_t) + \psi_{g2}\left(1 - \frac{\theta_2}{\rho}\right)\frac{1}{1-\theta_2 L}(\xi_t + \epsilon_t) \\
&= \psi_{g1}\widetilde{\theta}_{1t} + \psi_{g2}\widetilde{\theta}_{2t}
\end{aligned}$$

We can write

$$\begin{aligned}
\boldsymbol{a}_t &= \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \\
&= Q\widetilde{\theta}_t \\
&= \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix} \begin{bmatrix} \widetilde{\theta}_{1t} \\ \widetilde{\theta}_{2t} \end{bmatrix} \\
&= \begin{bmatrix} \psi_{11}\widetilde{\theta}_{1t} + \psi_{12}\widetilde{\theta}_{2t} \\ \psi_{21}\widetilde{\theta}_{1t} + \psi_{22}\widetilde{\theta}_{2t} \end{bmatrix}
\end{aligned}$$

Notice that we can write

$$\begin{aligned}
\widetilde{\theta}_{1t}(1-\theta_1 L) &= \left(1 - \frac{\theta_1}{\rho}\right)(\xi_t + \epsilon_t) \implies \widetilde{\theta}_{1t} = \theta_1 \widetilde{\theta}_{1t-1} + \left(1 - \frac{\theta_1}{\rho}\right)(\xi_t + \epsilon_t) \\
\widetilde{\theta}_{2t}(1-\theta_2 L) &= \left(1 - \frac{\theta_2}{\rho}\right)(\xi_t + \epsilon_t) \implies \widetilde{\theta}_{2t} = \theta_2 \widetilde{\theta}_{2t-1} + \left(1 - \frac{\theta_2}{\rho}\right)(\xi_t + \epsilon_t)
\end{aligned}$$

Which we can write as a system as

$$\widetilde{\theta}_t = \Lambda \widetilde{\theta}_{t-1} + \Gamma(\xi_t + \epsilon_t)$$

where

$$\Lambda = \begin{bmatrix} \theta_1 & 0 \\ 0 & \theta_2 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 1 - \frac{\theta_1}{\rho} \\ 1 - \frac{\theta_2}{\rho} \end{bmatrix}$$

Hence, we can write

$$
\begin{aligned}
\boldsymbol{a}_t &= Q\widetilde{\theta}_t \\
&= Q[\Lambda\widetilde{\theta}_{t-1} + \Gamma(\xi_t + \epsilon_t)] \\
&= Q\Lambda\widetilde{\theta}_{t-1} + Q\Gamma(\xi_t + \epsilon_t) \\
&= Q\Lambda Q^{-1}\boldsymbol{a}_{t-1} + Q\Gamma(\xi_t + \epsilon_t) \\
&= A\boldsymbol{a}_{t-1} + B\xi_t + B\epsilon_t
\end{aligned} \tag{A.59}
$$

□

Proof of Proposition 11. This proof mimicks the proof of Proposition 4, and extends it to allow for a public signal. In a similar way, we can write the aggregate action for households or firms as

$$
a_{gt} = \varphi_g \sum_{k=0}^{\infty} \beta_g^k \overline{\mathbb{E}}_{gt}\xi_{t+k} + \gamma_{g1}\overline{\mathbb{E}}_{gt}a_{1t} + (\beta_g\gamma_{g1} + \alpha_{g1}) \sum_{k=0}^{\infty} \beta_g^k \overline{\mathbb{E}}_{gt}a_{1t+k+1} + \gamma_{g2}\overline{\mathbb{E}}_{gt}a_{2t} + (\beta_g\gamma_{g2} + \alpha_{g2})) \sum_{k=0}^{\infty} \beta_g^k \overline{\mathbb{E}}_{gt}a_{2t+k+1} \tag{A.60}
$$

where $a_{1t} = y_t$, $a_{2t} = \pi_t$, $\xi_t = v_t$, $\overline{\mathbb{E}}_{1t}(\cdot) = \overline{\mathbb{E}}_{ct}(\cdot)$, $\overline{\mathbb{E}}_{2t}(\cdot) = \overline{\mathbb{E}}_{\pi t}(\cdot)$ and the following parametric restrictions are satisfied

$$
\begin{aligned}
\varphi_1 &= -\frac{1}{\varsigma} & \varphi_2 &= 0 \\
\beta_1 &= \beta & \beta_2 &= \beta\theta \\
\gamma_{11} &= -\frac{\phi_y}{\varsigma} & \gamma_{21} &= \kappa\theta \\
\gamma_{12} &= -\frac{\phi_\pi}{\varsigma} & \gamma_{22} &= 1-\theta \\
\alpha_{11} &= \delta - \beta & \alpha_{21} &= 0 \\
\alpha_{12} &= \frac{1}{\varsigma} & \alpha_{22} &= 0
\end{aligned}
$$

The best response of agent i in group g is specified as follows

$$
a_{igt} = \varphi_g \mathbb{E}_{igt}\xi_t + \beta_g \mathbb{E}_{igt}a_{igt+1} + \sum_{j=1}^{2} \gamma_{gj}\mathbb{E}_{igt}a_{jt} + \sum_{j=1}^{2} \alpha_{gj}\mathbb{E}_{igt}a_{jt+1} \tag{A.61}
$$

where a_{-gt} is the aggregate action of the other group at time t. Parameters $\{\beta_g\}$, $\{\gamma_{gk}\}$, $\{\alpha_{gk}\}$ help parameterize PE and GE considerations. Notice that GE effects run not only within groups but also across groups (the interaction of the two blocks of the NK model). Parameters $\{\varphi_g\}$ capture the direct exposure of group g to the exogenous shock.

Let $\boldsymbol{a}_t = (a_{gt})$ be a column vector collecting the aggregate actions of all groups (e.g., the vector of aggregate consumption and aggregate inflation)

$$\boldsymbol{a}_t = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Let $\boldsymbol{\varphi} = (\varphi_g)$ be a column vector containing the value of φ_g across groups

$$\boldsymbol{\varphi} = \begin{bmatrix} \varphi_1 \\ \varphi_2 \end{bmatrix}$$

Let $\boldsymbol{\beta} = \text{diag}(\beta_g)$ be a 2×2 diagonal matrix of discount factors, with off-diagonal elements equal to 0.

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{bmatrix}$$

Let $\boldsymbol{\gamma}$ be a 2×2 matrix collecting the (contemporaneous) interaction parameters γ_{gj}

$$\boldsymbol{\gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$$

Let $\boldsymbol{\alpha} = (\alpha_{gk})$ be a 2×2 matrix collecting the (future) interaction parameters α_{gj}

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}$$

Finally, let $\boldsymbol{\delta} \equiv \boldsymbol{\beta} + \boldsymbol{\alpha}$,

$$\boldsymbol{\delta} = \begin{bmatrix} \beta_1 + \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \beta_2 + \alpha_{22} \end{bmatrix}$$

Let us now have a look at the fundamental representation of the signal process. We know that

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

$$
\begin{aligned}
&= \frac{1}{1-\rho L}\eta_t, && \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \\
x_{igt} &= \xi_t + u_{igt}, && u_{igt} \sim \mathcal{N}(0, \sigma_g^2) \\
z_t &= \xi_t + \epsilon_t, && \epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)
\end{aligned}
$$

Notice that the signal process admits the following state-space representation

$$
\begin{aligned}
Z_t &= F Z_{t-1} + \Phi \widehat{\boldsymbol{s}}_{igt} \\
X_t &= H Z_t + \Psi \widehat{\boldsymbol{s}}_{igt}
\end{aligned}
$$

with $F = \rho$, $\Phi = \begin{bmatrix} 0 & 0 & \sigma_\eta \end{bmatrix}$, $Z_t = \xi_t$, $H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\Psi = \begin{bmatrix} \sigma_\epsilon & 0 & 0 \\ 0 & \sigma_g & 0 \end{bmatrix}$ and $X_t = \begin{bmatrix} z_t & x_{igt} \end{bmatrix}'$. Define $\tau_\eta \equiv \frac{1}{\sigma_\eta^2}$, $\tau_g \equiv \frac{1}{\sigma_g^2}$ and $\tau_\epsilon \equiv \frac{1}{\sigma_\epsilon^2}$. The signal system can be written as

$$
\begin{aligned}
X_t &= \begin{bmatrix} \tau_\epsilon^{-1/2} & 0 & \frac{\tau_\eta^{-1/2}}{1-\rho L} \\ 0 & \tau_g^{-1/2} & \frac{\tau_\eta^{-1/2}}{1-\rho L} \end{bmatrix} \begin{bmatrix} \widehat{\eta}_t \\ \widehat{u}_{igt} \end{bmatrix} \\
&= \boldsymbol{M}_g(L) \widehat{\boldsymbol{s}}_{igt}, \qquad \widehat{\boldsymbol{s}}_{igt} \sim \mathcal{N}(0, I)
\end{aligned}
$$

Denote λ_g as the inside root of $\det[\boldsymbol{M}_g(L)\boldsymbol{M}_g'(L)]$, which is given by

$$
\lambda_g = \frac{1}{2}\left[\frac{1}{\rho} + \rho + \frac{\tau_g + \tau_\epsilon}{\rho\tau_\eta} - \sqrt{\left(\frac{1}{\rho} + \rho + \frac{\tau_g + \tau_\epsilon}{\rho\tau_\eta}\right)^2 - 4}\right] \tag{A.62}
$$

We can also write

$$
\begin{aligned}
V_g^{-1} &= \frac{\tau_g \tau_\epsilon}{\rho\tau_\eta(\tau_g + \tau_\epsilon)} \begin{bmatrix} \frac{\rho\tau_g + \lambda_g\tau_\epsilon}{\tau_g} & \lambda_g - \rho \\ \lambda_g - \rho & \frac{\lambda_g\tau_g + \rho\tau_\epsilon}{\tau_\epsilon} \end{bmatrix} \\
B_g(L)^{-1} &= \frac{1}{1-\lambda_g L} \begin{bmatrix} 1 - \frac{\lambda_g\tau_g + \rho\tau_\epsilon}{\tau_g + \tau_\epsilon} L & \frac{\tau_g(\lambda_g - \rho)}{\tau_g + \tau_\epsilon} L \\ \frac{\tau_\epsilon(\lambda_g - \rho)}{\tau_g + \tau_\epsilon} L & 1 - \frac{\rho\tau_g + \lambda_g\tau_\epsilon}{\tau_g + \tau_\epsilon} L \end{bmatrix}
\end{aligned}
$$

Let us now move to the forecasting part. The forecast of a random variable f_t

$$
f_t = A(L)\widehat{\boldsymbol{s}}_t
$$

can be obtained using the Wiener-Hopf prediction filter

$$\mathbb{E}_{it} f_t = \left[A(L)M'(L^{-1})B(L^{-1})^{-1}\right]_+ B(L)^{-1} x_{it}$$

Based on this result, we can solve the model. Denote agent i in group g policy function

$$a_{igt} = h_{g1}(L) z_t + h_{g2}(L) x_{igt}$$

(in this model, agents only observe signals. As a result, the policy function can only depend on current and past private and public signals). The aggregate outcome in group g can then be expressed as follows

$$\begin{aligned}
a_{gt} &= \int a_{igt}\, di \\
&= \int h_{g1}(L) z_t + h_{g2}(L) x_{igt}\, di \\
&= \int h_{g1}(L) \left(\frac{\sigma_\eta}{1-\rho L} \widehat{\eta}_t + \sigma_\epsilon \widehat{\epsilon}_t \right) + h_{g2}(L) \left(\frac{\sigma_\eta}{1-\rho L} \widehat{\eta}_t + \sigma_g \widehat{u}_{igt} \right) di \\
&= [h_{g1}(L) + h_{g2}(L)] \frac{\sigma_\eta}{1-\rho L} \widehat{\eta}_t + h_{g1}(L) \sigma_\epsilon \widehat{\epsilon}_t
\end{aligned}$$

Let us now obtain the forecasts. Recall that, for

$$f_t = A(L) \widehat{\boldsymbol{s}}_t = \frac{a(L)}{\prod_{\tau=1}^{d} (L - \beta_\tau)} \widehat{\boldsymbol{s}}_t$$

The optimal forecast is given by

$$\begin{aligned}
\mathbb{E}_{it} f_t &= \left[A(L)M'(L^{-1})B'(L^{-1})^{-1}\right]_+ V^{-1} B(L)^{-1} x_t \\
&= \frac{a(L)}{\prod_{\tau=1}^{d}(L-\beta_\tau)} M'(L^{-1}) \rho_{xx}(L)^{-1} x_t - \sum_{k=1}^{u} \frac{a(\lambda_k) G(\lambda_k) V^{-1} B(L)^{-1}}{(L-\lambda_k) \prod_{\tau \neq k}^{u} (\lambda_k - \lambda_\tau) \prod_{\tau=1}^{d} (\lambda_k - \beta_\tau)} x_t - \\
&\quad - \sum_{k=1}^{d} \frac{a(\beta_k) G(\beta_k) V^{-1} B(L)^{-1}}{(L-\beta_k) \prod_{\tau=1}^{k} (\beta_k - \lambda_\tau) \prod_{\tau \neq k}^{d} (\beta_k - \beta_\tau)} x_t
\end{aligned}$$

Hence, applying this general example to our particular case

$$\mathbb{E}_{igt}\xi_t = \frac{\lambda_g}{\rho\tau_\eta(1-\lambda_g\rho)}\frac{1}{1-\lambda_g L}\begin{bmatrix}\tau_\epsilon & \tau_g\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix}$$

$$\mathbb{E}_{igt}a_{kt+1} = \begin{bmatrix}\frac{h_{k1}(L)}{L\tau_\eta} + h_{k2}(L)\frac{\lambda_g\tau_\epsilon}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2} & h_{k2}(L)\frac{\lambda_g\tau_g}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2}\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix} -$$

$$-\frac{\lambda_g(1-\rho L)h_{k2}(\lambda_g)}{(L-\lambda_g)(1-\lambda_g L)\rho(1-\rho\lambda_g)\tau_\eta^2}\begin{bmatrix}\tau_\epsilon & \tau_g\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix} -$$

$$-\frac{\lambda_g h_{k1}(0)}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2}\begin{bmatrix}\frac{(1-\lambda_g L)\tau_g+(1-\rho L)\tau_\epsilon}{L(\rho-\lambda_g)} & -\tau_g\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix}$$

$$\mathbb{E}_{igt}a_{kt} = \begin{bmatrix}\frac{h_{k1}(L)}{\tau_\eta} + h_{k2}(L)\frac{L\lambda_g\tau_\epsilon}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2} & h_{k2}(L)\frac{L\lambda_g\tau_g}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2}\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix} -$$

$$-\frac{\lambda_g^2(1-\rho L)h_{k2}(\lambda_g)}{(L-\lambda_g)(1-\lambda_g L)\rho(1-\rho\lambda_g)\tau_\eta^2}\begin{bmatrix}\tau_\epsilon & \tau_g\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix}$$

$$\mathbb{E}_{igt}\left(a_{igt+1}-a_{gt+1}\right) = \frac{\lambda_g h_{g2}(L)}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2}\begin{bmatrix}-\tau_\epsilon & \frac{(L-\rho)(1-\rho L)\lambda_g\tau_g+(L-\lambda_g)(1-\lambda_g L)\rho\tau_\epsilon}{L(\rho-\lambda_g)(1-\rho\lambda_g)}\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix} -$$

$$-\frac{\lambda_g(1-\rho L)h_{g2}(\lambda_g)}{(L-\lambda_g)(1-\lambda_g L)\rho(1-\rho\lambda_g)\tau_\eta^2}\begin{bmatrix}-\tau_\epsilon & \tau_g\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix} -$$

$$-\frac{\lambda_g h_{g2}(0)}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2}\begin{bmatrix}-\tau_\epsilon & \frac{(1-\rho L)\tau_g+(1-\lambda_g L)\tau_\epsilon}{L(\rho-\lambda_g)}\end{bmatrix}\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix}$$

Recall the best response for agent i in group g (A.61), which we rewrite for convenience

$$a_{igt} = \varphi_g\mathbb{E}_{igt}\xi_t + \beta_g\mathbb{E}_{igt}a_{igt+1} + \sum_{j=1}^{2}\gamma_{gj}\mathbb{E}_{igt}a_{jt} + \sum_{j=1}^{2}\alpha_{gj}\mathbb{E}_{igt}a_{jt+1}$$

Introducing the expectations just calculated, and rearranging terms,

$$\Bigg[h_{g1}(L)\left(1-\frac{\beta_g}{L\tau_\eta}\right) - \sum_{k=1}^{2}\frac{h_{k1}(L)}{\tau_\eta}\left(\gamma_{gk}+\frac{\alpha_{gk}}{L}\right) - \sum_{k=1}^{2}\frac{h_{k2}(L)\lambda_g\tau_\epsilon}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2}\left(\gamma_{gk}L+\alpha_{gk}\right),$$

$$h_{g2}(L)\left(1-\frac{\beta_g}{L\tau_\eta}\right) - \sum_{k=1}^{2}\frac{h_{k2}(L)\lambda_g\tau_g}{(L-\lambda_g)(1-\lambda_g L)\rho\tau_\eta^2}\left(\gamma_{gk}L+\alpha_{gk}\right)\Bigg]\begin{bmatrix}z_t \\ x_{igt}\end{bmatrix} =$$

$$= \Bigg[\frac{\varphi_g\lambda_g\tau_\epsilon}{\rho\tau_\eta(1-\rho\lambda_g)(1-\lambda_g L)} - h_{g1}(0)\frac{\beta_g\lambda_g[(1-\lambda_g L)\tau_g+(1-\rho L)\tau_\epsilon]}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2 L(\rho-\lambda_g)} + h_{g2}(0)\frac{\beta_g\lambda_g\tau_\epsilon}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2} -$$

$$-\sum_{k=1}^{2} h_{k1}(0)\frac{\alpha_{gk}\lambda_g[(1-\lambda_g L)\tau_g+(1-\rho L)\tau_\epsilon]}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2 L(\rho-\lambda_g)} - \sum_{k=1}^{2} h_{k2}(\lambda_g)\frac{\lambda_g(1-\rho L)\tau_\epsilon}{(L-\lambda_g)(1-\lambda_g L)\rho(1-\rho\lambda_g)\tau_\eta^2}(\alpha_{gk}+\lambda_g\gamma$$

$$\frac{\varphi_g\lambda_g\tau_g}{\rho\tau_\eta(1-\rho\lambda_g)(1-\lambda_g L)} + h_{g1}(0)\frac{\beta_g\lambda_g\tau_g}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2} - h_{g2}(0)\frac{\beta_g\lambda_g[(1-\rho L)\tau_g+(1-\lambda_g L)\tau_\epsilon]}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2 L(\rho-\lambda_g)} +$$

$$\left. + \sum_{k=1}^{2} h_{k1}(0)\frac{\alpha_{gk}\lambda_g\tau_g}{(1-\lambda_g L)(1-\rho\lambda_g)\tau_\eta^2} - \sum_{k=1}^{2} h_{k2}(\lambda_g)\frac{\lambda_g(1-\rho L)\tau_g}{(L-\lambda_g)(1-\lambda_g L)\rho(1-\rho\lambda_g)\tau_\eta^2}(\alpha_{gk}+\lambda_g\gamma_{gk})\right] \begin{bmatrix} z_t \\ x_{igt} \end{bmatrix}$$

We can write the above system of equations in terms of $\boldsymbol{h}(L)$ in matrix form

$$\boldsymbol{C}(L)\boldsymbol{h}(L) = \boldsymbol{d}(L) \tag{A.63}$$

where

$$C(L) = \begin{bmatrix} C_{11}(L) & C_{12}(L) & C_{13}(L) & C_{14}(L) \\ C_{21}(L) & C_{22}(L) & C_{23}(L) & C_{24}(L) \\ C_{31}(L) & C_{32}(L) & C_{33}(L) & C_{34}(L) \\ C_{41}(L) & C_{42}(L) & C_{43}(L) & C_{44}(L) \end{bmatrix}$$

$$h(L) = \begin{bmatrix} h_{11}(L) \\ h_{12}(L) \\ h_{21}(L) \\ h_{22}(L) \end{bmatrix}$$

$$D[L; h(\lambda), h(0)] = \begin{bmatrix} d_{11}(L) \\ d_{12}(L) \\ d_{21}(L) \\ d_{22}(L) \end{bmatrix}$$

where

$$C_{11}(L) = 1 - \frac{\beta_1 + \alpha_{11}}{L\tau_\eta} - \frac{\gamma_{11}}{\tau_\eta}$$

$$C_{12}(L) = -\frac{\lambda_1\tau_\epsilon(\alpha_{11} + \gamma_{11}L)}{(L-\lambda_1)(1-\lambda_1 L)\rho\tau_\eta^2}$$

$$C_{13}(L) = -\frac{\gamma_{12}}{\tau_\eta} - \frac{\alpha_{12}}{L\tau_\eta}$$

$$C_{14}(L) = -\frac{\lambda_1\tau_\epsilon(\alpha_{12} + \gamma_{12}L)}{(L-\lambda_1)(1-\lambda_1 L)\rho\tau_\eta^2}$$

$$C_{21}(L) = 0$$

$$C_{22}(L) = 1 - \frac{\beta_1}{L\tau_\eta} - \frac{\lambda_1\tau_1(\alpha_{11} + \gamma_{11}L)}{(L-\lambda_1)(1-\lambda_1 L)\rho\tau_\eta^2}$$

$$C_{23}(L) = 0$$

$$C_{24}(L) = -\frac{\lambda_1\tau_1(\alpha_{12} + \gamma_{12}L)}{(L-\lambda_1)(1-\lambda_1 L)\rho\tau_\eta^2}$$

$$C_{31}(L) = -\frac{\gamma_{21}}{\tau_\eta} - \frac{\alpha_{21}}{L\tau_\eta}$$

$$C_{32}(L) = -\frac{\lambda_2\tau_\epsilon(\alpha_{21} + \gamma_{21}L)}{(L-\lambda_2)(1-\lambda_2 L)\rho\tau_\eta^2}$$

$$C_{33}(L) = 1 - \frac{\beta_2 + \alpha_{22}}{L\tau_\eta} - \frac{\gamma_{22}}{\tau_\eta}$$

$$C_{34}(L) = -\frac{\lambda_2\tau_\epsilon(\alpha_{22} + \gamma_{22}L)}{(L-\lambda_2)(1-\lambda_2 L)\rho\tau_\eta^2}$$

$$C_{41}(L) = 0$$

$$C_{42}(L) = -\frac{\lambda_2\tau_2(\alpha_{21} + \gamma_{21}L)}{(L-\lambda_2)(1-\lambda_2 L)\rho\tau_\eta^2}$$

$$C_{43}(L) = 0$$

$$C_{44}(L) = 1 - \frac{\beta_2}{L\tau_\eta} - \frac{\lambda_2\tau_2(\alpha_{22} + \gamma_{22}L)}{(L-\lambda_2)(1-\lambda_2 L)\rho\tau_\eta^2}$$

and

$$\begin{aligned}
D_1(L) = {} & \frac{\varphi_1\lambda_1\tau_\epsilon}{\rho\tau_\eta(1-\rho\lambda_1)(1-\lambda_1 L)} - h_{11}(0)\frac{(\beta_1+\alpha_{11})\lambda_1[(1-\lambda_1 L)\tau_1 + (1-\rho L)\tau_\epsilon]}{(1-\lambda_1 L)(1-\rho\lambda_1)\tau_\eta^2 L(\rho-\lambda_1)} + \\
& + h_{12}(0)\frac{\beta_1\lambda_1\tau_\epsilon}{(1-\lambda_1 L)(1-\rho\lambda_1)\tau_\eta^2} - h_{21}(0)\frac{\alpha_{12}\lambda_1[(1-\lambda_1 L)\tau_1 + (1-\rho L)\tau_\epsilon]}{(1-\lambda_1 L)(1-\rho\lambda_1)\tau_\eta^2 L(\rho-\lambda_1)} - \\
& - [h_{12}(\lambda_1)(\alpha_{11} + \lambda_1\gamma_{11}) + h_{22}(\lambda_1)(\alpha_{12} + \lambda_1\gamma_{12})]\frac{\lambda_1(1-\rho L)\tau_\epsilon}{(L-\lambda_1)(1-\lambda_1 L)\rho\tau_\eta^2(1-\rho\lambda_1)}
\end{aligned}$$

$$\begin{aligned}
D_2(L) = {} & \frac{\varphi_1\lambda_1\tau_1}{\rho\tau_\eta(1-\rho\lambda_1)(1-\lambda_1 L)} + h_{11}(0)\frac{(\beta_1+\alpha_{11})\lambda_1\tau_1}{(1-\lambda_1 L)(1-\rho\lambda_1)\tau_\eta^2} - \\
& - h_{12}(0)\frac{\beta_1\lambda_1[(1-\rho L)\tau_1 + (1-\lambda_1 L)\tau_\epsilon]}{(1-\lambda_1 L)(1-\rho\lambda_1)\tau_\eta^2 L(\rho-\lambda_1)} + h_{21}(0)\frac{\alpha_{12}\lambda_1\tau_1}{(1-\lambda_1 L)(1-\rho\lambda_1)\tau_\eta^2} - \\
& - [h_{12}(\lambda_1)(\alpha_{11} + \lambda_1\gamma_{11}) + h_{22}(\lambda_1)(\alpha_{12} + \lambda_1\gamma_{12})]\frac{\lambda_1(1-\rho L)\tau_1}{(L-\lambda_1)(1-\lambda_1 L)\rho\tau_\eta^2(1-\rho\lambda_1)}
\end{aligned}$$

$$\begin{aligned}
D_3(L) = {} & \frac{\varphi_2\lambda_2\tau_\epsilon}{\rho\tau_\eta(1-\rho\lambda_2)(1-\lambda_2 L)} - h_{11}(0)\frac{\alpha_{21}\lambda_2[(1-\lambda_2 L)\tau_2 + (1-\rho L)\tau_\epsilon]}{(1-\lambda_2 L)(1-\rho\lambda_2)\tau_\eta^2 L(\rho-\lambda_2)} - \\
& - h_{21}(0)\frac{(\beta_2+\alpha_{22})\lambda_2[(1-\lambda_2 L)\tau_2 + (1-\rho L)\tau_\epsilon]}{(1-\lambda_2 L)(1-\rho\lambda_2)\tau_\eta^2 L(\rho-\lambda_2)} + h_{22}(0)\frac{\beta_2\lambda_2\tau_\epsilon}{(1-\lambda_2 L)(1-\rho\lambda_2)\tau_\eta^2} - \\
& - [h_{12}(\lambda_2)(\alpha_{21} + \lambda_2\gamma_{21}) + h_{22}(\lambda_2)(\alpha_{22} + \lambda_1 2\gamma_{22})]\frac{\lambda_2(1-\rho L)\tau_\epsilon}{(L-\lambda_2)(1-\lambda_2 L)\rho\tau_\eta^2(1-\rho\lambda_2)}
\end{aligned}$$

$$D_4(L) = \frac{\varphi_2\lambda_2\tau_2}{\rho\tau_\eta(1-\rho\lambda_2)(1-\lambda_2 L)} + h_{11}(0)\frac{\alpha_{21}\lambda_2\tau_2}{(1-\lambda_2 L)(1-\rho\lambda_2)\tau_\eta^2} + h_{21}(0)\frac{(\beta_2+\alpha_{22})\lambda_2\tau_2}{(1-\lambda_2 L)(1-\rho\lambda_2)\tau_\eta^2} -$$
$$- h_{22}(0)\frac{\beta_2\lambda_2[(1-\rho L)\tau_2 + (1-\lambda_2 L)\tau_\epsilon]}{(1-\lambda_2 L)(1-\rho\lambda_2)\tau_\eta^2 L(\rho-\lambda_2)} -$$
$$- [h_{12}(\lambda_2)(\alpha_{21}+\lambda_2\gamma_{21}) + h_{22}(\lambda_2)(\alpha_{22}+\lambda_1 2\gamma_{22})]\frac{\lambda_2(1-\rho L)\tau_2}{(L-\lambda_2)(1-\lambda_2 L)\rho\tau_\eta^2(1-\rho\lambda_2)}$$

From (A.63), the solution to the policy function is given by

$$\boldsymbol{h}(L) = \boldsymbol{C}(L)^{-1}\boldsymbol{d}(L) = \frac{\text{adj } \boldsymbol{C}(L)}{\det \boldsymbol{C}(L)}\boldsymbol{d}(L)$$

Hence, we need to obtain det $\boldsymbol{C}(L)$. Note that the degree of det $\boldsymbol{C}(L)$ is a polynomial of degree 8 on L. Denote the inside roots of det $\boldsymbol{C}(L)$ as $\{\zeta_1, \zeta_2, \zeta_3, \zeta_4, \zeta_5, \zeta_6\}$, and the outside roots as $\{\vartheta_1^{-1}, \vartheta_2^{-1}\}$. Because agents cannot use future signals, the inside roots have to be removed. Note that the number of free constants in $\boldsymbol{d}(L)$ is 6:

$$\Big\{h_{11}(0), h_{12}(0), h_{21}(0), h_{22}(0), \underbrace{h_{12}(\lambda_1)(\alpha_{11}+\lambda_1\gamma_{11}) + h_{22}(\lambda_1)(\alpha_{12}+\lambda_1\gamma_{12})}_{h(\lambda_1)},$$
$$\underbrace{h_{12}(\lambda_2)(\alpha_{21}+\lambda_2\gamma_{21}) + h_{22}(\lambda_2)(\alpha_{22}+\lambda_2\gamma_{22})}_{h(\lambda_2)}\Big\}$$

For a unique solution, it has to be the case that the number of outside roots is 2. By Cramer's rule, $h_{11}(L)$ is given by

$$h_{11}(L) = \frac{\det \begin{bmatrix} d_1(L) & C_{12}(L) & C_{13}(L) & C_{14}(L) \\ d_2(L) & C_{22}(L) & C_{23}(L) & C_{24}(L) \\ d_3(L) & C_{32}(L) & C_{33}(L) & C_{34}(L) \\ d_4(L) & C_{42}(L) & C_{43}(L) & C_{44}(L) \end{bmatrix}}{\det \boldsymbol{C}(L)}$$

and in a similar manner with the rest policy functions. The degree of the numerator is 7, as the highest degree of $D_g(L)$ is 1 degree less than that of $\boldsymbol{C}(L)$. By choosing the appropriate constants $\{h_{11}(0), h_{12}(0), h_{21}(0), h_{22}(0), h(\lambda_1), h(\lambda_2)\}$, the 6 inside roots will be removed. Therefore, the 6 constants are solutions to the following system of linear equations

$$\det \begin{bmatrix} d_1(\zeta_i) & C_{12}(\zeta_i) & C_{13}(\zeta_i) & C_{14}(\zeta_i) \\ d_2(\zeta_i) & C_{22}(\zeta_i) & C_{23}(\zeta_i) & C_{24}(\zeta_i) \\ d_3(\zeta_i) & C_{32}(\zeta_i) & C_{33}(\zeta_i) & C_{34}(\zeta_i) \\ d_4(\zeta_i) & C_{42}(\zeta_i) & C_{43}(\zeta_i) & C_{44}(\zeta_i) \end{bmatrix} = 0$$

for $i = 1, 2, ..., 6$. After removing the inside roots in the denominator, the degree of the numerator is 1 and the degree of the denominator is 2. The policy functions will be

$$h_{g1}(L) = \frac{\widetilde{\psi}_{g1,1} + \widetilde{\psi}_{g2,1}L}{(1-\vartheta_1 L)(1-\vartheta_2 L)}$$
$$h_{g2}(L) = \frac{\widetilde{\psi}_{g1,2} + \widetilde{\psi}_{g2,2}L}{(1-\vartheta_1 L)(1-\vartheta_2 L)}$$

and hence we have

$$\begin{aligned} a_{gt} &= [h_{g1}(L) + h_{g2}(L)]\xi_t + h_{g1}(L)\epsilon_t \\ &= \frac{(\widetilde{\psi}_{g1,1} + \widetilde{\psi}_{g1,2}) + (\widetilde{\psi}_{g2,1} + \widetilde{\psi}_{g2,2})L}{(1-\vartheta_1 L)(1-\vartheta_2 L)}\xi_t + \frac{\widetilde{\psi}_{g1,2} + \widetilde{\psi}_{g2,2}L}{(1-\vartheta_1 L)(1-\vartheta_2 L)}\epsilon_t \\ &= \psi_{g1}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 L}\xi_t + \psi_{g2}\left(1-\frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta_2 L}\xi_t + \phi_{g1}\left(1-\frac{\vartheta_1}{\rho}\right)\frac{1}{1-\vartheta_1 L}\epsilon_t + \phi_{g2}\left(1-\frac{\vartheta_2}{\rho}\right)\frac{1}{1-\vartheta} \\ &= \psi_{g1}\widetilde{\vartheta}^{\xi}_{1t} + \psi_{g2}\widetilde{\vartheta}^{\xi}_{2t} + \phi_{g1}\widetilde{\vartheta}^{\epsilon}_{1t} + \phi_{g2}\widetilde{\vartheta}^{\epsilon}_{2t} \end{aligned}$$

We can write

$$\begin{aligned} \boldsymbol{a}_t &= \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix} \\ &= Q_\xi \widetilde{\vartheta}^{\xi}_t + Q_\epsilon \widetilde{\vartheta}^{\epsilon}_t \\ &= \begin{bmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{bmatrix} \begin{bmatrix} \widetilde{\vartheta}^{\xi}_{1t} \\ \widetilde{\vartheta}^{\xi}_{2t} \end{bmatrix} + \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} \widetilde{\vartheta}^{\epsilon}_{1t} \\ \widetilde{\vartheta}^{\epsilon}_{2t} \end{bmatrix} \end{aligned}$$

Notice that we can write

$$\widetilde{\vartheta}^x_t = \Lambda \widetilde{\vartheta}^x_{t-1} + \Gamma x_t = (I - \Lambda L)^{-1}\Gamma x_t$$

for $x \in \{\xi, \epsilon\}$, where

$$\Lambda = \begin{bmatrix} \vartheta_1 & 0 \\ 0 & \vartheta_2 \end{bmatrix}, \qquad \Gamma = \begin{bmatrix} 1 - \frac{\vartheta_1}{\rho} \\ 1 - \frac{\vartheta_2}{\rho} \end{bmatrix}$$

Hence, we can write

$$\begin{aligned} \boldsymbol{a}_t &= Q_\xi (I - \Lambda L)^{-1} \Gamma \xi_t + Q_\epsilon (I - \Lambda L)^{-1} \Gamma \epsilon_t \\ &= Q_\xi \sum_{k=0}^{\infty} \Lambda^k \Gamma \xi_{t-k} + Q_\epsilon \sum_{k=0}^{\infty} \Lambda^k \Gamma \epsilon_{t-k} \end{aligned}$$

□

B Useful Mathematical Concepts

B.1 Wiener-Hopf Filter

Consider the non-causal prediction of $f_t = A(L)\widehat{\boldsymbol{s}}_{it}$ given the whole stream of signals

$$\begin{aligned} \mathbb{E}(f_t | x_i^\infty) &= \rho_{yx}(L) \rho_{xx}^{-1}(L) x_{it} \\ &= \rho_{yx}(L) \boldsymbol{B}(L^{-1})^{-1} \boldsymbol{V}^{-1} \boldsymbol{B}(L)^{-1} x_{it} \\ &= \rho_{yx}(L) \boldsymbol{B}(L^{-1})^{-1} \boldsymbol{V}^{-1} \boldsymbol{w}_{it} \\ &= \sum_{k=-\infty}^{\infty} h_k \boldsymbol{w}_{it-k} \end{aligned}$$

where $\rho_{yx}(z) = A(z)\boldsymbol{M}'(z^{-1})$ and $\rho_{xx}(z) = \boldsymbol{B}(z)\boldsymbol{V}\boldsymbol{B}'(z^{-1})$. Notice that we are using future values of $\boldsymbol{w}_{it}$. However, if the agent only observes events or signals up to time t, the best prediction is

$$\begin{aligned} \mathbb{E}(f_t | x_i^t) &= \left[\sum_{k=-\infty}^{\infty} h_k \boldsymbol{w}_{it-k} \right]_+ \\ &= \sum_{k=0}^{\infty} h_k \boldsymbol{w}_{it-k} \\ &= \left[\rho_{yx}(L) \boldsymbol{B}(L^{-1})^{-1} \right]_+ \boldsymbol{V}^{-1} \boldsymbol{w}_{it} \\ &= \left[\rho_{yx}(L) \boldsymbol{B}(L^{-1})^{-1} \right]_+ \boldsymbol{V}^{-1} \boldsymbol{B}(L)^{-1} x_{it} \end{aligned}$$

B.2 Annihilator Operator

The annihilator operator $[\cdot]_+$ eliminates the negative powers of the lag polynomial:

$$[A(z)]_+ = \left[\sum_{k=-\infty}^{\infty} a_k z^k \right]_+ = \sum_{k=0}^{\infty} a_k z^k$$

Suppose that we are interested in obtaining $[A(z)]_+$, where $A(z)$ takes this particular form, $A(z) = \frac{\phi(z)}{z-\lambda}$ with $|\lambda| < 1$, and $\phi(z)$ only contains positive powers of z. We can rewrite $A(z)$ as

$$A(z) = \frac{\phi(z) - \phi(\lambda)}{z - \lambda} + \frac{\phi(\lambda)}{z - \lambda}$$

Let us first have a look at the second term, We can write

$$\begin{aligned} \frac{\phi(\lambda)}{z - \lambda} &= -\frac{\phi(\lambda)}{\lambda} \frac{1}{1 - \lambda^{-1} z} \\ &= -\frac{\phi(\lambda)}{\lambda} (1 + \lambda^{-1} z + \lambda^{-2} z^2 + ...) \end{aligned}$$

which is not converging. Alternatively, we can write it as a converging series as

$$\begin{aligned} \frac{\phi(\lambda)}{z - \lambda} &= \phi(\lambda) z^{-1} \frac{1}{1 - \lambda z^{-1}} \\ &= \phi(\lambda) z^{-1} (1 + \lambda z^{-1} + \lambda^2 z^{-2} + ...) \end{aligned}$$

Notice that all the power terms are on the negative side of z. As a result,

$$\left[\frac{\phi(\lambda)}{z - \lambda} \right]_+ = 0$$

Let us now move to the first term. We can write

$$\begin{aligned} \phi(z) - \phi(\lambda) &= \sum_{k=0}^{\infty} \phi_k (z^k - \lambda^k) \\ &= \phi_0 \prod_{k=1}^{\infty} (z - \xi_k) \end{aligned}$$

where $\{\xi^k\}$ are the roots of this difference polynomial. Since we know that λ is a root of the

LHS, we can set $\xi^k = \lambda$ and write

$$\phi(z) - \phi(\lambda) = \phi_0(z-\lambda)\prod_{k=2}^{\infty}(z-\xi_k) \implies \frac{\phi(z)-\phi(\lambda)}{z-\lambda} = \prod_{k=2}^{\infty}(z-\xi_k)$$

which only contains positive powers of z. Hence, we have that

$$\left[\frac{\phi(z)}{z-\lambda}\right]_+ = \frac{\phi(z)-\phi(\lambda)}{z-\lambda}$$

Consider now instead the case $A(z) = \frac{\phi(z)}{(z-\lambda)(z-\beta)}$. Making use of partial fractions, we can write

$$\begin{aligned}\frac{\phi(z)}{(z-\lambda)(z-\beta)} &= \frac{1}{\lambda-\beta}\left[\frac{\phi(z)}{z-\lambda} - \frac{\phi(z)}{z-\beta}\right] \\ &= \frac{1}{\lambda-\beta}\left[\frac{\phi(z)-\phi(\lambda)}{z-\lambda} - \frac{\phi(z)-\phi(\beta)}{z-\beta} + \frac{\phi(\lambda)}{z-\lambda} - \frac{\phi(\beta)}{z-\beta}\right]\end{aligned}$$

Following the same steps as in the previous case, we can solve

$$\left[\frac{\phi(z)}{(z-\lambda)(z-\beta)}\right]_+ = \frac{\phi(z)-\phi(\lambda)}{(\lambda-\beta)(z-\lambda)} - \frac{\phi(z)-\phi(\beta)}{(\lambda-\beta)(z-\beta)}$$