Bias and Sensitivity under Ambiguity *

Zhen Huo† Marcelo Pedroni‡ Guangyu Pei§

July 17, 2023

Abstract

This paper characterizes the effects of ambiguity aversion under dispersed information. The equilibrium outcome is observationally equivalent to a Bayesian forecast of the fundamental with increased sensitivity to signals and a pessimistic bias. This equivalence result takes a simple form that accommodates dynamic information and strategic interactions. Applying the result, we show that ambiguity aversion helps rationalize the joint empirical pattern between the bias and persistence of inflation forecasts conditional on household income. In a policy game à la Barro and Gordon (1983) with ambiguity-averse agents, the policy rule features higher average inflation and increased responsiveness to fundamentals.

Keywords: Ambiguity, Incomplete information, Coordination, Bias
JEL classifications: E20, E32, F44

*We are grateful to Tan Wang for discussing our paper at China International Conference in Macroeconomics, and to Alex Kohlhas, Yueran Ma, Jianjun Miao, Yang Lu, Yulei Luo, Alessandro Pavan, and Michael Song for their feedback. We also acknowledge useful comments from seminar participants from CUHK, NUS, HKUST, HKU-UCL-ESRC Workshop, SHUFE, CityU of Hong Kong, AFR International Conference of Economics and Finance, Virtual Workshop on Imperfect Expectations and Macroeconomics, and China International Conference in Macroeconomics.

†Yale University; zhen.huo@yale.edu.
‡University of Amsterdam; m.pedroni@uva.nl.
§Chinese University of Hong Kong; guangyu.pei@cuhk.edu.hk.

1 Introduction

Workhorse macroeconomic models often assume that agents have perfect knowledge about the underlying data-generating process of the economy. However, households and firms often face model uncertainty or ambiguity[1] when making economic decisions. For example, after the pandemic, inflation has surged and average inflation targeting entered the policy discussion. It is less clear whether inflation will fluctuate around 2% or 4% in the near future, and each of these scenarios could be interpreted as a different model economy. Moreover, when economic decisions are interdependent, the coordination among market participants relies on their perceptions of how others would react to such uncertainty. In this environment, what are the macroeconomic effects of ambiguity? Can ambiguity help explain the deviations from rational expectations observed in survey data? How does the presence of ambiguity aversion in the private sector affect optimal policy design?

This paper makes two contributions. Theoretically, we show that despite the complexities associated with dispersed information, coordination motives, and ambiguity aversion, the equilibrium strategy is equivalent to a single-agent forecasting problem with two modifications: an amplified responsiveness to signals and a permanent pessimistic bias. This equivalence result applies to general information processes and takes a simple form, enabling the use of standard numerical algorithms to compute equilibrium strategies. On the applied side, we document that the bias in inflation forecasts decreases with household income, while the persistence of forecast errors increase with household income. This observed joint distribution can be naturally accounted for in a model where agents exhibit ambiguity aversion towards inflation shocks. In a policy game à la Barro and Gordon (1983), we show that policymakers respond to this behavioral pattern by increasing both their responsiveness to shocks and the unconditional level at which they set the inflation rate.

Framework. We consider an abstract Gaussian-quadratic economy. Agents' payoffs depend on an exogenous fundamental, their own actions, and the average actions of others, following Angeletos and Pavan (2007). The generic quadratic utility function can be regarded as the reduced form representation of a micro-founded economy that allows for general equilibrium (GE) effects. In addition, we impose minimal restrictions on the information structure that accommodates persistent learning and dispersed information. The main departure from the existing literature is that the fundamental is not only stochastic but also ambiguous in the sense that agents do not have perfect knowledge about its objective probability distribution. In the baseline specification, agents are ambiguity averse with preferences represented by the smooth model of ambiguity proposed by Klibanoff, Marinacci, and Mukerji (2005). We also extend the analysis to models with robust preferences (Hansen and Sargent, 2001a,b).

[1] According to Marinacci (2015), ambiguity refers to subjective uncertainty over probabilities due to a lack of ex-ante information to determine a specific model for the economy in the course of decision-making.

The interaction between ambiguity aversion and dispersed information makes the equilibrium difficult to characterize. It has been widely recognized that ambiguity-averse agents behave as if their beliefs about the fundamental are distorted (Ilut and Schneider, 2014; Bhandari, Borovička, and Ho, 2022) and that these distortions are affected by the equilibrium strategies. This results in a fixed-point problem between the strategies and beliefs of all agents in equilibrium. On top of that, the presence of imperfect coordination and persistent information leads to the infinite regress problem (Townsend, 1983), with higher-order beliefs potentially calling for an infinite-dimensional state space.

Equivalence results. Conceptually, the presence of ambiguity amounts to a more diffuse prior and ambiguity aversion makes the diffusion loom larger. As a result, agents divert attention from their priors to their signals, leading to a higher sensitivity to signals. Meanwhile, in the face of model uncertainty, ambiguity-averse agents are more concerned about models that yield lower payoffs and tend to place greater emphasis on adverse probability distributions. This consideration leads agents to behave as if they are pessimistically biased when making forecasts.

Our equivalence result formalizes this intuition and circumvents the aforementioned technical complexities. The equilibrium strategy is ultimately equivalent to that of a modified single-agent Bayesian forecasting the fundamental. We first introduce a (w, α)-modified signal process. In this auxiliary forecasting problem, the precision of idiosyncratic shocks is discounted by the degree of strategic complementarity, α, which captures the idea that the coordination motive reduces the importance of private information in inferring the aggregate outcome (Huo and Pedroni, 2020). In addition, the prior variance of the fundamental shock is amplified by w—an endogenous object that summarizes the extent to which the prior becomes more diffuse due to ambiguity aversion.

We establish that the equilibrium strategy coincides with the Bayesian forecasting rule using the (w, α)-modified signal process, with two notable adjustments governed by the endogenous variable w. First, there is an additional uniform overreaction to all signals. This overreaction distinguishes the equilibrium strategy from a Bayesian forecasting problem, resembling the departure from rationality implied by diagnostic expectations (Bordalo, Gennaioli, Ma, and Shleifer, 2020). Second, there is an additional bias that is independent of the signal realization, further differentiating the equilibrium allocation from its rational-expectations counterpart. To close the loop, we provide the condition that w needs to satisfy, which involves only unconditional moments about endogenous aggregate outcomes.

This result provides a concise summary of the effects of ambiguity aversion in a general equilibrium setting. The characterization enables the derivation of general comparative static results. For example, we show that an increase in the degree of strategic complementarity leads to greater bias. To understand the underlying intuition, one needs to invoke the fact that, with ambiguity aversion, equilibrium outcomes depend on *subjective* higher-order beliefs. With rational expectations, beliefs of higher order rely more on the common prior (Morris and Shin, 2002). With ambiguity aversion, the bias in others' beliefs is embedded in the common prior, which then accumulates as the order increases. A higher degree of strategic complementarity increases the relative weight of higher-order beliefs, thereby amplifying the bias.

Computationally, this result provides a tractable algorithm for computing the equilibrium strategies. In the absence of ambiguity, dynamic higher-order expectations require, in principle, the entire history of signals as state variables. However, a finite-state representation is possible when signals follow ARMA(p, q) processes (Woodford, 2003; Angeletos and La'O, 2010; Huo and Pedroni, 2020). Our results imply that a similar finite-state equilibrium representation is still possible even with ambiguity-averse agents. The original infinite-dimensional problem collapses to a one-dimensional problem to determine the endogenous amplifier of the prior variance, w. To solve the Bayesian forecasting problem with the (w, α)-modified signal process, the standard Kalman filter can be applied.

Survey evidence on inflation expectations. Our equivalence result helps explain the patterns observed in survey data regarding the joint behavior of bias and persistence in inflation forecasts. Using the Michigan Survey of Consumers (MSC) and the Survey of Consumer Expectations (SCE), we document that the bias and the persistence of inflation forecast errors jointly vary with household income levels. Specifically, higher-income households tend to exhibit lower forecast bias but more persistent forecast errors. This joint behavior of bias and persistence is at odds with the predictions of rational expectations models. Rationality would imply an average bias of zero, and the higher persistence of forecast errors of the rich would require that they have less precise information.

We demonstrate that the observed patterns naturally arise from a micro-founded optimal consumption problem in which households face ambiguity about the inflation process. With rigid nominal incomes, higher inflation erodes households' real purchasing power. Accordingly, ambiguity-averse households assign more weight to high-inflation models, resulting in an upward bias in inflation forecasts. At the same time, ambiguity aversion induces increased sensitivity to signals or a reduction in the weight put on the prior mean of inflation, which in turn reduces the persistence of forecast errors.

Importantly, households with lower nominal incomes are more exposed to variations in inflation and are more sensitive to ambiguous inflation shocks. In the cross-section of households, it is as if higher-income households were less ambiguity averse. Consequently, the inflation forecasts of the income-rich feature relatively lower levels of bias, lower sensitivity to signals, and higher persistence of forecast errors. With this rather parsimonious structure, our model matches the documented cross-sectional patterns reasonably well when calibrated to the data. Further, the predictions of the model are also broadly consistent with existing survey evidence on expectations, including under-reaction at the consensus level (Coibion and Gorodnichenko, 2015), over-reaction at the individual level (Bordalo, Gennaioli, Ma, and Shleifer, 2020), and delayed overshooting (Angeletos, Huo, and Sastry, 2021).

Optimal policy with ambiguity aversion. How should policy respond to the observed deviations from rational expectations? To shed light on this question, we explore a policy game à la Barro and Gordon (1983). Specifically, a policymaker attempts to strike a balance between minimizing the unemployment rate and minimizing deviations from the inflation target, subject to the Phillips curve. The optimal inflation policy is then a weighted average of the exogenous random inflation target and the average inflation expectation in the private sector. Departing from Barro and Gordon (1983), we allow agents to receive dispersed information and face ambiguity about the inflation target. This ambiguity may arise from factors such as ambiguous policy communication or a loose policy objective like average inflation targeting.

The optimal policy rule is an affine function of the exogenous inflation target. The slope represents the responsiveness to changes in the target, while the intercept determines the long-run average inflation rate. In the presence of imperfect information but no ambiguity, the slope is dampened due to the underreaction of the consensus forecasts, but the intercept coincides with the mean of the inflation target. Introducing ambiguity results in additional sensitivity to information and a permanent positive bias, so a steeper slope and a lifted intercept in the policy rule. The former brings the inflation dynamics closer to the first-best, while the latter pushes the policy away from it. We show that it is always beneficial to have some ambiguity about the inflation target as the initial benefits from increased sensitivity outweigh the initial costs associated with the higher bias. However, for high enough ambiguity this relationship flips, and additional ambiguity reduces welfare. It follows that an intermediate level of ambiguity is desirable.

Robust preferences Although our analysis focuses on the smooth model of ambiguity, the main insights on bias and sensitivity extend to models with robust preferences. Despite these being significantly different approaches to ambiguity, we find that under certain conditions, for a robust preferences model, a corresponding smooth model of ambiguity exists such that the equilibrium strategies in both models are identical up to a constant. There are subtle differences in determining w for the auxiliary (w, α)-modified signal process,[2] but the quantitative predictions of the two models in our applications are notably similar.

Related literature This paper contributes to the literature exploring the implications of ambiguity and ambiguity aversion in macroeconomic models. There are three prominent representations of ambiguity-averse preferences in the literature: (1) the multiple priors preference axiomatized by

[2]With robust preferences, the condition that w needs to satisfy involves conditional moments about individual variables, whereas the condition in the smooth model contains only unconditional moments of aggregate variables.

Gilboa and Schmeidler (1989); (2) the robust preferences model proposed by Hansen and Sargent (2001a,b); and (3) the smooth model of ambiguity axiomatized by Klibanoff, Marinacci, and Mukerji (2005). Each of these representations have been extensively used in macroeconomic applications. For example, in the context of business cycle models, Ilut and Schneider (2014), Bianchi, Ilut, and Schneider (2017), and Ilut and Saijo (2021) employ the multiple priors preference approach;[3] Luo and Young (2010), Bidder and Smith (2012), and Bhandari, Borovička, and Ho (2022) use robust preferences; and Backus, Ferriere, and Zin (2015) Altug, Collard, Çakmakl, Mukerji, and Özsöylev (2020), and Pei (2023) use the smooth model of ambiguity. In the asset pricing literature, Epstein and Wang (1994), Chen and Epstein (2002), Miao (2009) use the multiple priors preference approach; Hansen, Sargent, and Tallarini (1999) and Anderson, Hansen, and Sargent (2003) utilize robust preferences; and Ju and Miao (2012), Collard, Mukerji, Sheppard, and Tallon (2018), and Gallant, R Jahan-Parvar, and Liu (2018) employ the smooth model of ambiguity. Additionally, Michelacci and Paciello (2019) study the effects of monetary policy announcements under multiple priors preferences.

Most of the aforementioned works assume representative agents and abstract from incomplete information. We analyze the effects of ambiguity and ambiguity aversion within a flexible environment that accommodates not only GE considerations but also incomplete information and persistent learning. In this regard, our findings complement the literature on games with incomplete information (Morris and Shin, 2002; Woodford, 2003; Angeletos and Pavan, 2007; Angeletos and La'O, 2010), extending the analysis beyond the rational expectations benchmark. Our theoretical results also establish a link between the equilibrium outcomes in the smooth model of ambiguity and the robust preferences model.

Our paper is also related to an extensive body of literature examining systematic biases in agents' expectations using survey data. Elliott, Komunjer, and Timmermann (2008) present evidence of systematic bias in professional forecasters' expectations and suggest that an asymmetric loss function rationalizes the documented biases.[4] Similar evidence of biased expectations has been documented by Kohlhas and Robertson (2022) and Farmer, Nakamura, and Steinsson (2023) using the Survey of Professional Forecasters. Kohlhas and Robertson (2022) show that professional forecasters' expectations are biased while more accurate than commonly used time-series models, particularly in the short run. They propose a theory of cautious expectations in which agents estimate the optimal weight on observed signals using classical inference, resulting in a trade-off between bias and accuracy. Farmer, Nakamura, and Steinsson (2023) present evidence of bias in macroeconomic expectations using the Survey of Professional Forecasters (SPF) and note that the associated forecast errors are

[3] We refer to Ilut and Schneider (2022) for a more comprehensive review on the recent development of the applications of multiple priors preferences in macroeconomics.

[4] Pope and Schweitzer (2011) further demonstrate the bias originated from loss aversion can be persistent even in a high-stake context using data on professional golfers performance on the PGA TOUR.

serially correlated. Within a Bayesian paradigm that retains rationality, the authors demonstrate that the slow learning over the long-run trend with a unit root can rationalize the documented bias and persistence in forecast errors. In a similar context, Andolfatto, Hendry, and Moran (2008) argue that the bias in inflation expectations can arise due to small sample problem. In a recent study, using machine learning algorithms, Bianchi, Ludvigson, and Ma (2022) also document substantive bias in macroeconomic expectations of professional forecasters and reveal its cyclical properties.

The evidence of biased expectations extends beyond professional forecasters. Using the MSC, Bhandari, Borovička, and Ho (2022) document that households' inflation and unemployment rate forecasts feature pessimistic biases, which are counter-cyclical and co-move positively along the business cycle. Rozsypal and Schlafmann (2022) document a systematic bias in individual-level income expectations that varies with income levels.[5] They argue that an over-persistence bias rationalizes this evidence, which is broadly consistent with the approach proposed by Molavi (2022). Using survey evidence in UK, Michelacci and Paciello (2023) study how the pessimistic bias in inflation forecasts is related to households' wishes about inflation and nominal interest rates.

We contribute to this literature by presenting evidence on the joint behavior of bias and persistence of forecast errors across the income distribution, which cannot be easily rationalized by existing theories of expectation formation. Our theory provides a joint characterization of bias and sensitivity, directly addressing these empirical patterns.

Finally, our paper is related to the large literature on optimal policy under incomplete information (Adam, 2007; Lorenzoni, 2010; Paciello and Wiederholt, 2014; Amador and Weill, 2010; Angeletos and Lao, 2020; Angeletos and Sastry, 2021). This line of works has mainly focuses on the responsiveness and cyclicality of the policy instrument when agents are subject to informational frictions, while the mistakes made by the private sector are mostly temporary. Relative to the existing literature, the additional bias and sensitivity in our environment introduces a trade-off between the short-run responsiveness and the long-run bias in the policy design, and we illustrate the different welfare implications with alternative micro-foundations for the presence of biases.

2 An Illustrative Example

In this section, we discuss the interaction between ambiguity and imperfect information in a single-agent environment. We show how ambiguity aversion increases sensitivity to signals and leads to biased forecasts, relative to the Bayesian benchmark. We then extend these insights to a more general setting in the next section, where we introduce recurrent shocks and general equilibrium considerations.

[5] See Dominitz and Manski (1997), Dominitz (1998), Das and van Soest (1999), and Massenot and Pettinicchi (2019) for more related studies on the pessimistic bias of expectations on individual income.

2.1 A pure forecasting problem

Consider an inference problem about some exogenous economic fundamental. Assume the fundamental ξ is drawn from a Gaussian distribution with mean $\bar{\mu}$ and variance σ_ξ^2,

$$\xi \sim \mathcal{N}(\bar{\mu}, \sigma_\xi^2).$$

Agent i does not observe the fundamental perfectly, but receives a private noisy signal x_i about it:

$$x_i = \xi + \epsilon_i, \quad \text{with} \quad \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2).$$

In the absence of ambiguity, the distributions of underlying shocks are common knowledge. Agents perceive that there is a single model, or a single probability distribution, that describes the stochasticity of random variables. Uncertainty arises from the realization of shocks within a model, which is referred to as risk.

Bayesian expectations To begin, we revisit the standard Bayesian benchmark. In this case, agents simply want to minimize the mean-squared error (MSE) of their forecast. With Gaussian shocks, we can focus on linear strategies,[6]

$$g(x_i) = s x_i + b,$$

where the s represents the sensitivity to signals and b is some constant. Given a particular strategy characterized by the pair (s, b), the MSE can be expressed as the sum of two terms

$$\mathbb{E}\left[(g(x_i) - \xi)^2\right] = \mathcal{R}(s) + (b - (1-s)\,\bar{\mu})^2\,, \quad \text{where} \quad \mathcal{R}(s) = s^2\sigma_\epsilon^2 + (1-s)^2\,\sigma_\xi^2. \tag{2.1}$$

The first term, $\mathcal{R}(s)$, represents the cost of within-model uncertainty (risk)—a weighted sum of the variance of the fundamental and of the noise. For any sensitivity s, the second term can always be set to zero by appropriately choosing the constant b. It follows that the optimal sensitivity, s^{RE}, is obtained by minimizing the cost of risk:

$$s^{\text{RE}} = \operatorname{argmin} \mathcal{R}(s) = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\epsilon^2}.$$

The optimal strategy is then given by

$$g(x_i) = s^{\text{RE}} x_i + (1 - s^{\text{RE}})\mu, \tag{2.2}$$

[6]With Gaussian shocks, the optimal conditional expectation is linear in signals.

a simple weighted average between the prior mean and the signal, following Bayesian updating. The optimal sensitivity corresponds to the familiar signal-to-noise ratio.

2.2 Ambiguity and Ambiguity Aversion

Ambiguity When agents face ambiguity, they could perceive multiple plausible models that describe the economy, each corresponding to a different distribution of underlying shocks. This generalization accommodates the possibility that agents may have doubts about what is the right model of the economy. For instance, in the aftermath of the Pandemic, consumers may be uncertain about whether inflation will fluctuate around an average level of 2% or 4% in the upcoming years. Similarly, during the slow recovery from the Great recession, firms may wonder whether their sales growth will remain stagnant or rebound to its pre-recession level.

We restrict our attention to the case where agents face ambiguity about the prior mean of the aggregate fundamental ξ.[7] Objectively, ξ is distributed according to $\mathcal{N}(\bar{\mu}, \sigma_\xi^2)$. Subjectively, however, agents' priors do not necessarily coincide with the objective distribution. They believe there can be multiple possible prior means, which themselves follow a normal distribution centered around the objective mean:

$$\xi \sim \mathcal{N}(\mu, \sigma_\xi^2), \quad \text{where} \quad \mu \sim \mathcal{N}(\bar{\mu}, \sigma_\mu^2).$$

In this case, different models are indexed by different μ, while σ_μ^2 parameterizes the ex-ante uncertainty about the models. When $\sigma_\mu^2 = 0$, we return to the Bayesian benchmark. Without loss of generality, we assume that $\bar{\mu} = 0$. The ex-ante probability density function of the perceived distribution of μ satisfies

$$p(\mu) \propto \exp\left(-\frac{1}{2}\sigma_\mu^{-2}\mu^2\right).$$

Ambiguity Aversion We now specify the preference of agents towards ambiguity. Specifically, we are interested in the case where agents are ambiguity averse, meaning they dislike ambiguity more than risk. To this end, given a strategy $g(x_i)$, we adopt the following loss function $\mathcal{L}(g)$, à la Klibanoff, Marinacci, and Mukerji (2005)

$$\mathcal{L}(g) = \phi^{-1}\left(\int_\mu \phi\bigg(\mathbb{E}^\mu\left[(g(x_i)-\xi)^2 - \chi\xi\right]\bigg)p(\mu)\mathrm{d}\mu\right). \tag{2.3}$$

The integral over μ reflects the fact that agents face additional uncertainty about the prior mean. For each μ, $\mathbb{E}^\mu[\cdot]$ denotes the mathematical expectations in an economy where the prior mean is

[7] We also explore the case where the ambiguity is about the variance of the fundamental in Appendix E. There, the effect of ambiguity shows up in a similar but more restrictive way: the equilibrium strategy features higher sensitivity but zero bias.

given by μ.

There are two additional modifications relative to the previous forecasting problem. First, the transformation $\phi(\cdot)$ introduces an additional cost associated with ambiguity about μ. When $\phi(\cdot)$ is linear, the problem reduces to

$$\mathcal{L}(g) = \int_{\mu} \mathbb{E}^{\mu}\left[(g(x_i) - \xi)^2 - \chi\xi\right] p(\mu)\mathrm{d}\mu. \tag{2.4}$$

This corresponds to the *ambiguity neutral* case, where the uncertainty about μ is treated in the same way as the within-model uncertainty about ξ and ϵ_i. Namely, adding uncertainty about μ is equivalent to drawing a compound lottery.

The distinction between ambiguity and risk becomes meaningful when the transformation $\phi(\cdot)$ is convex. Then, agents incur additional losses when their perceived model is found to be incorrect, and agents are said to have ambiguity aversion. In what follows, we assume that $\phi(\cdot)$ takes a constant absolute ambiguity aversion (CAAA) form, that is,

$$\phi\left(x\right) = \frac{1}{\lambda}\exp\left(\lambda x\right), \tag{2.5}$$

where $\lambda \geq 0$ measures the degree of ambiguity aversion.

Second, the payoff directly depends on the level of the fundamental, captured by the term $\chi\xi$. For example, firms benefit from higher TFP regardless of their production decisions, while consumers benefit from higher real income independent of their consumption-saving decisions. The direct dependence on the exogenous fundamental is common in economic problems but can usually be ignored as it is inconsequential when agents are ambiguity neutral or when there is no ambiguity. In contrast, this dependence leads to biased forecast when agents are ambiguity averse.

Sensitivity and bias To see how ambiguity aversion modifies agents' strategies, it is useful to decompose the loss function into the costs due to risk and ambiguity. Given a strategy $g(x_i) = sx_i + b$, this decomposition can be expressed as

$$\mathcal{L}(g) = \underbrace{\mathcal{R}(s)}_{\text{cost of risk}} + \underbrace{\frac{1}{\lambda}\log\int_{\mu}\exp\left(\lambda\left[\left(b-\left(1-s\right)\mu\right)^2 - \chi\mu\right]\right)p(\mu)\mathrm{d}\mu}_{\text{cost of ambiguity}}. \tag{2.6}$$

First, ambiguity aversion leads to a higher sensitivity towards signals. Intuitively, a more diffused prior leads agents to rely more heavily on signals, and ambiguity aversion amplifies this effect. To see the forces more clearly, note that the cost of risk, $\mathcal{R}(s)$, remains the same as in equation (2.1). When there is no ambiguity, an agent can simply choose $s = s^{\mathrm{RE}}$ to minimize $\mathcal{R}(s)$. When uncertainty about the prior mean is present, a trade-off arises between the cost of risk and the cost of ambiguity. At one extreme, if an agent only wants to minimize the cost of ambiguity, they would set $s = 1$ to eliminate the impact of ambiguity on forecast errors. Striking a balance between the two types of cost implies an enhanced sensitivity towards signals.

Second, ambiguity aversion introduces a bias in agents' strategies and forecasts. In equation (2.6), the direct effect of the exogenous fundamental, $\chi\mu$, is not symmetric around zero. Consider the case where $\chi > 0$. The losses in a model indexed with a negative prior mean are higher than the gains in a model indexed with a positive prior mean. Similar to the idea of self-insurance, an agent finds it optimal to mitigate losses in "rainy days" by assigning more weight to models with a negative μ. This affects the choice of the constant b, leading to biased forecasts. Ex ante, this incentive makes it appear as if agents have a more pessimistic view of the world.

The adjustment of sensitivity and bias in an agent's strategy is driven by their payoffs, as they place more weight on models that generate higher expected costs due to ambiguity aversion. Simultaneously, the magnitude of these costs is determined by the chosen strategy. This interdependence leads to a fixed-point problem. According to the following proposition, the optimal strategy can be understood as if agents faced no ambiguity but had a modified prior belief.

Proposition 2.1. *The optimal linear strategy is equivalent to the Bayesian expectation with a more diffused prior belief,* $\xi \sim \mathcal{N}(0, \tilde{\sigma}_\mu^2(s^*) + \sigma_\xi^2)$*, and a bias*

$$g(x_i) = s^* x_i + \mathcal{B},$$

where the sensitivity s^* *satisfies*

$$s^* = \frac{\tilde{\sigma}_\mu^2(s^*) + \sigma_\xi^2}{\tilde{\sigma}_\mu^2(s^*) + \sigma_\xi^2 + \sigma_\epsilon^2}, \qquad \text{with} \qquad \tilde{\sigma}_\mu^2(s) \equiv \frac{\sigma_\mu^2}{1 - 2\lambda\sigma_\mu^2(1-s)^2}, \tag{2.7}$$

and the bias is given by

$$\mathcal{B} = -\chi\lambda\sigma_\mu^2(1 - s^*). \tag{2.8}$$

Let us unpack these expressions. First of all, given the additional variance, $\tilde{\sigma}_\mu^2(s^*)$, in the prior belief, the optimal sensitivity is identical to the one under Bayesian expectations (2.2). The determination of the optimal sensitivity, s^*, requires solving a fixed-point problem. In the presence of ambiguity aversion, agents increase their sensitivity to reduce the ambiguity cost. At the same time, however, the degree to which they want to penalize the ambiguity cost is endogenously determined by the employed sensitivity. The system (2.7) succinctly summarizes these forces.

Next, even though the ex-ante belief about the distribution of μ is centered around zero, agents behave as if the prior mean is biased, captured by $\mathcal{B}$. The "as-if" shift of the mean is increasing in how much agents directly care about the fundamental, the degree of ambiguity aversion, and the amount of ambiguity.

The presence of ambiguity aversion predicts a joint pattern of sensitivity and bias. In the sequel, we show that this prediction is robust to different information structures and preferences and is supported by survey evidence on expectations.

3 General Equivalence Result

In this section, we present our main theoretical results. We extend the intuition from the illustrating example to an environment with more general preferences that also accommodate strategic interactions between agents. In addition, we allow for flexible persistent information structures and learning dynamics. We show that the optimal action under ambiguity aversion is akin to a pure forecasting problem of the fundamental with a modified prior belief.

3.1 Environment

Objective environment The economy is populated by a continuum of agents indexed by i. Agents care about a common fundamental, ξ_t, which follows the stochastic process:

$$\xi_t = a(L)\eta_t, \quad \text{with} \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2),$$

where $a(L)$ is a polynomial in the lag operator L, and η_t is the innovation to the aggregate fundamental.[8]

Agents receive dispersed information about the fundamental. The vector of signals observed by individual agent i every period is given by

$$x_{it} = m(L)\eta_t + n(L)\epsilon_{it}, \quad \text{with} \quad \epsilon_{it} \sim \mathcal{N}(0, \Sigma), \tag{3.1}$$

where $m(L)$ and $n(L)$ are polynomial matrices in L that determine the dynamics of the signal process, and ϵ_{it} is a vector of idiosyncratic noises that wash out in aggregate. Unlike in Section 2, the fundamental as well as the signals can be persistent, which implies that the entire history of signals will be relevant for inference. To ensure that the random variables are stationary and signals do not contain future information, we make the following standard assumption.

Assumption 1. *All elements of $a(L)$, $m(L)$, and $n(L)$ contain only L with non-negative powers and*

[8] So far, η_t is the only aggregate shock in the economy. We extend the analysis to allow for multiple aggregate shocks in Appendix B.2.

are square-summable.

Ambiguity In the objective environment, η_t is normally distributed with mean zero. Subjectively, agents believe that η_t is drawn from a Gaussian distribution with the same volatility, σ_η^2, but there is uncertainty about its prior mean, denoted by μ_t. The ambiguity about ξ_t is captured by the perception that

$$\eta_t \sim \mathcal{N}(\mu_t, \sigma_\eta^2), \quad \text{and} \quad \mu_t \sim \mathcal{N}(0, \sigma_\mu^2). \tag{3.2}$$

As in Section 2, the value of σ_μ^2 determines the degree of ambiguity.

Preference towards risk We first specify preferences about within-model risk. As a baseline, we consider a class of economies with quadratic utility given by

$$u(k_{it}, K_t, \xi_t) = -\frac{1}{2}\left[(1-\alpha)\left(k_{it}-\xi_t\right)^2 + \alpha\left(k_{it}-K_t\right)^2\right] - \chi\xi_t - \frac{1}{2}\gamma\xi_t^2, \tag{3.3}$$

where k_{it} denotes agent i's action and K_t denotes the aggregate outcome of the economy,

$$K_t \equiv \int k_{it} \mathrm{d}i.$$

The first component of the utility function, $(1-\alpha)(k_{it}-\xi_t)^2 + \alpha(k_{it}-K_t)^2$, captures the payoff directly associated with the individual agent's action. Agents aim to align their actions with the exogenous fundamental and the aggregate outcome. The degree of strategic complementarity, controlled by the parameter α, influences the strength of general equilibrium considerations. The second component, $\chi\xi_t + \frac{1}{2}\gamma\xi_t^2$, captures the non-strategic impact of the fundamental on the agent's utility. This component only affects the agent's optimal strategy when there is ambiguity and agents are ambiguity averse.

The utility specification in equation (3.3) can be considered a quadratic approximation of a generic utility function, $u(k_{it}, K_t, \xi_t)$, as specified in Angeletos and Pavan (2007).[9] This specification accommodates strategic interactions among agents and a flexible dependence on fundamentals, but excludes the dependence on $(K_t-\xi_t)^2$, K_t, and $K_t\xi_t$. In the language of Angeletos and Pavan (2007), specification (3.3) pertains to economies that are efficient under both complete and incomplete information. However, in general, the underlying economy may be inefficient, and in such cases, dependence on the aforementioned terms may arise. Nevertheless, our main observational equivalence result still applies in these cases, as discussed in Appendix B.1.

[9]The approximation is centered around the non-stochastic steady state: either the deterministic or the ambiguous steady state depending on the model environment.

Preference towards ambiguity Agents are assumed to be ambiguity averse, with preferences represented by the smooth model of ambiguity,

$$\phi^{-1}\left(\int_{\mu^t}\phi\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it},K_t,\xi_t\right)\right]\right)p\left(\mu^t\right)\mathrm{d}\mu^t\right). \tag{3.4}$$

The expectation operator $\mathbb{E}^{\mu^t}\left[u\left(k_{it},k_t,\xi_t\right)\right]$ denotes the ex-ante expected utility under the model indexed by $\mu^t\equiv\{\mu_t,\mu_{t-1},\ldots\}$, and $p\left(\mu^t\right)$ denotes the prior belief about μ^t derived from (3.2). We continue to assume the functional form:

$$\phi(x)=-\frac{1}{\lambda}\exp(-\lambda x),$$

which permits the tractability of the inference problem.

Remark on ex-ante strategy Similar to the illustrative example from Section 2, each agent i chooses their preferred strategy, which is now a contingency plan denoted by $k_{it}=g(x_i^t)$, for every possible history of private signals, $x_i^t\equiv\{x_{it},x_{it-1},\ldots\}$. We assume that all agents can commit to following their strategies, determined ex ante, when taking actions ex post. The assumption of full commitment is not restrictive on its own. It is equivalent to assuming that the conditional preferences of agents upon receiving a history of private signals, x_i^t, are dynamic consistent.[10] In the smooth model of ambiguity, it is the smooth rule of updating, proposed by Hanany and Klibanoff (2009), that ensures dynamic consistency. Moreover, the full-commitment allocation coincides with that of the ex-ante equilibrium defined in Hanany, Klibanoff, and Mukerji (2020), which is sequentially optimal when conditional preferences are updated using the smooth rule of updating.[11]

Finally, we make the following assumption to ensure that the problem is well-defined.[12]

Assumption 2. $\gamma\geq 0$ *and* $\lambda\gamma\frac{\sigma_\mu^2}{\sigma_\eta^2}\mathbb{V}(\xi_t)<1$.

3.2 Subjective Beliefs and Equilibrium

We begin by defining a Nash equilibrium in our environment.

[10]We restrict attention to the set of preferences that preserve closure, namely each member of that set remains in the set after updating. This property holds naturally for expected utility preferences with Bayesian updating.

[11]In fact, if the smooth rule of updating is employed, the equilibrium under full commitment corresponds to the sequential equilibrium with ambiguity, the equilibrium refinement proposed by Hanany, Klibanoff, and Mukerji (2020) that mimics the notion of a sequential equilibrium under expected utility.

[12]The first restriction, $\gamma\geq 0$, ensures that in the non-stochastic steady state under complete information, the utility function is concave in the fundamental. The second assumption stipulates that the level of ambiguity or the degree of ambiguity aversion should not exceed a certain threshold. It ensures that the ex-ante objective (3.4) is finite for at least one strategy, so that the agent's choice set is non-empty (see the Proof of Lemma A.5 in Appendix A for a more detailed discussion).

Definition 3.1. *A linear Nash equilibrium is a strategy $g(x_i^t)$ such that $k_{it} = g(x_i^t)$ maximizes the objective (3.4), and the aggregate outcome is consistent with individual actions, $K_t = \int g(x_i^t)\mathrm{d}i$.*

Through out, we focus on linear strategies where $g(x_i^t)$ is a linear function of the history of signals.[13] We can show that such an equilibrium always exists.

Proposition 3.1. *The linear Nash equilibrium exists.*

Without ambiguity aversion, the problem reduces to a standard beauty contest, and the optimal strategy can be written as an average of the expected fundamental and the expected aggregate outcome, as in Morris and Shin (2002). With ambiguity aversion, a similar result holds, albeit the expectations need to be based on the endogenous subjective beliefs. Consider the first-order condition for maximizing (3.4) with respect to the individual action k_{it}:

$$\int_{\mu^t} \mathbb{E}^{\mu^t}\left[\frac{\partial u(k_{it}, K_t, \xi_t)}{\partial k_{it}} \mid x_i^t\right] \phi'\left(\mathbb{E}^{\mu^t}[u\left(k_{it}, K_t, \xi_t\right)]\right) p\left(\mu^t \mid x_i^t\right) \mathrm{d}\mu^t = 0. \tag{3.5}$$

Notice that when evaluating the payoff implications of an action, relative to the Bayesian kernel $p(\mu^t|x_i^t)$, agents behave as if they were using an expectation kernel distorted by $\phi'(\mathbb{E}^{\mu^t}[u\left(k_{it}, K_t, \xi_t\right)])$, which we refer to as the agents' subjective beliefs. This distortion reflects the fact that, whenever a model generates lower ex-ante expected utility in equilibrium, agents would regard it as the more likely model in their posterior belief relative to the Bayesian posterior. Recall that we have already employed this line of argument when explaining the bias term in the illustrating example. In the special case in which agents are ambiguity neutral, $\phi'(\cdot) = 1$, and the subjective kernel coincides with the Bayesian one. The following proposition summarizes this discussion.

Proposition 3.2. *Taking the law of motion of K_t as given, individual i's best response satisfies*

$$k_{it} = (1-\alpha)\,\mathcal{F}_{it}\left[\xi_t\right] + \alpha\,\mathcal{F}_{it}\left[K_t\right], \tag{3.6}$$

where $\mathcal{F}_{it}\left[\cdot\right]$ denotes agent i's subjective expectation operator, that is

$$\mathcal{F}_{it}\left[\cdot\right] \equiv \int_{\mu^t} \mathbb{E}^{\mu^t}[\,\cdot \mid x_i^t]\,\hat{p}\left(\mu^t \mid x_i^t\right) \mathrm{d}\mu^t, \quad \textit{with} \quad \hat{p}\left(\mu^t \mid x_i^t\right) \propto \phi'\left(\mathbb{E}^{\mu^t}[u\left(k_{it}, K_t, \xi_t\right)]\right) p\left(\mu^t \mid x_i^t\right).$$

Importantly, since agents' payoffs depend on the aggregate outcome K_t, both the magnitude and dynamics of belief distortions hinge on the equilibrium coordination motive. Consequently, besides

[13] The focus on linear strategies should be viewed as a refinement rather than a restriction of the equilibrium, since the linear-strategy equilibrium satisfies the optimality condition (3.5). We cannot rule out, however, the possibility that all agents collectively choose a non-linear strategy.

needing to anticipate the actions and beliefs of others, the way that agents form their beliefs is also affected by the aggregate outcome. The fact that the equilibrium outcome and the distorted subjective beliefs have to be jointly determined makes solving the equilibrium significantly more involved.

The equilibrium outcome can also be expressed as a weighted sum of higher-order subjective expectations.

Corollary 3.1. *In equilibrium, the aggregate outcome is a function of a weighted sum of infinite subjective higher-order expectations of* ξ_t*,*

$$K_t = (1-\alpha)\sum_{j=0}^{\infty} \alpha^j \overline{\mathcal{F}}_t^{j+1}\left[\xi_t\right], \tag{3.7}$$

where $\overline{\mathcal{F}}_t^1[\cdot] \equiv \int \mathcal{F}_{it}[\cdot]\mathrm{d}i$*, and* $\overline{\mathcal{F}}_t^{j+1}[\cdot] \equiv \int \mathcal{F}_{it}\left[\overline{\mathcal{F}}_t^j[\cdot]\right]\mathrm{d}i$*.*

This result indicates that when forming their own beliefs, agents must take into account the possible bias and altered sensitivity in others' forecasts, and this incentive is regulated by the degree of strategic complementarity. In Section 3.4, we show that this representation helps uncover the interaction between coordination and ambiguity aversion.

3.3 Equilibrium Characterization

In this subsection, we provide an equivalence result that circumvents the determination of higher-order subjective beliefs. We show that solving for the equilibrium strategy described above can be reduced to solving a single-agent Bayesian forecasting problem with a modified information structure. This forecasting problem can then be tackled using the Kalman filter, which facilitates the development of a convenient toolbox for solving models featuring ambiguity and persistent learning. The straightforward characterization it provides also allows us to derive valuable comparative statics results under general conditions.

Auxiliary forecasting problem First consider the following auxiliary inference problem about the fundamental with Bayesian forecasters, which we later link back to the economy with ambiguity.

Definition 3.2. *The* (w,α)*-modified signal process is given by*

$$\tilde{\xi}_t = a(L)\tilde{\eta}_t, \qquad \textit{with} \quad \tilde{\eta}_t \sim \mathcal{N}(0,(1+w)\sigma_\eta^2), \tag{3.8}$$

$$\tilde{x}_{it} = m(L)\tilde{\eta}_t + n(L)\tilde{\epsilon}_{it}, \qquad \textit{with} \quad \tilde{\epsilon}_{it} \sim \mathcal{N}(0,(1-\alpha)^{-1}\Sigma), \tag{3.9}$$

where w *is a non-negative scalar and* α *is the degree of complementarity. Let the optimal Bayesian* *forecast be given by*

$$\tilde{\mathbb{E}}_{it}[\tilde{\xi}_t] = p(L; w, \alpha)\, \tilde{x}_{it}. \tag{3.10}$$

This signal process imposes two modifications relative to the original process. The volatility of the innovation to the fundamental is amplified by a factor of $(1 + w)$, and the covariance matrix of the idiosyncratic noise is amplified by $(1 - \alpha)^{-1}$. Intuitively, ambiguity aversion leads to a more diffuse prior about η_t, while the coordination motive reduces agents' incentives to rely on private information. These two considerations from the original environment are captured by the modifications to the shock processes in this auxiliary forecasting problem.

Sensitivity and bias Next, we introduce a notion of aggregate sensitivity and bias in our multivariate setting, which helps characterize the equilibrium strategy.

Definition 3.3. *Define the aggregate sensitivity to signals as*

$$\mathcal{S} \equiv 1 - \frac{\mathbb{COV}\,(\xi_t - K_t, \xi_t)}{\mathbb{V}\,(\xi_t)}, \tag{3.11}$$

and the bias as

$$\mathcal{B} \equiv \mathbb{E}[\xi_t] - \mathbb{E}[K_t]. \tag{3.12}$$

With multiple signals, the sensitivity to signals can no longer be measured by the loading on a particular signal. To better understand the definition in equation (3.11), notice that under perfect information, the aggregate outcome K_t simply mirrors the exogenous fundamental ξ_t, and $\mathcal{S} = 1$. When information is incomplete, however, agents receive noisy signals and their aggregate response is dampened, reducing $\mathcal{S}$. Moving to the definition of bias in equation (3.12), notice that without ambiguity, the unconditional mean of the aggregate outcome coincides with that of the fundamental, resulting in $\mathcal{B} = 0$, even if information is incomplete. Ambiguity aversion, however, can lead to a permanent bias in agents' actions.

These definitions for sensitivity and bias can also be viewed as the counterparts of the regression coefficients in the following equation:

$$\xi_t - K_t = \beta_0 + \beta_1 \xi_t + \text{residuals}, \tag{3.13}$$

where β_0 corresponds to the bias, $\mathcal{B}$, and β_1 corresponds to the reverse of the sensitivity, $1-\mathcal{S}$.[14] These two moments contain rich information about agents' subjective beliefs and about the expectation

[14] When K_t stands for the consensus forecasts about ξ_t, regression (3.13) resembles the ones explored in Kohlhas and Walther (2019).

formation process. As in Section 2, the levels of bias and sensitivity are jointly determined in equilibrium.

Equilibrium strategy Despite the complex interactions between persistent information, coordination motives, and ambiguity aversion, the equilibrium strategy ultimately takes a relatively simple form. This form can be connected to the notion of aggregate sensitivity and bias. Let $\tau_\mu \equiv \sigma_\mu^2/\sigma_\eta^2$ be a normalized measure of the amount of ambiguity.

Proposition 3.3. *The linear strategy in equilibrium takes the following form*

$$g(x_i^t) = (1+r)p(L; w, \alpha)x_{it} + \mathcal{B}. \tag{3.14}$$

1. *The polynomial matrix $p(L; w, \alpha)$ is the Bayesian forecasting rule in (3.10) with the (w, α)-modified signal process and w satisfies*

$$w = \left[\frac{1}{\tau_\mu} - \lambda\left(1-\alpha\right)\left(\mathbb{V}\left(\xi_t - K_t\right) + \frac{\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)^2\left(1-\mathcal{S}\right)^2}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\right)\right]^{-1} \geq \tau_\mu. \tag{3.15}$$

2. *The additional amplification, r, satisfies*

$$r = \gamma\frac{\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\frac{w}{1+w}(1-\mathcal{S}) \geq 0. \tag{3.16}$$

3. *The level of bias, $\mathcal{B}$, satisfies*

$$\mathcal{B} = \chi\frac{\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}(1-\mathcal{S}). \tag{3.17}$$

To understand this result, let us first ignore ambiguity ($\tau_\mu = 0$). In this case, the equilibrium strategy reduces to $p(L; 0, \alpha)$. This means agents' actions are equivalent to forming Bayesian expectations about ξ_t, with the volatility of idiosyncratic noise adjusted to accommodate coordination considerations. This equilibrium characterization bypasses the dependence on higher-order beliefs and includes the single-agent results from Huo and Pedroni (2020).

When ambiguity is present ($\tau_\mu > 0$), prior uncertainty looms larger. This effect is captured by the amplified variance of the fundamental, by a factor of $(1+w)$, in equation (3.8). Following the new forecasting rule $p(L; w, \alpha)$, agents' reliance on their signals increases when making forecasts. This channel makes it appear as if agents are overconfident in their signals, as in Broer and Kohlhas (2019). Moreover, remember that the term $-\chi\xi_t - \frac{1}{2}\gamma\xi_t^2$ in the utility function captures the non-strategic impact of the fundamental on the agent's payoff. A positive γ introduces an additional reason for agents to react to signals: the dependence on ξ_t^2 intensifies the impact of extreme realizations of the prior mean μ_t on welfare. This pushes agents' actions further away from the Bayesian benchmark through its effect on r. This additional overreaction is similar to the diagnostic expectations explored in Bordalo, Gennaioli, Ma, and Shleifer (2020), but in our environment the magnitude of this overreaction is endogenously determined in equilibrium.

Additionally, if $\chi \neq 0$, the equilibrium strategy exhibits a permanent bias. Keep in mind that agents' subjective beliefs, $\mathcal{F}_{it}[\cdot]$, assign more weight to models that generate lower ex-ante expected utility, leading to a pessimistic bias.[15] Condition (3.17) indicates that the magnitude of the bias $\mathcal{B}$ and the level of sensitivity $\mathcal{S}$ are endogenously connected and constrained by equilibrium conditions. It is worth noting that condition (3.17) does not imply a direct inverse relationship between sensitivity and bias, as both are ultimately functions of deep parameters in the model. For instance, in Section 2, we observe that when the degree of ambiguity aversion varies, both bias and sensitivity move in the same direction. The joint pattern of bias and sensitivity yields testable implications, which we explore in Section 4.

Computation Even without ambiguity, solving for the equilibrium presents a significant challenge due to the combination of persistent information and coordination motives. Agents' strategies are infinite-dimensional objects with the entire history of signals as state variables. Although a finite-state representation is possible when signals follow finite ARMA(p, q) processes (Woodford, 2003; Angeletos and La'O, 2010; Huo and Pedroni, 2020), it remains unclear whether these results could be extended to models with ambiguity aversion. The following corollary confirms they can:

Corollary 3.2. *When the fundamental and signals follow finite ARMA processes, the equilibrium strategy admits a finite-state representation.*

Proposition 3.3 shows that the finite-state representation of the equilibrium can be achieved with ambiguity-averse agents, persistent information, and general-equilibrium considerations. It also outlines a clear computation method:[16]

1. For a particular pair of (w, α), the Bayesian forecasting rule $p(L; w, \alpha)$ can be obtained using standard algorithms such as the Kalman filter.

2. The value of r can be determined using condition (3.16). Together with the equilibrium strategy (3.14), this leads to the outcome K_t and the sensitivity $\mathcal{S}$. At this stage, the bias can be ignored as it does not affect any of the terms in the formulas.

[15] Notice that the sign of the bias depends on the sign of χ.

[16] We could also accommodate the case where the signals contain endogenous aggregate variables. Since each agent behaves in a competitive way, they still treat the information as exogenous even if signals are endogenous. Therefore, adding endogenous signals amount to adding another layer of fixed point problem on top of our equivalence result.

3. Condition (3.15) can then be used to iterate on the value of w until convergence.

4. Finally, the bias can be obtained from equation (3.17).

In summary, this complex, seemingly infinite-dimensional problem effectively reduces to a one-dimensional fixed-point problem about w. In Section 4, we leverage this result to characterize the dynamics of inflation forecasts under ambiguity, which boils down to solving a single cubic equation.

3.4 Role of General Equilibrium Considerations

How do GE considerations interact with ambiguity aversion? To answer this question, it is useful to revisit the higher-order expectation representation in equation (3.7). A change in α not only affects the responses of expectations at each order, but also shifts the relative weight assigned to each expectation. However, these intensive and extensive margins do not necessarily move in the same direction in shaping equilibrium outcomes.

To illustrate the intuition with, we start with a simple example with static information, and then extend the findings to general information processes. Suppose that the utility function is given by

$$u\left(k_i, K, \xi\right) = -\frac{1}{2}\left[\left(1-\alpha\right)\left(k_i-\xi\right)^2 + \alpha\left(k_i-K\right)^2\right] - \chi\xi,$$

and that agents perceive ambiguity about the fundamental according to

$$\xi \sim \mathcal{N}\left(\mu, \sigma_\xi^2\right), \quad \text{and} \quad \mu \sim \mathcal{N}\left(0, \sigma_\mu^2\right),$$

where μ is objectively zero. Agents also observe a private signal $x_i \sim \mathcal{N}\left(\xi, \sigma_\epsilon^2\right)$. In this simple setup, we can further evaluate the effects of GE considerations on sensitivity and bias, by examining the properties of subjective higher-order expectations.

Proposition 3.4. *In the example economy, subjective higher-order expectations obey the following structure*

1. *The m-th order subjective expectation defined in condition (3.7) is given by*

$$\overline{\mathcal{F}}^m[\xi] = \kappa_m\xi + \beta_m, \quad \text{with} \quad \kappa_m = \left(\frac{(1+w)\sigma_\xi^2}{(1+w)\sigma_\xi^2 + \sigma_\epsilon^2}\right)^m, \quad \text{and} \quad \beta_m = \beta_{m-1} + (\kappa_m - \kappa_{m-1})\lambda\chi\sigma_\mu^2;$$

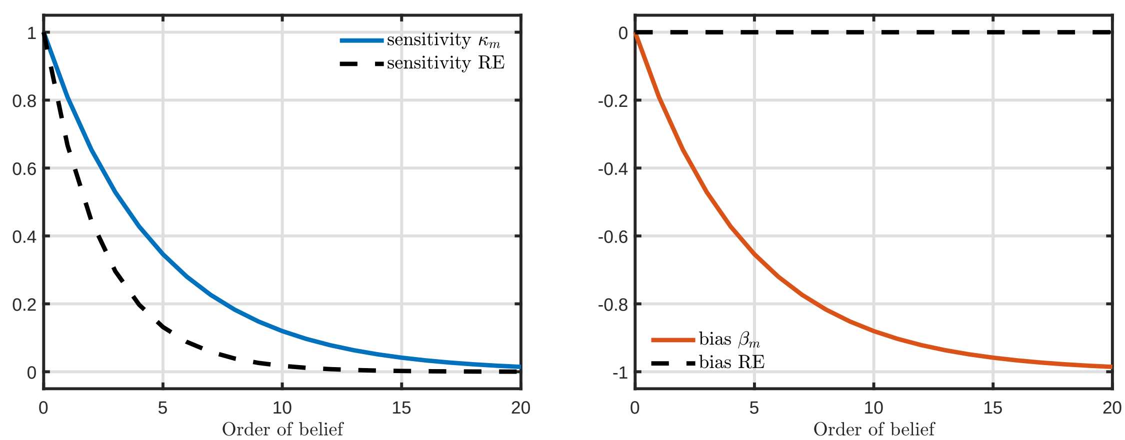
2. *The endogenous multiplier w is not monotonic in α;*

3. The aggregate outcome is given by

$$K = \mathcal{S}\,\xi + \mathcal{B} = (1-\alpha)\sum_{m=0}^{\infty}\alpha^m \kappa_m\,\xi + (1-\alpha)\sum_{m=0}^{\infty}\alpha^m\beta_m.$$

Part 1 implies that for a fixed α, as the order increases, the sensitivity of the subjective higher-order expectations, κ_m, decreases with the order. This is displayed in the left panel of Figure 3.1 and resembles the rational expectations result without ambiguity (Morris and Shin, 2002). Interestingly, the right panel of Figure 3.1 shows that the bias, β_m, is increasing in the order of expectation. This is a result of the accumulation of pessimism in ascending the hierarchy of beliefs—a footprint of ambiguity aversion. When forecasting the beliefs of others, agents internalize the bias of others in addition to their own.

FIGURE 3.1: Subjective Higher-Order Beliefs



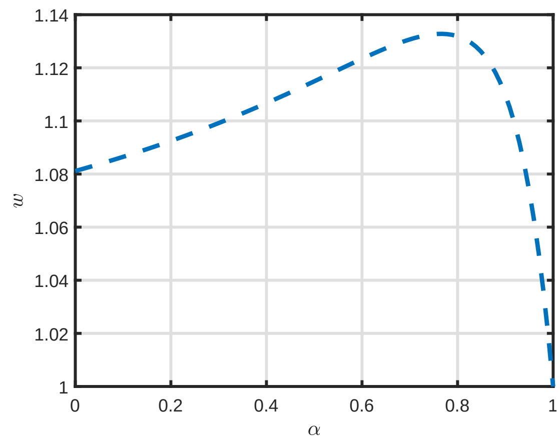
(A) Sensitivity of HOB (B) Bias of HOB

Note: This figure reports sensitivity (Panel A) and bias (Panel B) associated with the entire belief hierarchy as a function of the order of beliefs.

Part 2 suggests that there are competing forces shaping the effect of a change in α on w. Ceteris paribus, a higher α shifts agents' attention from forecasting the fundamental to forecasting others' actions. As a result, private information about the fundamental becomes less relevant, and agents reduce how much they respond to signals. As can be seen in condition (3.15), the endogenous amplification force captured by w vanishes when α approaches 1.

However, this is not the only force at play. Condition (3.15) also reveals that a reduction in the sensitivity $\mathcal{S}$ could contribute to a higher w. This is due to the fact that, when agents reduce the responsiveness to signals, they also increase their reliance on their ambiguous prior, and this is costly since agents are ambiguity averse. Endogenously, this puts upwards pressure on w to undo this effect. Overall, these two forces leave the comparative statics of w with respect to α ambiguous. As shown in Figure 3.2, w exhibits a humped shape with respect to the strength of GE considerations.

FIGURE 3.2: Comparative Statics of w over α.



This nonmonotonicity is inherited by both κ_m and β_m. Together with Part 1 of Proposition 3.4, this implies that sensitivities and biases at every order mirror the pattern of w and are not monotonic with respect to the GE consideration α.

Part 3 states that the observed sensitivity and bias of the aggregate outcome is a weighted average of κ_m and β_m. Mechanically, as α increases, a larger weight is assigned to higher-order expectations relative to first-order expectations.

Ultimately, when α changes, the reweighting channel dominates the ambiguous effects on w, leading to overall lower sensitivity and increased bias. These observations hold beyond this simple example. Taking advantage of Proposition 3.3, we obtain the following result for a general information structure.

Proposition 3.5. *When $\gamma = 0$, the sensitivity $\mathcal{S}$ is decreasing in α, and the magnitude of bias $|\mathcal{B}|$ is increasing in α.*

3.5 Extension

The results discussed so far depend on assumptions about the payoff function and the information structure. However, the observational equivalence result and the insights regarding bias and sensitivity are more general. In Appendix B.1, we allow the economy to be inefficient under both complete and incomplete information. Essentially, our equivalence result applies to general quadratic payoff structures. Our optimal-policy application in Section 5 utilizes this set of results.

In another extension, Appendix B.2 shows how to generalize these insights to settings with multiple aggregate shocks, including common noises. This extension can be used to study the interaction between ambiguity and the effects of non-fundamental driven fluctuations, for example.

4 Application: Inflation Forecasts

Under full information rational expectations (FIRE), optimal forecasts are unbiased and forecast errors are serially uncorrelated. In contrast, using survey data on expectations, the literature has demonstrated that the average forecasts of agents are biased and that forecast errors are serially correlated.[17] In this section, we document additional survey evidence on the joint behavior of bias and persistence across the income distribution, which cannot be easily rationalized by existing theories of expectation formation. We then show that these facts emerge naturally from a micro-founded model with ambiguity-averse consumers.

4.1 Data and Facts

The Michigan Survey of Consumers (MSC) collects data on household inflation expectations asking what is their "price expectations for the next 12 months." It also provides information about household income, which allows us to allocate surveyed households in each quarter into groups based on their income percentiles. For each income group g, the average inflation forecast error in quarter t is calculated as the average inflation forecast error across every household $i \in \mathbb{I}_g$:

$$\overline{\mathrm{FE}}_{g,t} \equiv \int_{\mathbb{I}_g} \left(\pi_{t,f} - \mathcal{F}_{i,g,t}[\pi_t]\right) \mathrm{d}i,$$

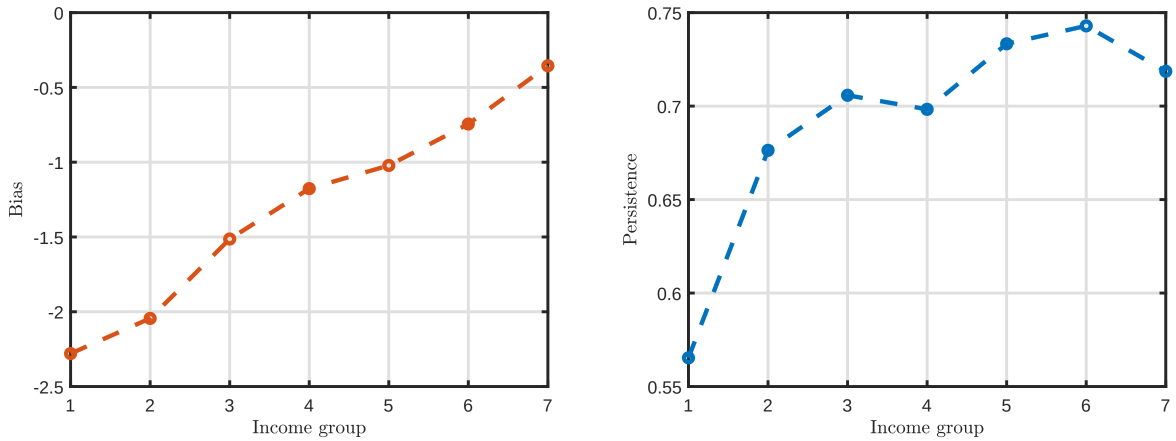
where we continue to use $\mathcal{F}_{i,g,t}$ to denote households' subjective expectations. The bias and persistence of forecast errors for each income group are given by their across-time average and autocorrelation:

$$\mathrm{Bias}_g \equiv \frac{1}{T}\sum_{t=1}^{T} \overline{\mathrm{FE}}_{g,t}, \qquad \text{and} \qquad \mathrm{Persistence}_g \equiv \mathrm{Corr}\left(\overline{\mathrm{FE}}_{g,t}, \overline{\mathrm{FE}}_{g,t-1}\right). \tag{4.1}$$

Figure 4.1 presents the bias (panel A) and persistence (panel B) of households' one-year-ahead inflation forecasts. In line with the existing literature, households' inflation forecasts are biased upwards leading to negative average forecast errors. What is interesting is the joint behavior of bias and persistence in the cross-section of households' income. As households move up the income ladder, bias in inflation expectation decreases, whereas the persistence of inflation forecast errors increases.

[17] See, for example, Farmer, Nakamura, and Steinsson (2023) for evidence of bias and autocorrelated forecast errors in the survey of professional forecasters; Kohlhas and Robertson (2022) for evidence of bias in professional forecasters'

Figure 4.1: Bias and Persistence of Forecast Error in the Survey Data



(A) Bias in forecast

(B) Persistence of forecast error

Note: This figure reports bias (Panel A) and persistence (Panel B) of households' inflation forecasts in the cross-section of the income distribution. Bias and persistence of each income percentile are calculated by the mean and serial correlation of forecast errors of households' inflation expectations for the next 12 months. Data are obtained from the Michigan Survey of Consumers between 1987:I and 2020:IV.

In Appendix F, we further document several additional results: (1) In the Survey of Consumer Expectations (SCE), conducted by the Federal Reserve of New York, the joint pattern of bias and persistence of forecast errors is similar to the one observed in the MSC, as shown in Table F.1. (2) At the individual level, after controlling for other observed characteristics such as age and resident state, the magnitude of the bias continues decreases with income, as shown in Table F.2.

The documented patterns of bias and persistence of forecast errors present challenges to the assumption of rational expectations. Within the rational expectation paradigm, under either full or noisy information, forecast errors should be zero on average. Noisy rational expectations can generate persistent forecast errors, but matching the increasing pattern of persistence across household income would require having richer households being less informed about the economy. Broer, Kohlhas, Mitman, and Schlafmann (2021) show that the dependence of sensitivity to information on income level can be rationalized when consumers are rationally inattentive, but this logic cannot explain the observed average bias. A notable exception is Farmer, Nakamura, and Steinsson (2023). They demonstrate that slow learning over the long-run trend with a unit root can create bias and persistence in forecast errors. However, the question of why bias and persistence move in opposite directions as household income changes remains unaddressed.

Another possibility is that households effectively report forecasts about the inflation rates of their own

forecasts, and Bhandari, Borovička, and Ho (2022) for evidence of bias in household forecasts.

consumption baskets, and that individual inflation rates differ by households' income level. However, as shown in Kaplan and Schulhofer-Wohl (2017), the annual inflation difference between the top and bottom income groups is about 1%, and the persistence of inflation rates across households income is virtually identical.[18]

In what follows, we set up a two-period consumption model with sticky nominal income and ambiguity about the inflation process. Applying the results developed in Section 3, we show that the inflation forecasts that arise from this a model are consistent with the type of cross-sectional distribution of bias and persistence documented above.

4.2 Model

Household problem There is a finite number of household groups indexed by $g \in 1, \dots, N$, with each group differing in their nominal income level, denoted by Y_g. Within each income group, there is a continuum of households indexed by i. We consider a simple, stylized consumption-saving problem where households only plan their consumption path for periods t and $t+1$. The utility function of the household is given by

$$U\left(C_{i,g,t}, C_{i,g,t+1}\right) = \frac{C_{i,g,t}^{1-\nu} - 1}{1-\nu} + \beta \frac{C_{i,g,t+1}^{1-\nu} - 1}{1-\nu},$$

where ν controls the degree of risk aversion of households.

Nominal income is rigid between t and $t+1$. A household in group g receives nominal income $P_t Y_g / 2$ in each period. The budget constraint is therefore given by

$$P_t C_{i,g,t} + P_{t+1} C_{i,g,t+1} = P_t Y_g.$$

Let $\pi_{t+1} \equiv (P_{t+1} - P_t)/P_t$ denote the inflation rate. The budget constraint of the household i can then be rewritten as

$$C_{i,g,t} + \left(1 + \pi_{t+1}\right) C_{i,g,t+1} = Y_g.$$

That is, a higher inflation rate makes consumption tomorrow more expensive relative to consumption today, adversely affecting households due to the nominal rigidity in their income.[19] In this problem, households only face uncertainty about the inflation rate. Once the belief about future inflation is determined, the optimal consumption plan follows. The following lemma directs us to concentrate on inflation expectations.

[18] See Table 3 in the main text and Figure 8 in the online appendix of Kaplan and Schulhofer-Wohl (2017).

[19] The assumption of complete rigid nominal income is only for simplicity. What is important is that relative to goods prices, nominal income changes at a slower rate.

Lemma 4.1. *Around the zero-inflation steady state:*

1. *The optimal consumption change, $c_{i,g,t}$, is proportional to the household's subjective expectation about inflation*
$$c_{i,g,t} = \frac{\beta^{1/\nu}}{1+\beta^{1/\nu}} \mathcal{F}_{i,g,t}\left[\pi_{t+1}\right]; \tag{4.2}$$

2. *The quadratic approximation of the utility function $U(C_{i,g,t}, C_{i,g,t+1})$ is given by*
$$U \approx Q(\mathcal{F}_{i,g,t}\left[\pi_{t+1}\right], \pi_{t+1}) = const - \delta_g \left(\mathcal{F}_{i,g,t}\left[\pi_{t+1}\right] - \pi_{t+1}\right)^2 - \chi_g \pi_{t+1} - \gamma_g \pi_{t+1}^2, \tag{4.3}$$
where δ_g, χ_g, and γ_g are positive when $\nu > 1$ and satisfy
$$\delta_g, \chi_g, \gamma_g \propto Y_g^{1-\nu}.$$

The first part of Lemma 4.1 states that there is a one-to-one mapping between the optimal choice of consumption and the household's subjective expectation about the inflation rate. A higher inflation expectation implies that future goods become more expensive relative to current goods, and households have an incentive to increase their current consumption.

Condition (4.2) allows us to express the utility in terms of expected inflation and actual inflation, since future consumption depends on the realized inflation. This quadratic approximation of the utility function is nested within the general specification (3.3) in Section 3.1. The utility function (4.3) reveals two important properties that are crucial in matching the data later on: (1) ceteris paribus, higher inflation lowers household welfare since both χ_g and γ_g are positive; (2) a higher level of income reduces households' exposure to variations in inflation since $(\delta_g, \chi_g, \gamma_g)$ are decreasing in Y_g.

Subjective expectations The inflation rate, π_t, follows an exogenous AR(1) process,
$$\pi_t = \rho \pi_{t-1} + \eta_t, \quad \text{with} \quad \eta_t \sim \mathcal{N}\left(0, \sigma_\eta^2\right),$$
and at each period t, household i receives a noisy private signal,
$$x_{i,g,t} = \pi_t + \epsilon_{i,g,t}, \quad \text{with} \quad \epsilon_{i,g,t} \sim \mathcal{N}\left(0, \sigma_\epsilon^2\right).$$
The information set of household i at time t is $\mathcal{I}_{i,g,t} = \{x_{i,g,t}, x_{i,g,t-1}, \ldots\}$. So far, this specification is similar to the structure used in the literature that studies survey evidence on inflation forecasts (Coibion and Gorodnichenko, 2015; Bordalo, Gennaioli, Ma, and Shleifer, 2020). It is also worth noting that the received information is independent of the household income level.

We depart from rational expectations by allowing agents perceive ambiguity about the mean of the innovation to the inflation,

$$\eta_t \sim \mathcal{N}\left(\mu_t, \sigma_\eta^2\right), \quad \text{and} \quad \mu_t \sim \mathcal{N}\left(0, \sigma_\mu^2\right),$$

where, again, σ_μ^2 corresponds to the amount of ambiguity. Let $h(x_{i,g}^t)$ denote households' strategy in forming expectations, i.e., $\mathcal{F}_{i,g,t}[\pi_{t+1}] = h(x_{i,g}^t)$. The mapping from subjective inflation expectations to optimal consumption in period t allows us to transform the original optimal consumption problem into an optimal forecasting problem that can be embedded into our general theoretical framework. Following the specification from Section 3, households maximize the following objective function:

$$\phi^{-1}\left(\int_{\mu^t} \phi\left(\mathbb{E}_{i,g,t}^{\mu^t}\left[Q\left(h\left(x^t\right), \pi_{t+1}\right)\right]\right) p\left(\mu^t\right) \mathrm{d}\mu^t\right). \tag{4.4}$$

By establishing the equivalence between the optimal forecasting problem (4.4) and the original optimal consumption problem, we are implicitly assuming that households respond to the survey questions about inflation by reporting beliefs that are consistent with their consumption decisions. The assumption is common in studies that connect survey data on expectations to structural models with ambiguity-averse agents, such as Bhandari, Borovička, and Ho (2022) and Pei (2023).

4.3 Inflation Forecasts under Ambiguity

Bayesian benchmark We begin by muting ambiguity ($\sigma_\mu^2 = 0$) and considering the predictions of this model under standard Bayesian expectations, characterized in the following proposition.

Proposition 4.1. *The Bayesian forecast is a weighted sum of the prior and the new signal,*

$$\mathbb{E}_{i,g,t}[\pi_{t+1}] = \omega\, \mathbb{E}_{i,g,t-1}[\pi_t] + (\rho - \omega)x_{i,g,t}, \tag{4.5}$$

and the average forecast error satisfies

$$\pi_{t+1} - \overline{\mathbb{E}}_{g,t}[\pi_{t+1}] = \frac{1}{1-\omega L}\eta_{t+1},$$

where the persistence ω is given by

$$\omega = \frac{1}{2}\left(\rho + \frac{\sigma_\epsilon^2 + \sigma_\eta^2}{\rho\sigma_\epsilon^2} - \sqrt{\left(\rho + \frac{\sigma_\epsilon^2 + \sigma_\eta^2}{\rho\sigma_\epsilon^2}\right)^2 - 4}\right).$$

Without ambiguity, households underreact to their signals due solely to the noise in their observations. As a result, the aggregate forecast errors are persistent over time, as discussed in Coibion and Gorodnichenko (2015). However, the persistence ω does not correlate with household income levels. Moreover, the unconditional mean of household forecasts coincides with that of actual inflation, leaving no room for permanent bias. These properties are inconsistent with the empirical facts documented in Section 4.1.

Ambiguity-averse households When households are ambiguity averse, their income level matters for their subjective beliefs. The following characterization directly applies the results developed in Section 3.

Proposition 4.2. *With ambiguity aversion, the individual subjective forecast satisfies*

$$\mathcal{F}_{i,g,t}[\pi_{t+1}] = \vartheta_g \mathcal{F}_{i,g,t-1}[\pi_t] + (1 + r_g)(\rho - \vartheta_g)x_{i,g,t} + (1 - \vartheta_g)\,\mathcal{B}_g, \tag{4.6}$$

and the average forecast error obeys

$$\pi_{t+1} - \overline{\mathcal{F}}_{g,t}[\pi_{t+1}] = \frac{1 + r_g}{1 - \vartheta_g L}\eta_{t+1} - \frac{r_g}{1 - \rho L}\eta_{t+1} - \mathcal{B}_g,$$

where the persistence ϑ_g is given by

$$\vartheta_g = \frac{1}{2}\left(\rho + \frac{(1 + w_g)\sigma_\eta^2 + \sigma_\epsilon^2}{\rho\sigma_\epsilon^2} - \sqrt{\left(\rho + \frac{(1 + w_g)\sigma_\eta^2 + \sigma_\epsilon^2}{\rho\sigma_\epsilon^2}\right)^2 - 4}\right) < \omega, \tag{4.7}$$

with $r_g > 0$, $w_g > 0$, and $\mathcal{B}_g > 0$.

Condition (4.6) shows that the subjective inflation expectation follows a law of motion similar to the Bayesian one but is subject to several important modifications. First, households consistently overestimate the inflation rate by an amount determined by the bias term $\mathcal{B}_g$.

Second, the persistence of forecast errors is smaller than in the Bayesian case, $\vartheta_g < \omega$, for all g. Households react to their signals as if they are overconfident à la Broer and Kohlhas (2019), as the perceived signal-to-noise ratio $(1 + w_g)\sigma_\eta^2/\sigma_\epsilon^2$ is larger. Thus, households rely less on their prior means, which lowers the persistence of their forecasts.

Third, households display an additional overreaction to their current signals relative to a Bayesian rule, captured by the term $1+r_g$. The forecasting rule (4.6) shares similar properties to the one under diagnostic expectations (Bordalo, Gennaioli, Ma, and Shleifer, 2020) in which forecasters overweight representative states. This additional response to the signal distinguishes it from a Bayesian rule.

In the setup with ambiguity, the magnitudes of overreaction and bias endogenously depend on how inflation enters households' payoff functions. As emphasized in Lemma 4.1, the parameters that govern a household's exposure to inflation decrease with their income level. This reduced exposure is isomorphic to a lower degree of ambiguity aversion. As a result, higher-income households are effectively less concerned about ambiguity, and their forecasting strategy is more aligned with the Bayesian benchmark. It follows that, qualitatively, richer households are less biased and their forecast errors are more persistent, which is consistent with the pattern observed in the data.

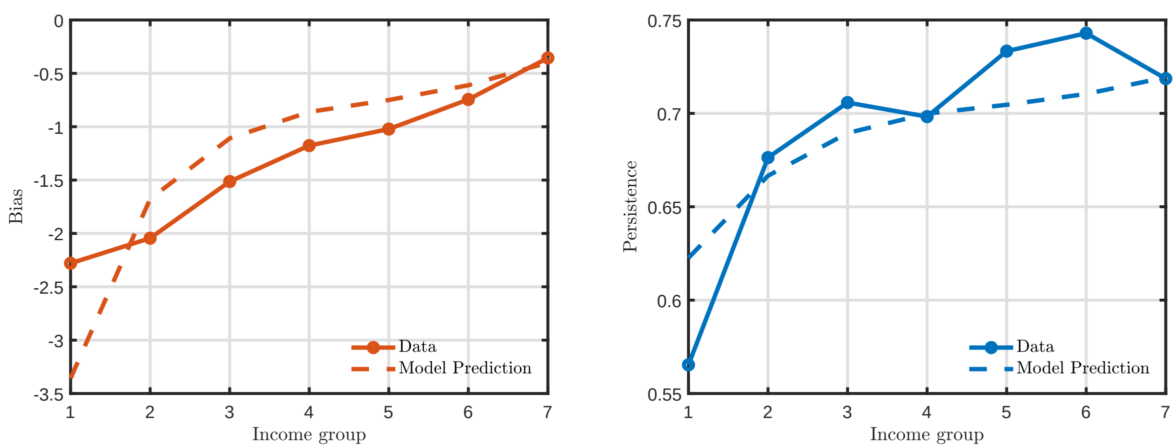
4.4 Results

To bring the model to the data, we set $\rho = 0.88$ and $\sigma_\eta = 0.72$ to fit the actual process for inflation, and fix the household's discount factor to $\beta = 0.99$. We normalize average income to 1, and set Y_g to match the share of income of each group in the MSC. The remaining parameters are the standard deviation of private information, σ_ϵ, the amount of ambiguity, σ_μ, the degree of ambiguity aversion, λ, and the degree of relative risk aversion, ν. We calibrate these parameters to match the persistence and bias of inflation forecast errors displayed in Figure 4.1. Operationally, we minimize

$$\sum_g \left(\text{Bias}_g^{\text{data}} - \text{Bias}_g^{\text{model}}\right)^2 + \left(\text{Persistence}_g^{\text{data}} - \text{Persistence}_g^{\text{model}}\right)^2.$$

The bias and persistence in the data are computed from the MSC using equation (4.1), and their model counterparts are the theoretical moments derived from Proposition 4.2.

FIGURE 4.2: Goodness of Fit.



(A) Bias in forecast (B) Persistence of forecast error

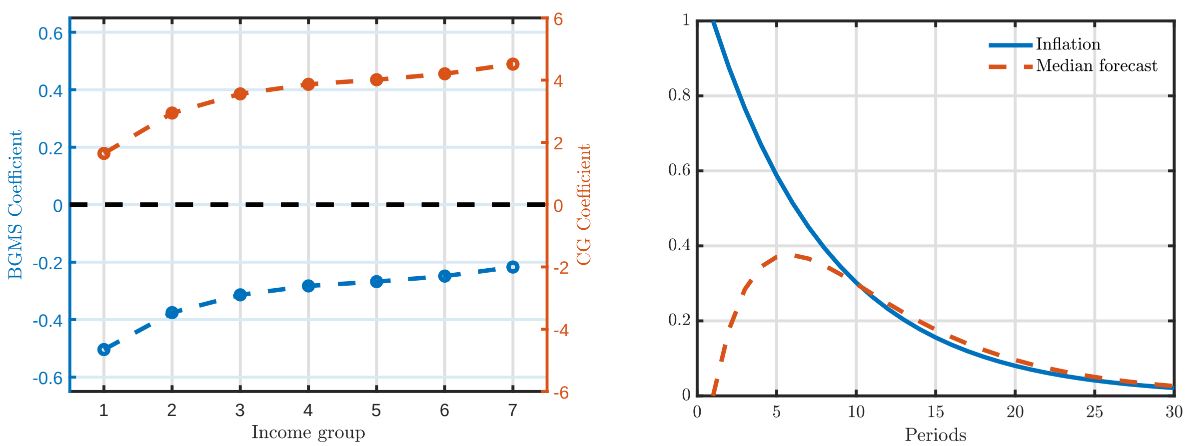
Table 4.1 presents the calibrated parameters. The model fit requires both signal noise and perceived ambiguity. Figure 4.2 displays the goodness of fit of our calibrated model. Given the rather parsimonious structure, our model captures the cross-sectional patterns of bias and persistence of inflation forecast errors reasonably well: richer households tend to have less bias in inflation forecasts but more persistent forecast errors.

Table 4.1: Calibrate Parameters

Param.	Value	Related to
σ_ϵ	4.00	std of noise in private signals
σ_μ	0.75	amount of ambiguity
λ	0.25	degree of ambiguity aversion
ν	1.48	risk aversion

It is challenging to capture these cross-sectional patterns in the Bayesian-expectation model without ambiguity aversion, where the bias is zero and the persistence does not depend on income. If households perceive ambiguity ($\sigma_\mu > 0$) but their preference is ambiguity-neutral ($\lambda = 0$), the persistence of forecast errors would increase with income but bias would still not be present.

Figure 4.3: Conditional and Unconditional Moments of Subjective Beliefs



(A) CG and BGMS coefficients (B) IRF of median forecast

Note: Panel A reports the theoretical CG and BGMS regression coefficients for each income group. Panel B reports the theoretical impulse response of the the consensus inflation expectation for the median income group.

Connection with existing survey evidence The predictions of our model are also broadly consistent with existing survey evidence on expectations. Due to dispersed, noisy information, the consensus forecast underreacts to new information, which implies a positive correlation between forecast error and forecast revision, as documented in Coibion and Gorodnichenko (2015) using the SPF. At the individual level, this correlation would necessarily be equal to zero under rational expectations. Bordalo, Gennaioli, Ma, and Shleifer (2020) show that, in the SPF, forecasters tend to overreact to their signals, resulting in a negative correlation at the individual level. The left panel of Figure 4.3 displays our model's prediction for these two coefficients for different income groups, which are consistent with the existing evidence. Notably, the CG coefficient increases with household income while the BGMS coefficient decreases in magnitude, which are the footprints of ambiguity aversion. The right panel shows the response of the average inflation forecast for the median income group to an inflationary shock over time. Due to the additional overreaction, associated with $1 + r_g$, the average forecast actually overshoots the true inflation. The implied sign-switching pattern is consistent with the empirical findings in Angeletos, Huo, and Sastry (2021).[20]

5 Application: Optimal Policy under Ambiguity

When agents in the private sector are ambiguity-averse, how should policy respond to changes in fundamentals, and how does additional sensitivity and bias affect policy design? In this section, we explore a policy game that builds on Barro and Gordon (1983) to shed light on these questions.

5.1 Environment

Time-consistent policy rule The policymaker chooses the inflation rate, π, so as to minimize the following social-loss function

$$\mathcal{L} = \mathbb{E}[U^2 + \omega(\pi - \pi^*)^2],$$

where U is the unemployment rate, π is the endogenous inflation rate, and π^* is an exogenous random inflation target that is drawn according to $\pi^* \sim \mathcal{N}(\overline{\pi}, \sigma_\pi^2)$.[21] The parameter ω balances the preference for lower unemployment against smaller deviations from the inflation target.

The policymaker faces a static Phillips curve that specifies how the unemployment rate reacts to average inflation surprises,

$$U = -\beta(\pi - \overline{\mathcal{F}}[\pi]),$$

where β is the slope of the Phillips curve. The unemployment rate is lower when actual inflation exceeds the average expected inflation. Importantly, agents in the private sector may not be rational, and the expectation operator corresponds to agents' subjective expectations. We have implicitly normalized the natural unemployment rate to zero, $U^* = 0$. Therefore, if $\pi = \overline{\mathcal{F}}[\pi] = \pi^*$, the policy maker would achieve the first-best outcome.

We consider a discretionary scenario in which the policymaker cannot commit to a policy ex ante.

[20] Note that in the IRF, we do not include the bias term. When conducting the projection method in Angeletos, Huo, and Sastry (2021), such bias terms are absorbed by the constant regressor.

[21] We normalize $\overline{\pi} = 0$ later on, but the main results do not hinge on this normalization.

As in Barro and Gordon (1983), the time-consistent inflation policy is given by

$$\pi = (1-\alpha)\,\pi^* + \alpha\,\overline{\mathcal{F}}[\pi], \quad \text{with} \quad \alpha \equiv \frac{\beta^2}{\omega + \beta^2}. \tag{5.1}$$

That is, the inflation rate is a weighted average between the exogenous inflation target and the economy-wide subjective inflation expectation.

Subjective expectations The exogenous inflation target cannot be perfectly observed by the public, and each agent i receives a private, noisy signal about it:

$$x_i = \pi^* + \epsilon_i, \quad \text{with} \quad \epsilon_i \sim \mathcal{N}(0, \sigma_\epsilon^2).$$

The noise here captures the notion that agents are inattentive to signals due to either attention costs or cognitive constraints. In addition to the informational frictions, agents in the private sector perceive ambiguity about the exogenous inflation target π^*. More specifically, agents believe that the mean of the target is ambiguous, that is,

$$\pi^* \sim \mathcal{N}(\mu, \sigma_\pi^2), \quad \text{and} \quad \mu \sim \mathcal{N}(\overline{\pi}, \sigma_\mu^2),$$

where σ_μ controls the amount of ambiguity.

The payoff of an agent depends on their subjective expectation and on inflation itself,

$$u(\mathcal{F}_i[\pi], \pi) = -(\mathcal{F}_i[\pi] - \pi)^2 - \chi\pi.$$

That is, agents care about the accuracy of their forecast, and an increase in the inflation rate directly reduces their utility. This utility function can be seen as a reduced-form version of the micro-founded structure introduced in Section 4,[22] which allows for biased subjective beliefs, in line with the data. In contrast with Section 4, inflation here is an endogenous equilibrium object that depends on the average subjective inflation expectation. As a result, the coordination motive is at play in shaping the subjective expectations.

[22]For simplicity, we abstract from the complications due to the second-order term and effectively assume households are homogeneous except for the dispersed information.

5.2 Optimal inflation policy

In equilibrium, the optimal policy that satisfies condition (5.1) can be characterized by a pair of policy parameters, $\mathcal{R}$ and $\mathcal{C}$, such that

$$\pi = \mathcal{R}\pi^* + \mathcal{C},$$

where $\mathcal{R}$ represents the responsiveness to the inflation target, and $\mathcal{C}$ determines the average level of inflation. How do aggregate inflation forecasts affect the optimal inflation policy?

We start with a rational expectations benchmark with informational frictions but in which agents do not perceive ambiguity.

Proposition 5.1. *With rational expectations ($\sigma_\mu^2 = 0$), the optimal policy rule is given by*

$$\mathcal{R}^{RE} = 1 - \alpha + \alpha \frac{\sigma_\pi^2}{\sigma_\pi^2 + (1-\alpha)^{-1}\sigma_\epsilon^2} \leq 1, \quad \textit{and} \quad \mathcal{C}^{RE} = 0.$$

Without ambiguity, the discretionary policy (5.1) can be viewed as a beauty contest game. When individual agents form expectations, they need to forecast the forecast of others, and the strength of these considerations is regulated by α. In the end, the average expectation is given by

$$\overline{\mathbb{E}}[\pi] = \frac{\sigma_\pi^2}{\sigma_\pi^2 + (1-\alpha)^{-1}\sigma_\epsilon^2}\,\pi^*,$$

where $(1-\alpha)^{-1}$ captures the discounting of private signals due to the coordination motive. Effectively, agents form their expectations using (w,α)-modified signals as in Section 3 with $w=0$. Although the responsiveness is dampened due to dispersed information, the inflation rate remains proportional to the target, and $\mathcal{C}^{\mathrm{RE}} = 0$. With full information rational expectations (FIRE), the dampening effect vanishes, and the endogenous inflation perfectly tracks the target: $\mathcal{R}^{\mathrm{FIRE}} = 1$.

Next, consider the case in which agents are ambiguity averse.

Proposition 5.2. *With ambiguity aversion, the aggregate subjective expectation is given by*

$$\overline{\mathcal{F}}[\pi] = \mathcal{S}\pi^* + \mathcal{B},$$

where

$$\mathcal{S} = \frac{(1+w)\,\sigma_\pi^2}{(1+w)\,\sigma_\pi^2 + (1-\alpha)^{-1}\sigma_\epsilon^2}, \qquad \mathcal{B} = \chi\lambda\,(1-\alpha+\alpha\mathcal{S})\,(1-\mathcal{S})\,\sigma_\mu^2,$$

and

$$w = \frac{\sigma_\mu^2/\sigma_\pi^2}{1 - 2\lambda\,(1-\alpha)^2\,(1-\mathcal{S})^2\,\sigma_\mu^2}.$$

The implied inflation policy satisfies

$$\mathcal{R} = 1 - \alpha + \alpha\mathcal{S} \in [\mathcal{R}^{RE}, 1], \quad \textit{and} \quad \mathcal{C} = \alpha\mathcal{B}.$$

Proposition 5.2 echoes our earlier emphasis on the effects of ambiguity aversion: the sensitivity to signals is amplified through the endogenously perceived prior, captured by the factor $(1+w)$, and the expectation is permanently biased upwards by $\mathcal{B}$.

The coordination motive, controlled by α, affects the equilibrium outcome in two ways. First, the policymaker internalizes the behavioral patterns of the private sector and adjusts the actual inflation process accordingly. This leads to higher responsiveness, $\mathcal{R} \geq \mathcal{R}^{\mathrm{RE}}$, and a lifted intercept, $\mathcal{C} > 0$. The extent to which the inflation policy inherits this behavior from the private sector is mechanically increasing in the degree of coordination motive α. Second, the effect of ambiguity on an agent's subjective belief also hinges on the strength of the coordination motive, as both w and $\mathcal{B}$ depend on α directly. Different from the analysis in Section 3, the economy here is inefficient even with complete information, and the socially optimal coordination level à la Angeletos and Pavan (2007) is $\widehat{\alpha} \equiv 1 - (1-\alpha)^2$, which is the key statistics that determines effects of ambiguity via w.[23]

FIGURE 5.1: Inflation Policy with and without Ambiguity

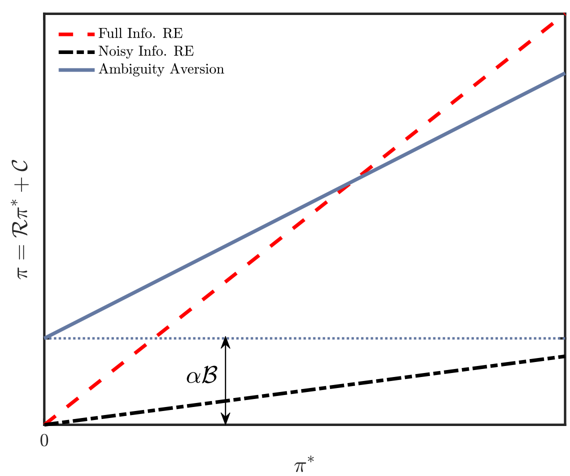


Figure 5.1 depicts the increased slope and shifted intercept that result from introducing ambiguity aversion. The red-dashed line represents the inflation policy under FIRE, which aligns exactly with the 45-degree line. The black-broken line displays the policy rule with noisy information but without ambiguity. Relative to FIRE, the only change is the dampened responsiveness. In both cases, there is no bias. When agents are ambiguity averse, we get the blue-solid line. The slope of the

[23] We provide a detailed analysis of inefficient economies in Appendix B.1.

policy approaches the one under FIRE, but this enhanced responsiveness is accompanied by a higher intercept due to the presence of bias in agents' forecasts.

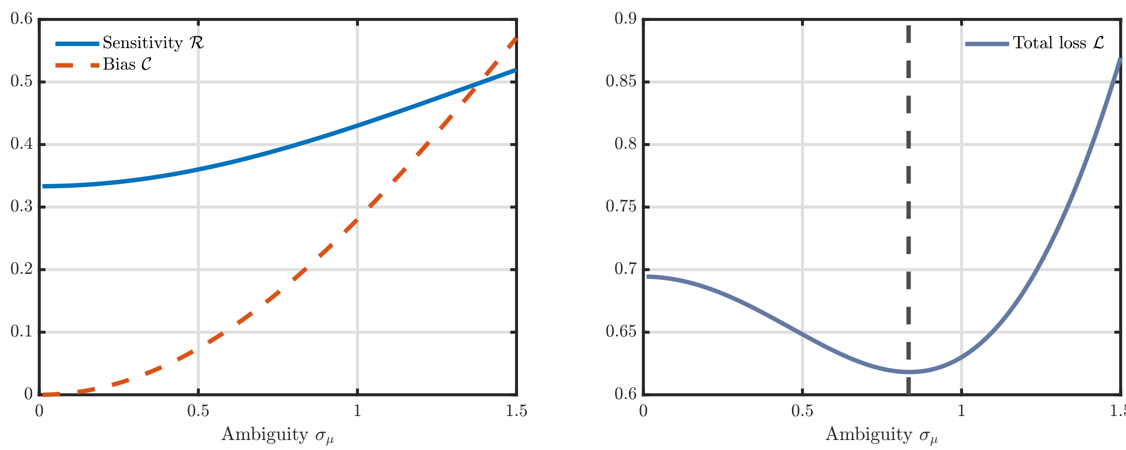
5.3 Ambiguity and social welfare

Given an inflation policy rule, social welfare losses can be expressed as follows:

$$\mathcal{L} = \mathbb{E}[U^2 + \omega(\pi - \pi^*)^2] = \frac{\omega}{\alpha}\left[(1-\mathcal{R})^2\sigma_\pi^2 + \mathcal{C}^2\right].$$

The inflation policy under FIRE, with $\mathcal{R} = 1$ and $\mathcal{C} = 0$, achieves the first best. Further, either an increase in responsiveness, towards $\mathcal{R} = 1$, or a reduction in the magnitude of the intercept, $\mathcal{C}$, would improve welfare. Intuitively, the policymaker would like the public to pay attention to the inflation target, facilitating the implementation of the desired inflation level. Any irrelevant deviation of the average expectation from π^* is socially inefficient.

FIGURE 5.2: Bias, Sensitivity, and Social Welfare



(A) Bias and sensitivity (B) Social loss $\mathcal{L}$

When information is noisy, the responsiveness of the private sector is dampened and the introduction of ambiguity can help approach the policy under FIRE. As can be observed in Figure 5.1, the combined effect of higher responsiveness and higher bias can lead to a better approximation of the FIRE first-best policy. This result is formalized in the following proposition.

Proposition 5.3. *Fixing the amount of noise σ_ϵ^2, there exists an intermediate level of ambiguity $\sigma_\mu^2 > 0$ that minimizes the social loss $\mathcal{L}$.*

Figure 5.2 helps illustrate the basic idea. The left panel shows how the responsiveness, $\mathcal{R}$, and the intercept, $\mathcal{C}$, vary with the amount of ambiguity, σ_μ. On one hand, the introduction of ambiguity enhances the responsiveness to signals, as the private sector shifts attention from their ambiguous priors to their signals, which contain relevant information about the policy target. On the other hand, it generates bias, as the private sector becomes increasingly concerned about the realization of a high-inflation model. These two competing effects generate a U-shape social loss function, which is minimized at an intermediate level of ambiguity.

This mechanism could justify, for instance, the recent switch by the Fed from a fixed inflation target to average inflation targeting (AIT). Jia and Wu (2022) argue that the adoption of AIT by the Fed in 2020 introduced ambiguity about the target to the extent that the time horizon used to compute the average is unclear.[24]

Alternative type of bounded rationality So far, we have rationalized the bias in agents' forecasts with ambiguity aversion. However, other types of bounded rationality could lead to the observed bias and this could lead to different policy implications.

As an example, consider a heterogeneous prior approach à la Angeletos, Collard, and Dellas (2018). Specifically, suppose that each agent i understands that $\pi^* \sim \mathcal{N}(\overline{\pi}, \sigma_\pi^2)$, but believes that all other agents perceive the inflation target with a bias, $\pi^* \sim \mathcal{N}(\overline{\pi} + \mathcal{B}, \sigma_\pi^2)$. Further, suppose all agents continue to receive a noisy signal about the inflation target. In this case, without ambiguity, the optimal inflation policy also features a non-zero intercept, but the welfare implication is significantly different.

Proposition 5.4. *With heterogeneous priors, the inflation policy rule is given by*

$$\mathcal{R} = \mathcal{R}^{RE}, \quad \textit{and} \quad \mathcal{C} = \alpha \left(\mathcal{R}^{RE} - 1 \right) \mathcal{B}.$$

The social loss monotonically increases with $\mathcal{B}$.

In contrast to the results derived above, if the observed bias in inflation forecasts is attributed to heterogeneous priors, a higher bias necessarily implies lower welfare. Since there is no concomitant effect on the sensitivity to signals, bias simply leads to less accurate average expectations.

6 Connection with Robust Preference

In this section, we document an intimate connection between the smooth model of ambiguity and the robust preferences model (Hansen and Sargent, 2001a,b). Although these two approaches to model uncertainty are conceptually different,[25] the main theoretical insights developed earlier about

[24] Jia and Wu (2022) also show that the ambiguity generated by this switch can be beneficial as it allows the Fed to increase its credibility. This benefit of ambiguity is different from the one we highlight. Their setup abstracts from ambiguity aversion, which plays a crucial role in our results.

[25] We refer to Hansen and Marinacci (2016) for a comprehensive discussion.

sensitivity and bias, along with the observational equivalence to Bayesian forecasts, also apply to models with robust preferences.

In parallel with the smooth model of ambiguity, we consider an efficient economy in which the utility function is given by

$$u(k_{it}, K_t, \xi_t) = -\frac{1}{2}\left[(1-\alpha)(k_{it}-\xi_t)^2 + \alpha(k_{it}-K_t)^2\right] - \chi\xi_t - \frac{1}{2}\gamma\xi_t^2,$$

and the signals follow the same general processes described in Section 3.1.

We model robust preferences following Hansen and Sargent (2005). Agents worry about potential model misspecification and consider a set of alternatives:

$$\begin{aligned} \max_{k_{it}} \min_{m_{it}} \quad & \mathbb{E}_{it}\left[u(k_{it}, K_t, \xi_t)\, m_{it} + \frac{1}{\varpi} m_{it} \log m_{it}\right] \\ \text{s.t.} \quad & m_{it} > 0, \quad \text{and} \quad \mathbb{E}_{it}[m_{it}] = 1. \end{aligned} \tag{6.1}$$

Here, each random variable m_{it} introduces a distorted distribution, generating an alternative model, where $\mathbb{E}_{it}[m_{it} \log m_{it}]$ corresponds to the relative entropy. The parameter ϖ controls the extent to which agents desire robustness. Agents then choose their strategies to optimize the worst-case scenario across the set of models under consideration.

Just as in the smooth model of ambiguity, under robust preferences, the subjective expectations and strategies of agents are jointly determined. Despite these complex interactions, the equilibrium strategy ultimately takes a simple form similar to the one in the smooth model.

Proposition 6.1. *The linear strategy under robust preferences takes the following form*

$$g\left(x_i^t\right) = (1+r)\, p\,(L; w, \alpha)\, x_{it} + \mathcal{B}.$$

1. *The polynomial matrix $p(L; w, \alpha)$ is the Bayesian forecasting rule with the (w, α)-modified signal process and w satisfies*
$$w = \frac{\varkappa_2}{1-\alpha}; \tag{6.2}$$
2. *The additional amplification, r, satisfies*
$$r = \frac{\varkappa_1 - \varkappa_2}{1-\alpha+\varkappa_2}; \tag{6.3}$$
3. *The level of bias, $\mathcal{B}$, satisfies*
$$\mathcal{B} = \chi\frac{r}{\gamma}. \tag{6.4}$$

The two endogenous scalars $(\varkappa_1, \varkappa_2)$ are such that

$$\varkappa_1 - \varkappa_2 = \varpi\gamma\,(1 - \alpha + \varkappa_1)\,\mathbb{V}_{it}(\xi_t) + \varpi\gamma\,(\alpha - \varkappa_2)\,\mathbb{COV}_{it}(K_t, \xi_t),$$
$$\varkappa_2 = \frac{\alpha\,(1-\alpha)}{\gamma}\,(\varkappa_1 - \varkappa_2) - \varpi\alpha\,(1-\alpha)\,\mathbb{DISP}(k_{it}),$$

where $\mathbb{V}_{it}(\xi_t)$, $\mathbb{COV}_{it}(K_t, \xi_t)$, and $\mathbb{DISP}(k_{it})$ denote the conditional volatility of the fundamental, the conditional covariance between the aggregate action and the fundamental, and the unconditional cross-sectional dispersion of individual actions, respectively.

Under robust preferences, agents want their actions to be robust across various models. In contrast, the smooth model of ambiguity is isomorphic to a setup where agents look for robustness across various priors about a fixed set of possible models (Hansen and Marinacci, 2016). Despite the conceptual differences between these two approaches, Proposition 6.1 shows that they are, in a sense, observationally equivalent. The following corollary in turn provides the condition under which the two models are equivalent in terms of individual strategies, with only a difference in the amount of bias.

Corollary 6.1. *Fix the objective environment. For a robust preferences model that satisfies*

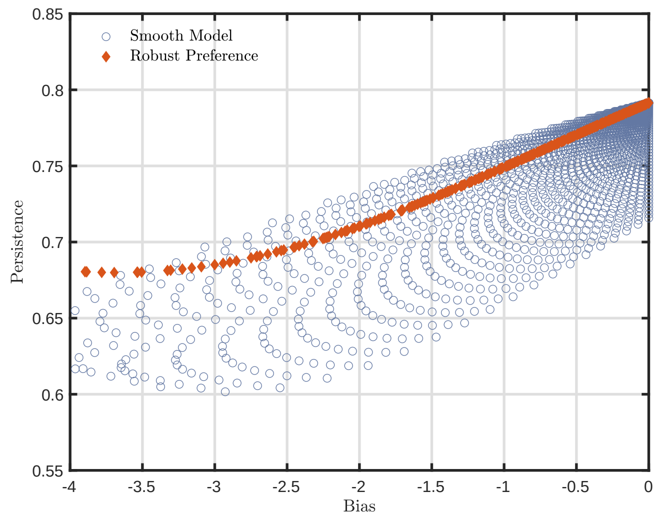
1. $w \geq 0, r \geq 0, \mathcal{S} \leq 1$,
2. $(1-\mathcal{S})\left(\frac{\gamma w}{(1+w)r} - \frac{(1-\alpha)(1+w)r}{w}\right) + \gamma > (1-\alpha)\frac{\mathbb{V}(\xi_t - K_t)}{\mathbb{V}(\xi_t)}$,

there exists a smooth model of ambiguity such that the equilibrium strategies in these two models are identical up to a constant.

Although the equilibrium strategies share a similar transformation in terms of Bayesian expectations, there are still some subtle differences between the two approaches. In the smooth model of ambiguity, we focus on the case where the model uncertainty is only about the prior mean of the fundamental. With robust preferences, the model uncertainty is about the entire stochasticity of the environment. When determining the endogenous amplification parameter w, the former only requires unconditional moments of aggregate variables, while the latter requires conditional moments of both aggregate and individual variables.

Another difference between the two approaches is that the smooth model allows a separation between the attitude towards ambiguity, λ, and the amount of ambiguity, σ_μ, whereas the robust preferences model is more parsimonious, with a single parameter, ϖ, to control the overall concern about model misspecification. To demonstrate the quantitative difference between the two approaches, we revisit the inflation-expectations application. Specifically, we keep the objective environment fixed,

Figure 6.1: A Comparison between SM and RP



Note: This figure compares the predictions on the bias and persistence of forecast errors between the smooth model and the robust preferences model. Each blue circle corresponds a smooth model with a choice of (λ, σ_μ^2). Each red dot corresponds to a robust preferences model with a choice of ϖ. The objective environment is fixed the same as that in Section 4.

including the exogenous stochastic process of inflation, ρ and σ_η, the private signal noise, σ_ϵ, and the preference parameters, β and ν. We examine the bias and persistence of forecast errors for households with median income. For the smooth model, we vary both the degree of ambiguity aversion, λ, and the amount of ambiguity, σ_μ^2. For the robust preferences model, we vary the parameter controlling the preference for robustness, ϖ.

Figure 6.1 presents the joint distribution of bias and persistence of forecast errors. Each blue circle corresponds to the outcome under a specific pair of (λ, σ_μ^2) in the smooth model, while each red diamond point represents the outcome under a particular choice of ϖ in the robust preferences model. Both approaches yield the same qualitative prediction: higher bias is associated with lower persistence. Moreover, due to the additional degree of freedom in the smooth model, the scatter plot covers the line implied by the robust preferences model. Quantitatively, the two approaches also produce comparable predictions.

7 Conclusion

In this paper, we study the effects of ambiguity in a general equilibrium environment with incomplete information. We provide an equivalence result that characterizes the equilibrium strategy as the solution to an adjusted single-agent Bayesian forecasting problem. Ambiguity aversion induces additional sensitivity to signals and a pessimistic bias, with both effects depending endogenously on agents' payoffs and the market structure. These properties allow us to match salient patterns observed in survey evidence on expectations, which are difficult to explain with rational expectation models. We also show that additional sensitivity and bias in subjective beliefs can change the effect of policy instruments, and that the exact micro-foundation for such belief distortions matters for policy design.

While our focus thus far has been on payoff structures that are homogeneous across agents, it would be interesting to explore the extent to which our main insights can be extended to network games. For example, firms in production supply chains may have concerns about their upstream suppliers and downstream customers perceiving different models. Another direction for future research is to explore how policymakers should incorporate substantial deviations of subjective beliefs from rational ones into the design of monetary and fiscal policies in quantitative models.

References

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

Altug, S., F. Collard, C. Çakmakl, S. Mukerji, and H. Özsöylev (2020): "Ambiguous business cycles: A quantitative assessment," *Review of Economic Dynamics*, 38, 220–237.

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

Anderson, E. W., L. P. Hansen, and T. J. Sargent (2003): "A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection," *Journal of the European Economic Association*, 1(1), 68–123.

Andolfatto, D., S. Hendry, and K. Moran (2008): "Are inflation expectations rational?," *Journal of Monetary Economics*, 55(2), 406–422.

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

Angeletos, G.-M., Z. Huo, and K. Sastry (2021): "Imperfect Macroeconomic Expectations: Evidence and Theory," in *NBER Macroeconomics Annual 2020*, vol. 35, chap. 1.

Angeletos, G.-M., and J. La'O (2010): "Noisy Business Cycles," *NBER Macroeconomics Annual*, 24(1), 319–378.

Angeletos, G.-M., and J. Lao (2020): "Optimal monetary policy with informational frictions," *Journal of Political Economy*, 128(3), 1027–1064.

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

Angeletos, G.-M., and K. A. Sastry (2021): "Managing expectations: Instruments versus targets," *The Quarterly Journal of Economics*, 136(4), 2467–2532.

Backus, D., A. Ferriere, and S. Zin (2015): "Risk and ambiguity in models of business cycles," *Journal of Monetary Economics*, 69, 42–63.

Barro, R. J., and D. B. Gordon (1983): "Rules, discretion and reputation in a model of monetary policy," *Journal of monetary economics*, 12(1), 101–121.

Bhandari, A., J. Borovička, and P. Ho (2022): "Survey Data and Subjective Beliefs in Business Cycle Models," *Working Paper*.

Bianchi, F., C. L. Ilut, and M. Schneider (2017): “Uncertainty Shocks, Asset Supply and Pricing over the Business Cycle,” *The Review of Economic Studies*, 85(2), 810–854.

Bianchi, F., S. C. Ludvigson, and S. Ma (2022): “Belief Distortions and Macroeconomic Fluctuations,” *American Economic Review*, 112(7), 2269–2315.

Bidder, R., and M. Smith (2012): “Robust animal spirits,” *Journal of Monetary Economics*, 59(8), 738 – 750.

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

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

Broer, T., A. Kohlhas, K. Mitman, and K. Schlafmann (2021): “Information and wealth heterogeneity in the macroeconomy,” .

Chen, Z., and L. Epstein (2002): “Ambiguity, Risk, and Asset Returns in Continuous Time,” *Econometrica*, 70(4), 1403–1443.

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

Collard, F., S. Mukerji, K. Sheppard, and J.-M. Tallon (2018): “Ambiguity and the historical equity premium,” *Quantitative Economics*, 9(2), 945–993.

Das, M., and A. van Soest (1999): “A panel data model for subjective information on household income growth,” *Journal of Economic Behavior & Organization*, 40(4), 409–426.

Dominitz, J. (1998): “Earnings Expectations, Revisions, and Realizations,” *The Review of Economics and Statistics*, 80(3), 374–388.

Dominitz, J., and C. F. Manski (1997): “Using Expectations Data To Study Subjective Income Expectations,” *Journal of the American Statistical Association*, 92(439), 855–867.

Elliott, G., I. Komunjer, and A. Timmermann (2008): “Biases in Macroeconomic Forecasts: Irrationality or Asymmetric Loss?,” *Journal of the European Economic Association*, 6(1), 122–157.

Epstein, L. G., and T. Wang (1994): “Intertemporal Asset Pricing under Knightian Uncertainty,” *Econometrica*, 62(2), 283–322.

Farmer, L., E. Nakamura, and J. Steinsson (2023): “Learning About the Long Run,” *Journal of Political Economy (forthcoming).*

Gallant, A. R., M. R Jahan-Parvar, and H. Liu (2018): "Does Smooth Ambiguity Matter for Asset Pricing?," *The Review of Financial Studies*, 32(9), 3617–3666.

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18(2), 141–153.

Hanany, E., and P. Klibanoff (2009): "Updating Ambiguity Averse Preferences," *The B.E. Journal of Theoretical Economics*, 9.

Hanany, E., P. Klibanoff, and S. Mukerji (2020): "Incomplete Information Games with Ambiguity Averse Players," *American Economic Journal: Microeconomics*, 12(2), 135–87.

Hansen, L. P., and M. Marinacci (2016): "Ambiguity Aversion and Model Misspecification: An Economic Perspective," *Statistical Science*, 31(4), 511–515.

Hansen, L. P., and T. J. Sargent (2001a): "Acknowledging Misspecification in Macroeconomic Theory," *Review of Economic Dynamics*, 4(3), 519 – 535.

——— (2001b): "Robust Control and Model Uncertainty," *American Economic Review*, 91(2), 60–66.

Hansen, L. P., and T. J. Sargent (2005): "Robust estimation and control under commitment," *Journal of Economic Theory*, 124(2), 258–301, Learning and Bounded Rationality.

Hansen, L. P., T. J. Sargent, and T. D. Tallarini (1999): "Robust Permanent Income and Pricing," *The Review of Economic Studies*, 66(4), 873–907.

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

Ilut, C., and H. Saijo (2021): "Learning, confidence, and business cycles," *Journal of Monetary Economics*, 117, 354–376.

Ilut, C., and M. Schneider (2014): "Ambiguous Business Cycles," *American Economic Review*, 104(8), 2368–2399.

Ilut, C. L., and M. Schneider (2022): "Modeling Uncertainty as Ambiguity: a Review," Working Paper 29915, National Bureau of Economic Research.

Jia, C., and J. C. Wu (2022): "Average Inflation Targeting: Time Inconsistency and Intentional Ambiguity," NBER Working Papers 29673, National Bureau of Economic Research, Inc.

Ju, N., and J. Miao (2012): "Ambiguity, Learning, and Asset Returns," *Econometrica*, 80(2), 559–591.

KAPLAN, G., AND S. SCHULHOFER-WOHL (2017): "Inflation at the household level," *Journal of Monetary Economics*, 91, 19–38.

KLIBANOFF, P., M. MARINACCI, AND S. MUKERJI (2005): "A smooth model of decision making under ambiguity," *Econometrica*, 73(6), 1849–1892.

KOHLHAS, A., AND D. ROBERTSON (2022): "Cautious Expectations," *Working Paper.*

KOHLHAS, A., AND A. WALTHER (2019): "Asymmetric Attention," *IIES miméo.*

LORENZONI, G. (2010): "Optimal monetary policy with uncertain fundamentals and dispersed information," *The Review of Economic Studies*, 77(1), 305–338.

LUO, Y., AND E. R. YOUNG (2010): "Risk-sensitive consumption and savings under rational inattention," *American Economic Journal: Macroeconomics*, 2(4), 281–325.

MALMENDIER, U., AND S. NAGEL (2016): "Learning from inflation experiences," *The Quarterly Journal of Economics*, 131(1), 53–87.

MARINACCI, M. (2015): "Model Uncertainty," *Journal of the European Economic Association*, 13(6), 1022–1100.

MASSENOT, B., AND Y. PETTINICCHI (2019): "Can households see into the future? Survey evidence from the Netherlands," *Journal of Economic Behavior & Organization*, 164, 77–90.

MIAO, J. (2009): "Ambiguity, Risk and Portfolio Choice under Incomplete Information," *Annals of Economics and Finance*, 10(2), 257–279.

MICHELACCI, C., AND L. PACIELLO (2019): "Ambiguous Policy Announcements," *The Review of Economic Studies*, 87(5), 2356–2398.

——— (2023): "Ambiguity Aversion and Heterogeneity in Households' Beliefs"," *The American Economic Journal: Macroeconomics (accepted).*

MOLAVI, P. (2022): "Simple Models and Biased Forecasts," *Working Paper.*

MORRIS, S., AND H. S. SHIN (2002): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1534.

PACIELLO, L., AND M. WIEDERHOLT (2014): "Exogenous information, endogenous information, and optimal monetary policy," *Review of Economic Studies*, 81(1), 356–388.

PEI, G. (2023): "Pessimism, Disagreement, and Economic Fluctuations," *Working Paper.*

Pope, D. G., and M. E. Schweitzer (2011): "Is Tiger Woods Loss Averse? Persistent Bias in the Face of Experience, Competition, and High Stakes," *American Economic Review*, 101(1), 129–57.

Rozsypal, F., and K. Schlafmann (2022): "Overpersistence Bias in Individual Income Expectations and its Aggregate Implications," *American Economic Journal: Macroeconomics (Forthcoming)*.

Townsend, R. M. (1983): "Forecasting the forecasts of others," *The Journal of Political Economy*, pp. 546–588.

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

Online Appendix

A Proofs of Main Results

In this appendix, we present the proofs of the main results from Section 3. We start by proving Proposition 3.2, which yields the fixed point conditions that characterize the equilibrium. We proceed by proving the general equivalence result, Proposition 3.3, based on which we can prove the existence of equilibrium, Proposition 3.1, as well as the comparative statics of sensitivity $\mathcal{S}$ and bias $\mathcal{B}$ with respect to the coordination motive α, Proposition 3.5.

Proof of Proposition 3.2. The equilibrium concept from Definition 3.1 is equivalent to the notion of ex-ante equilibrium from Hanany, Klibanoff, and Mukerji (2020). It is equivalent to the characterization of sequential equilibria with ambiguity (SEA) when conditional preferences are updated using the smooth rule of updating proposed in Hanany and Klibanoff (2009). The key for the equilibrium refinement of SEA is to ensure dynamic consistency, in the sense that ex-ante contingent plans are respected ex-post with the arrival of new information. Specifically, conditional on the realization of any possible history of private information, x_i^t, the optimal strategy of agent i maximizes their conditional preference, given by

$$\phi^{-1}\left(\int_{\mu^t}\phi\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it},K_t,\xi_t\right)\mid x_i^t\right]\right)\tilde{p}\left(\mu^t\mid x_i^t\right)\mathrm{d}\mu^t\right), \tag{A.1}$$

where $\mathbb{E}^{\mu^t}\left[u\left(k_{it},K_t,\xi_t\right)\mid x_i^t\right]$ denotes the expected utility conditional on x_i^t under a particular model μ^t. The interim belief system is characterized by a posterior belief $\tilde{p}\left(\mu^t \mid x_i^t\right)$ that follows the smooth rule of updating:

$$\tilde{p}\left(\mu^t \mid x_i^t\right) \propto \underbrace{\frac{\phi'\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it}^*, K_t^*, \xi_t\right)\right]\right)}{\phi'\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it}^*, K_t^*, \xi_t\right) \mid x_i^t\right]\right)}}_{\text{Weights}} \underbrace{p\left(x_i^t \mid \mu^t\right) p\left(\mu^t\right)}_{\text{Bayesian Kernel}},$$

where $\{k_{it}^*\left(x_i^t\right)\}_{x_i^t, i}$ denotes the equilibrium strategy profiles in the cross-section of the economy and $K_t^* \equiv \int_i k_{it}^* \mathrm{d}i$ denotes the equilibrium aggregate action.

The first-order condition of maximizing (A.1) with respect to k_{it} yields

$$\int_{\mu^t} \phi'\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it}, K_t, \xi_t\right) \mid x_i^t\right]\right) \frac{\partial \mathbb{E}^{\mu^t}\left[u\left(k_{it}, K_t, \xi_t\right) \mid x_i^t\right]}{\partial k_{it}} \tilde{p}\left(\mu^t \mid x_i^t\right) \mathrm{d}\mu^t = 0.$$

Since

$$\frac{\partial \mathbb{E}^{\mu^t}\left[u\left(k_{it}, K_t, \xi_t\right) \mid x_i^t\right]}{\partial k_{it}} = k_{it} - (1-\alpha)\, \mathbb{E}^{\mu^t}\left[\xi_t \mid x_i^t\right] - \alpha \mathbb{E}^{\mu^t}\left[K_t \mid x_i^t\right],$$

the first-order condition can be used to solve for the optimal strategies $\{k_{it}^*\left(x_i^t\right)\}_{x_i^t, i}$,

$$k_{it}^*(x_i^t) = \int_{\mu^t} \left((1-\alpha)\, \mathbb{E}^{\mu^t}\left[\xi_t \mid x_i^t\right] + \alpha \mathbb{E}^{\mu^t}\left[K_t \mid x_i^t\right]\right) \hat{p}\left(\mu^t \mid x_i^t\right) \mathrm{d}\mu^t,$$

with

$$\hat{p}\left(\mu^t \mid x_i^t\right) \equiv \frac{\phi'\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it}^*, K_t^*, \xi_t\right)\right]\right) p\left(x_i^t \mid \mu^t\right) p\left(\mu^t\right)}{\int_{\mu^t} \phi'\left(\mathbb{E}^{\mu^t}\left[u\left(k_{it}^*, K_t^*, \xi_t\right)\right]\right) p\left(x_i^t \mid \mu^t\right) p\left(\mu^t\right) \mathrm{d}\mu^t},$$

which completes the proof. □

Proof of Proposition 3.3. Following Huo and Pedroni (2020), we first consider a truncated version of our model. After solving this truncated version, the appropriate limits yield the desired result.[26]

Fix t and define

$$\vartheta \equiv \xi_t = \sum_{k=0}^{\infty} a_k \eta_{t-k}, \quad \text{and} \quad x_i \equiv x_i^t.$$

Let ϑ_q denote the MA(q) truncation of ϑ, such that

$$\vartheta_q = \sum_{k=0}^{q} a_k \eta_{t-k},$$

and let $x_{p,i}^J \equiv \{x_{p,it}, \cdots, x_{p,it-J}\}$, with $x_{p,it-k}$ denoting the MA(p) truncation of x_{it-k}.

Consider the truncated problem of forecasting the the fundamental ϑ_q given $x_{p,i}^N$. To further ease notation,

[26] See Online Appendix A.1 of Huo and Pedroni (2020) for detailed proofs.

define

$$\eta \equiv \begin{bmatrix} \eta_t \\ \vdots \\ \eta_{t-T} \end{bmatrix}, \quad \mu \equiv \begin{bmatrix} \mu_t \\ \vdots \\ \mu_{t-T} \end{bmatrix}, \quad \epsilon_i \equiv \begin{bmatrix} \epsilon_{it} \\ \vdots \\ \epsilon_{it-T} \end{bmatrix}, \quad \text{and} \quad \nu_i \equiv \begin{bmatrix} \eta \\ \epsilon_i \end{bmatrix}$$

Let R denote the length of x_{it}, and N the length of ϵ_{it}. It follows that, there exists a vector a with length $u \equiv T+1$, and a matrix B with dimensions $n \times m$, where $n \equiv R(T+1)$ and $m \equiv (1+N)(T+1)$, such that the truncated fundamental and the private signals are given by

$$\theta \equiv \vartheta_q = A\nu_i, \quad \text{with} \quad A \equiv \left[a', 0'_{m-u,1}\right], \quad \text{and} \quad x_i \equiv x_{p,i}^N = B\nu_i,$$

where $0_{m-u,1}$ is an $(m-u) \times 1$ vector of zeros. In the objective environment, ν_i is normally distributed,

$$\nu_i \sim \mathcal{N}(0, \Omega), \quad \text{with} \quad \Omega = \begin{bmatrix} \sigma_\eta^2 \, \mathrm{I}_u & 0 \\ 0 & \Xi \end{bmatrix},$$

where I_u denotes the identity matrix of size u and Ξ denotes the variance-covariance matrix of the $(m-u) \times 1$ vector of idiosyncratic shocks, ϵ_i. Subjectively, agents believe that η is drawn from a Gaussian distribution with variance-covariance matrix $\sigma_\eta^2 \, \mathrm{I}_u$ but there is uncertainty about its prior mean, denoted by μ. Ambiguity is then captured by the perception that

$$\eta \sim \mathcal{N}\left(\mu, \sigma_\eta^2 \, \mathrm{I}_u\right), \quad \text{and} \quad \mu \sim \mathcal{N}(0, \Omega_\mu), \quad \text{with} \quad \Omega_\mu \equiv \sigma_u^2 \, \mathrm{I}_u.$$

From Proposition 3.2, we know that the optimal strategy satisfies

$$k_i = \int_\mu \Big((1-\alpha)\, \mathbb{E}^\mu [\theta \mid x_i] + \alpha \mathbb{E}^\mu [K \mid x_i] \Big) \hat{p}(\mu \mid x_i) \, \mathrm{d}\mu,$$

with

$$\hat{p}(\mu \mid x_i) \propto \exp\left(-\lambda \mathbb{E}^\mu \left[u(k_i, K, \theta)\right]\right) p(x_i \mid \mu)\, p(\mu).$$

We proceed by using a guess-and-verify strategy. First, we guess that

$$k_i = h' B \nu_i + h_0.$$

We can show that ex-ante expected utility, under a particular model μ, is such that

$$\begin{aligned} \mathbb{E}^\mu [u(k_i, K, \theta)] = & -\mu' \left[\frac{1}{2}(1-\alpha)\, \mathcal{K}(A' - B'h)(A - h'B)\mathcal{K}' + \frac{1}{2}\gamma \mathcal{K} A' A \mathcal{K}'\right] \mu \\ & + \left[\frac{1}{2}(1-\alpha)\, h_0 (A - h'B)\mathcal{K}' + \frac{1}{2}\chi A \mathcal{K}'\right] \mu + \mu' \left[\frac{1}{2}(1-\alpha)\, h_0 \mathcal{K}(A' - B'h) + \frac{1}{2}\chi \mathcal{K} A'\right] \\ & \underbrace{-\frac{1}{2}(1-\alpha)(A - h'B)\,\Omega\,(A' - B'h) - \frac{1}{2}(1-\alpha)\, h_0^2 - \frac{1}{2}\alpha h' B\,(\mathrm{I}_m - \Lambda)\,\Omega B' h - \frac{1}{2}\gamma A \Omega A}_{\text{independent of } \mu}, \end{aligned} \tag{A.2}$$

where matrices $\mathcal{K}$ and Λ are such that

$$\mathcal{K} \equiv [\mathrm{I}_u, 0_{u,m-u}], \quad \text{and} \quad \Lambda \equiv \mathcal{K}'\mathcal{K}.$$

At the same time, we have that

$$p(\mu|x_i) \propto \exp\left(-\frac{1}{2}\mu'\left(\mathcal{K}B'(B\Omega B')^{-1}B\mathcal{K}'\right)^{-1}\mu + \frac{1}{2}\mu'\mathcal{K}(B\Omega B')^{-1}x_i + \frac{1}{2}x_i'(B\Omega B')^{-1}\mathcal{K}'\mu\right).$$

It follows that

$$\hat{p}(\mu \mid x_i) \propto \exp\left(-\frac{1}{2}\mu' S^{-1}\mu + \frac{1}{2}\mu' S^{-1}(Mx_i + \pi) + \frac{1}{2}(Mx_i + \pi)' S^{-1}\mu\right),$$

where matrices M, π, and S are such that

$$M \equiv S\mathcal{K}(B\Omega B')^{-1}, \quad \pi \equiv S\left[-\lambda(1-\alpha)h_0\mathcal{K}(A' - B'h) + \lambda\chi\mathcal{K}A'\right],$$

and

$$S \equiv \left(\mathcal{K}B'(B\Omega B')^{-1}B\mathcal{K}' + \Omega_\mu^{-1} - \lambda\left[(1-\alpha)\mathcal{K}(A' - B'h)(A - h'B)\mathcal{K}' + \gamma\mathcal{K}A'A\mathcal{K}'\right]\right)^{-1}.$$

Accordingly, we can show that the subjective expectations are such that

$$\int_\mu \mathbb{E}^\mu[\theta \mid x_i]\,\hat{p}(\mu \mid x_i)\,\mathrm{d}\mu = \mathrm{T}\,x_i + (A - \mathrm{T}B)\mathcal{K}'\left[S\mathcal{K}B'(B\Omega B')^{-1}x_i + \pi\right],$$

and

$$\int_\mu \mathbb{E}^\mu[K \mid x_i]\,\hat{p}(\mu \mid x_i)\,\mathrm{d}\mu = \mathrm{H}\,x_i + h'(B\Lambda - \mathrm{H}B)\mathcal{K}'\left[S\mathcal{K}B'(B\Omega B')^{-1}x_i + \pi\right] + h_0,$$

where matrices T and H are given by

$$\mathrm{T} \equiv A\Omega B'(B\Omega B')^{-1}, \quad \text{and} \quad \mathrm{H} \equiv B\Lambda\Omega B'(B\Omega B')^{-1}.$$

Therefore, matching coefficients leads to the following equilibrium conditions for h and h_0,

$$h' = (1-\alpha)\mathrm{T} + \alpha h'\mathrm{H} + \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'S\mathcal{K}B'(B\Omega B')^{-1}, \tag{A.3}$$

and

$$(1-\alpha)h_0 = \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'\pi, \tag{A.4}$$

In what follows, we first focus on equation (A.3). Through a sequence of lemmas, we show that this fixed-point problem for h can be recast as the solution of a pure forecasting problem. We then proceed to characterize h_0 using equation (A.4).

The next lemmas are organized as follows. Lemma A.1 rewrites the equilibrium condition for h described above as a beauty-contest problem with a modified variance-covariance matrix. Lemma A.2 establishes that h can be obtained by solving a forecasting problem with a modified variance-covariance matrix. Lemma A.3 maps this forecasting problem into the sum of two more easily interpreted forecasting problems. Lemma A.4 then combines these problems into the one that appears in the proposition. After the lemmas we take the limits of the truncated forecasting problem as $T \to \infty$.

Lemma A.1. *Define*

$$\hat{\Omega} \equiv \Omega + \mathcal{K}' W \mathcal{K}, \quad \hat{\mathrm{T}} \equiv A\hat{\Omega}B' \left(B\hat{\Omega}B'\right)^{-1}, \quad \hat{\mathrm{H}} \equiv B\Lambda\hat{\Omega}B' \left(B\hat{\Omega}B'\right)^{-1},$$

and

$$W \equiv \left(\Omega_{\mu}^{-1} - \lambda \left[(1-\alpha)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}' + \gamma\mathcal{K}A'A\mathcal{K}'\right]\right)^{-1}.$$

Then, the equilibrium h solves the following fixed-point problem

$$h' = (1-\alpha)\hat{\mathrm{T}} + \alpha h'\hat{\mathrm{H}}.$$

Proof. Using the Woodbury matrix identity, we have that

$$\begin{aligned}
\left(B\hat{\Omega}B'\right)^{-1} &= \left(B\Omega B' + B\mathcal{K}'W\mathcal{K}B'\right)^{-1} \\
&= \left(B\Omega B'\right)^{-1} - \left(B\Omega B'\right)^{-1} B\mathcal{K}' \left(\mathcal{K}B' \left(B\Omega B'\right)^{-1} B\mathcal{K}' + W^{-1}\right)^{-1} \mathcal{K}B' \left(B\Omega B'\right)^{-1} \\
&= \left(B\Omega B'\right)^{-1} - \left(B\Omega B'\right)^{-1} B\mathcal{K}'S\mathcal{K}B' \left(B\Omega B'\right)^{-1},
\end{aligned} \tag{A.5}$$

Then, if $\hat{h}$ is such that $\hat{h}' = (1-\alpha)\hat{\mathrm{T}} + \alpha\hat{h}'\hat{\mathrm{H}}$, we have that

$$\begin{aligned}
\hat{h}' &= (1-\alpha) A\hat{\Omega}B' \left(B\hat{\Omega}B'\right)^{-1} + \alpha\hat{h}'B\Lambda\hat{\Omega}B' \left(B\hat{\Omega}B'\right)^{-1} \\
&= (1-\alpha) A\left(\Omega + \mathcal{K}'W\mathcal{K}\right)B' \left(B\hat{\Omega}B'\right)^{-1} + \alpha\hat{h}'B\Lambda\left(\Omega + \mathcal{K}W\mathcal{K}'\right)B' \left(B\hat{\Omega}B'\right)^{-1} \\
&= (1-\alpha) A\Omega B' \left(B\hat{\Omega}B'\right)^{-1} + (1-\alpha) A\mathcal{K}'W\mathcal{K}B' \left(B\hat{\Omega}B'\right)^{-1} \\
&\quad + \alpha\hat{h}'B\Lambda\Omega B' \left(B\hat{\Omega}B'\right)^{-1} + \alpha\hat{h}'B\Lambda\mathcal{K}'W\mathcal{K}B' \left(B\hat{\Omega}B'\right)^{-1}.
\end{aligned}$$

Using equation (A.5), it follows that

$$\begin{aligned}
\hat{h}' = & \ (1-\alpha) A\Omega B' (B\Omega B')^{-1} - (1-\alpha) A\Omega B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
& + (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} - (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
& + \alpha\hat{h}' B\Lambda\Omega B' (B\Omega B')^{-1} - \alpha\hat{h}' B\Lambda\Omega B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
& + \alpha\hat{h}' B\Lambda\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} - \alpha\hat{h}' B\Lambda\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
= & \ \underbrace{(1-\alpha) A\Omega B' (B\Omega B')^{-1}}_{(1-\alpha)\mathrm{T}} + \underbrace{\alpha\hat{h}' B\Lambda\Omega B' (B\Omega B')^{-1}}_{\alpha\hat{h}'\mathrm{H}} \\
& - \underbrace{(1-\alpha) A\Omega B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1}}_{(1-\alpha)\mathrm{T}B\mathcal{K}'S\mathcal{K}B'(B\Omega B')^{-1}} - \underbrace{\alpha\hat{h}' B\Lambda\Omega B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1}}_{\alpha\hat{h}'\mathrm{H}B\mathcal{K}'S\mathcal{K}B'(B\Omega B')^{-1}} \\
& + (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} - (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
& + \alpha\hat{h}' B\Lambda\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} - \alpha\hat{h}' B\Lambda\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1}.
\end{aligned}$$

Further, notice that the terms in the second-to-last line can be rewritten as

$$\begin{aligned}
& (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} - (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
= & \ (1-\alpha) A\mathcal{K}' W \left(\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' + W^{-1}\right) \left(\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' + W^{-1}\right)^{-1} \mathcal{K} B' (B\Omega B')^{-1} \\
& - (1-\alpha) A\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' \left(\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' + W^{-1}\right)^{-1} \mathcal{K} B' (B\Omega B')^{-1} \\
= & \ (1-\alpha) A\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1},
\end{aligned}$$

and, similarly, the terms in the last line can be rewritten as

$$\begin{aligned}
& \alpha\hat{h}' B\Lambda\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} - \alpha\hat{h}' B\Lambda\mathcal{K}' W\mathcal{K} B' (B\Omega B')^{-1} B\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1} \\
& = \alpha\hat{h}' B\Lambda\mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1}.
\end{aligned}$$

Therefore, we have that

$$\hat{h}' = (1-\alpha)\mathrm{T} + \alpha\hat{h}'\mathrm{H} + \left[(1-\alpha)(A - \mathrm{T}B) + \alpha\hat{h}'(B\Lambda - \mathrm{H}B)\right] \mathcal{K}' S\mathcal{K} B' (B\Omega B')^{-1},$$

which is equivalent to the expression for h in equation (A.3). □

Lemma A.2. *Define*

$$\Omega_\Gamma \equiv \Gamma\hat{\Omega}, \quad \textit{with} \quad \Gamma \equiv \begin{bmatrix} \mathrm{I}_u & 0_{u,m-u} \\ 0_{m-u,u} & \frac{\mathrm{I}_{m-u}}{1-\alpha} \end{bmatrix}.$$

Then, the equilibrium h satisfies

$$h' = A\Omega_\Gamma B' (B\Omega_\Gamma B')^{-1}.$$

Proof. Follows directly from Lemma A.1 and Theorem 1 in Huo and Pedroni (2020). □

Lemma A.3. *Define*

$$\Delta \equiv \Gamma\Omega + \hat{w}\tau_\mu^{-1}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}, \quad \textit{and} \quad \tilde{\Omega}_\mu \equiv \left(\Omega_\mu^{-1} - \lambda\gamma\mathcal{K}A'A\mathcal{K}'\right)^{-1},$$

with the scalar $\hat{w}$ *given by*

$$\hat{w} \equiv \frac{\tau_\mu}{1 - \lambda(1-\alpha)\left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)}.$$

Then, the equilibrium h *satisfies*

$$h' = A\Delta B'\left(B\Delta B'\right)^{-1}.$$

Proof. It follows from Lemma A.2 that

$$\left(A - h'B\right)\Omega_\Gamma B' = 0,$$

and from the definition of Ω_Γ and $\tilde{\Omega}_\mu$ we have that

$$\Omega_\Gamma = \Gamma\Omega + \mathcal{K}'\left(\tilde{\Omega}_\mu^{-1} - \lambda(1-\alpha)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)^{-1}\mathcal{K}.$$

It is then sufficient to show that

$$\left(A - h'B\right)\left(\Gamma\Omega + \hat{w}\tau_\mu^{-1}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\right) = \left(A - h'B\right)\left(\Gamma\Omega + \mathcal{K}'\left(\tilde{\Omega}_\mu^{-1} - \lambda\left(1-\alpha\right)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)^{-1}\mathcal{K}\right),$$

or, equivalently,

$$\begin{aligned}
\hat{w}\tau_\mu^{-1}\left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K} &= \left(A - h'B\right)\mathcal{K}'\left(\tilde{\Omega}_\mu^{-1} - \lambda\left(1-\alpha\right)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)^{-1}\mathcal{K} \\
&= \left(A - h'B\right)\mathcal{K}'\left(\mathrm{I}_u - \lambda\left(1-\alpha\right)\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)^{-1}\tilde{\Omega}_\mu\mathcal{K}.
\end{aligned}$$

Thus, it is sufficient to establish that

$$\hat{w}\tau_\mu^{-1}\left(A - h'B\right)\mathcal{K}' = \left(A - h'B\right)\mathcal{K}'\left(\mathrm{I}_u - \lambda\left(1-\alpha\right)\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)^{-1}.$$

It follows that

$$\hat{w}\tau_\mu^{-1}\left(A - h'B\right)\mathcal{K}'\left(\mathrm{I}_u - \lambda\left(1-\alpha\right)\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right) = \left(A - h'B\right)\mathcal{K}',$$

which can be rewritten as

$$\hat{w}\tau_\mu^{-1}\left(1 - \lambda\left(1-\alpha\right)\left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)\right)\left(A - h'B\right)\mathcal{K}' = \left(A - h'B\right)\mathcal{K}'.$$

The definition of $\hat{w}$ then yields the result. □

Lemma A.4. *Define*

$$\bar{\Delta} \equiv \Gamma\Omega + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu\mathcal{K},$$

with the scalar $\hat{w}$ given by

$$\hat{w}=\frac{\tau_\mu}{1-\lambda\left(1-\alpha\right)\left(A-h'B\right)\mathcal{K}'\left(\Omega_\mu+\frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\mathcal{K}\left(A'-B'h\right)}.$$

Also, let the scalar $\hat{r}$ be given by

$$\hat{r}\equiv\frac{\hat{w}}{1+\hat{w}}\left(\frac{\lambda\gamma}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\left(A-h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A'.$$

Then, the equilibrium h satisfies

$$h'=(1+\hat{r})A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1}.$$

Proof. From the definition of $\tilde{\Omega}_\mu$ and Δ in Lemma A.3, we have that

$$\tilde{\Omega}_\mu\equiv\left(\Omega_\mu^{-1}-\lambda\gamma\mathcal{K}A'A\mathcal{K}'\right)^{-1}=\Omega_\mu+\frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'},$$

and

$$\Delta\equiv\Gamma\Omega+\hat{w}\tau_\mu^{-1}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}=\bar{\Delta}+\hat{w}\tau_\mu^{-1}\mathcal{K}'\left(\frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\mathcal{K}=\bar{\Delta}+s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K},$$

with $s\equiv\lambda\gamma\tau_\mu^{-1}/(1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A')$. Hence, it follows from the result in Lemma A.3 that

$$h'=A\left(\bar{\Delta}+s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'\left[B\left(\bar{\Delta}+s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'\right]^{-1},$$

and, therefore,

$$h'\left[B\left(\bar{\Delta}+s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'\right]=A\left(\bar{\Delta}+s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'.$$

Rearranging, we get

$$h'B\bar{\Delta}B'+s\hat{w}h'B\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}B'=A\bar{\Delta}B'+s\hat{w}A\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}B',$$

and right-multiplying both sides by $\left(B\bar{\Delta}B'\right)^{-1}$ yields

$$\begin{aligned}h'&=A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1}+s\hat{w}\left(A-h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1}\\&=A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1}+(1+\hat{w})\hat{r}\tau_\mu^{-1}A\mathcal{K}'\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1}.\end{aligned}$$

Then, from the definition of $\bar{\Delta}$ and using the fact that $\Omega_\mu=\tau_\mu\mathcal{K}\Omega\mathcal{K}'$ and $A\Gamma\Omega=\tau_\mu^{-1}A\mathcal{K}'\Omega_\mu\mathcal{K}$, it follows that

$$A\bar{\Delta}=A\left(\Gamma\Omega+\hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu\mathcal{K}\right)=\left(1+\hat{w}\right)\tau_\mu^{-1}A\mathcal{K}\Omega_\mu\mathcal{K}'.$$

Plugging this back into the equation for h' we obtain the desired result,

$$h'=\left(1+\hat{r}\right)A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1}.$$

$\square$

Parts 1 and 2 of Proposition 3.3. Given the result in Lemma A.4, we are left with taking the limit as $T \to \infty$ of the truncated problem. In particular, we have that

$$\lim_{T\to\infty} A\bar{\Delta}B' \left(B\bar{\Delta}B'\right)^{-1} = p\left(L; w, \alpha\right), \quad \lim_{T\to\infty} A\mathcal{K}'\Omega_\eta \mathcal{K}A' = \mathbb{V}(\xi_t),$$

$$\lim_{T\to\infty} \left(A - h'B\right)\mathcal{K}'\Omega_\eta\mathcal{K}(A' - Bh') = \mathbb{V}(\xi_t - K_t), \quad \lim_{T\to\infty} \left(A - h'B\right)\mathcal{K}'\Omega_\eta\mathcal{K}A' = \mathbb{COV}\left(\xi_t - K_t, \xi_t\right),$$

$$\lim_{T\to\infty} \frac{\left(A - h'B\right)\mathcal{K}'\Omega_\eta\mathcal{K}A'}{A\mathcal{K}'\Omega_\eta\mathcal{K}A'} = 1 - \mathcal{S}.$$

Let $w \equiv \lim_{T\to\infty}\hat{w}$, and $r \equiv \lim_{T\to\infty}\hat{r}$. Then, we can show that

$$\begin{aligned} r &= \lim_{T\to\infty} \frac{\hat{w}}{1+\hat{w}} \frac{\lambda\gamma\tau_\mu A\mathcal{K}'\Omega_\eta\mathcal{K}A'}{1-\lambda\gamma\tau_\mu A\mathcal{K}'\Omega_\eta\mathcal{K}A'} \frac{\left(A-h'B\right)\mathcal{K}'\Omega_\eta\mathcal{K}A'}{A\mathcal{K}'\Omega_\eta\mathcal{K}A'} \\ &= \frac{w}{1+w}\frac{\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\left(1-\mathcal{S}\right), \end{aligned} \tag{A.6}$$

and

$$\begin{aligned} w &= \lim_{T\to\infty} \frac{\tau_\mu}{1-\lambda\left(1-\alpha\right)\left(A-h'B\right)\mathcal{K}'\left(\Omega_\mu + \frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\mathcal{K}\left(A'-B'h\right)} \\ &= \lim_{T\to\infty} \frac{\tau_\mu}{1-\lambda\left(1-\alpha\right)\tau_\mu\left(\left(A-h'B\right)\mathcal{K}'\Omega_\eta\mathcal{K}\left(A'-B'h\right) + \hat{r}\frac{1+\hat{w}}{\hat{w}}A\mathcal{K}'\Omega_\eta\mathcal{K}\left(A'-B'h\right)\right)} \\ &= \frac{\tau_\mu}{1-\lambda\left(1-\alpha\right)\tau_\mu\left(\mathbb{V}\left(\xi_t-K_t\right) + r\frac{1+w}{w}\left(1-\mathcal{S}\right)\mathbb{V}\left(\xi_t\right)\right)}. \end{aligned}$$

Solving for w, we obtain

$$w = \frac{\tau_\mu}{1-\lambda\left(1-\alpha\right)\tau_\mu\left(\mathbb{V}\left(\xi_t-K_t\right) + \frac{\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2}{1-\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)}\right)}. \tag{A.7}$$

Lemma A.5 below establishes that $w \geq \tau_\mu$ and $r \geq 0$, which completes the proof of parts 1 and 2 of Proposition 3.3.

Lemma A.5. *If w and r satisfy equations (A.6) and (A.7), then $w \geq \tau_\mu$ and $r \geq 0$.*

Proof. The ex-ante objective of an agent i must obtain finite values under an equilibrium strategy $k_i = h'B\nu_i + h_0$. The ex-ante objective is given by

$$\begin{aligned} \mathcal{V} &= -\frac{1}{\lambda}\ln\left(\int_\mu \exp\left(-\lambda\mathbb{E}^\mu\left[u\left(-k_i, K, \theta\right)\right]\right)p\left(\mu\right)\mathrm{d}\mu\right) \\ &= \text{constant} - \frac{1}{\lambda}\ln\left(\int_\mu \exp\left(-\frac{1}{2}\mu'\bar{S}\mu + \mu'\bar{\pi}' + \bar{\pi}\mu\right)\mathrm{d}\mu\right), \end{aligned}$$

with the matrix $\bar{S}$ and the vector $\bar{\pi}$ given by

$$\begin{aligned}\bar{S} &\equiv \Omega_\mu^{-1} - \lambda\left(1-\alpha\right)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}' - \lambda\gamma\mathcal{K}A'A\mathcal{K}',\\ \bar{\pi} &\equiv -\lambda\frac{1}{2}\left(1-\alpha\right)h_0\left(A - h'B\right)\mathcal{K}' - \lambda\frac{1}{2}\chi A\mathcal{K}',\end{aligned}$$

where we used the fact that $\mathbb{E}^\mu\left[u\left(k_i, K, \theta\right)\right]$ is given by equation (A.2) and

$$p(\mu) = \left(2\pi\right)^{-u/2}\det\left(\Omega_\mu\right)^{-1/2}\exp\left(-\frac{1}{2}\mu'\Omega_\mu^{-1}\mu\right).$$

Thus, a necessary condition for $\mathcal{V}$ to be finite in equilibrium is for $\bar{S}$ to be positive definite; otherwise, the integral would become explosive.[27] Since

$$\tilde{\Omega}_\mu^{-1} = \Omega_\mu^{-1} - \lambda\gamma\mathcal{K}A'A\mathcal{K}',$$

it must be that

$$\tilde{\Omega}_\mu^{-1} - \lambda\left(1-\alpha\right)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}' \text{ is positive definite.}$$

Defining the vector $F \equiv \left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu$, it follows that

$$\begin{aligned}0 &\leq F\left(\tilde{\Omega}_\mu^{-1} - 2\lambda\left(1-\alpha\right)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)F'\\ &= \left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)\left(1 - \lambda\left(1-\alpha\right)\left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)\right).\end{aligned}$$

Let $x \equiv \left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)$, then we have that

$$x(1 - \lambda(1-\alpha)x) \geq 0 \quad \text{or} \quad x \geq \lambda(1-\alpha)x^2 \geq 0.$$

Hence, we have that $x \geq 0$, and $1 - \lambda(1-\alpha)x \geq 0$, which implies that

$$\hat{w} = \frac{\tau_\mu}{1 - \lambda(1-\alpha)x} \geq \tau_\mu,$$

and, since $w = \lim_{T\to\infty}\hat{w}$, it follows that $w \geq \tau_\mu$.

Next, for a contradiction, suppose that $r < 0$. Then, it follows from equation (A.6) and Assumption 2 that $\mathbb{COV}\left(\xi_t - K_t, \xi_t\right) < 0$. Further, we have that

$$\mathbb{COV}\left(\xi_t - K_t, \xi_t\right) = \mathbb{V}\left(\xi_t\right) - \left(1+r\right)\mathbb{COV}\left(\hat{K}_t, \xi_t\right),$$

[27] The same argument applies to how Assumption 2 ensures the problem is well defined. Specifically, a well-defined problem requires the choice set to be non-empty, which is equivalent to requiring $\bar{S}$ to be positive definite for at least one h. The necessary and sufficient condition for the existence of an h that makes $\bar{S}$ positive definite is that $\tilde{\Omega}_\mu$ is positive definite. Notice that $\tilde{\Omega}_\mu = \Omega_\mu + \frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}$. It is then straightforward to see that $1 - \lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A' > 0$ is the sufficient condition to ensure that $\tilde{\Omega}_\mu$ is positive definite. Taking the limit as $T \to \infty$, this is equivalent to Assumption 2.

where $\hat{K}_t \equiv K_t/(1+r)$ is the average optimal forecast of the fundamental ξ_t under the (w,α)-modified signal process (net of the bias $\mathcal{B}$, which is uncorrelated with ξ_t),[28] so that it must be that

$$0 \leq \mathbb{COV}\left(\hat{K}_t, \xi_t\right) \leq \mathbb{V}(\xi_t).$$

Hence, $\mathbb{COV}(\xi_t - K_t, \xi_t) < 0$ implies $r > 0$ and we have a contradiction. Therefore, $r \geq 0$. □

Part 3 of Proposition 3.3. Next, we switch focus to the level of the $\mathcal{B} \equiv \lim_{T\to\infty} h_0$. From equation (A.4) and the definition of π, we have that

$$(1-\alpha) h_0 = \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'S\left[-\lambda(1-\alpha) h_0 \mathcal{K}(A' - B'h) + \lambda\chi\mathcal{K}A'\right].$$

It is straightforward to see there exists a unique h_0 that satisfies this equation. We postulate that there exists $\tilde{\mu}$ such that

$$(1-\alpha) h_0 = \left[(1-\alpha) A + \alpha h' B\Lambda - h'B\right]\mathcal{K}'\tilde{\mu},$$

so that solving for $\tilde{\mu}$ pins down the unique h_0. To proceed, first replace the guess for h_0 on the RHS of equation (A.4),

$$\begin{aligned}
\mathrm{RHS} \equiv\ & \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'S\left[-\lambda(1-\alpha) h_0\mathcal{K}(A' - B'h) + \lambda\chi\mathcal{K}A'\right] \\
=\ & \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'S \\
& \times\left\{-\lambda\mathcal{K}(A' - B'h)\left[(1-\alpha)(A - h'B) + \alpha h'B(\Lambda - \mathrm{I}_m)\right]\mathcal{K}'\tilde{\mu} + \lambda\chi\mathcal{K}A'\right\}
\end{aligned}$$

Next, for the LHS of the equation, we have that

$$\mathrm{LHS} \equiv (1-\alpha) h_0 = \left[(1-\alpha) A + \alpha h'B\Lambda - h'B\right]\mathcal{K}'\tilde{\mu},$$

and, substituting the last h using equation (A.3), it follows that

$$\begin{aligned}
\mathrm{LHS} =\ & \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\left[\mathrm{I_m} - \mathcal{K}'\mathrm{S}\mathcal{K}\mathrm{B}'\left(\mathrm{B}\Omega\mathrm{B}'\right)^{-1}\mathrm{B}\right]\mathcal{K}'\tilde{\mu} \\
=\ & \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'S\left[S^{-1} - \mathcal{K}B'\left(B\Omega B'\right)^{-1}B\mathcal{K}'\right]\tilde{\mu} \\
=\ & \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{H}B)\right]\mathcal{K}'S \\
& \times\left\{\Omega_\mu^{-1} - \lambda\left[(1-\alpha)\mathcal{K}(A' - B'h)(A - h'B)\mathcal{K}' + \gamma\mathcal{K}A'A\mathcal{K}'\right]\right\}\tilde{\mu},
\end{aligned}$$

where the last equality uses the definition of S. Putting these results together, we have that

$$\begin{aligned}
\mathrm{LHS} - \mathrm{RHS} =\ & \left[(1-\alpha)(A - \mathrm{T}B) + \alpha h'(B\Lambda - \mathrm{HB})\right]\mathcal{K}'S \\
& \times\left[\Omega_\mu^{-1}\tilde{\mu} + \alpha\lambda\mathcal{K}(A' - B'h) h'B(\Lambda - \mathrm{I_m})\mathcal{K}'\tilde{\mu} - \lambda\gamma\mathcal{K}A'A\mathcal{K}'\tilde{\mu} - \lambda\chi\mathcal{K}A'\right].
\end{aligned}$$

[28] More precisely, notice that $\hat{K}_t = p(L; w, \alpha)\int x_{it} - \mathcal{B}/(1+r)$, and that it follows from Definition 3.2 that $\int \tilde{x}_{it} = \sqrt{1 + w\tau_\mu}\int x_{it}$ and $\tilde{\xi}_t = \sqrt{1 + w\tau_\mu}\,\xi_t$. Therefore, $\hat{K}_t = \int\tilde{\mathbb{E}}_{it}[\xi_t] - \mathcal{B}/(1+r)$ and $\mathbb{COV}\left(\hat{K}_t, \xi_t\right) = \mathbb{COV}\left(\int\tilde{\mathbb{E}}_{it}[\xi_t], \xi_t\right)$.

Since $\alpha\lambda\mathcal{K}\left(A'-B'h\right)h'B\left(\Lambda-\mathrm{I_m}\right)\mathcal{K}'=0$, a sufficient condition for $\mathrm{LHS}-\mathrm{RHS}=0$ is

$$\Omega_\mu^{-1}\tilde{\mu}-\lambda\gamma\mathcal{K}A'A\mathcal{K}'\tilde{\mu}-\lambda\chi\mathcal{K}A'=0,$$

which, using the Sherman-Morrison formula, implies that

$$\tilde{\mu}=\chi\lambda\left(\Omega_\mu^{-1}-\lambda\gamma\mathcal{K}A'A\mathcal{K}'\right)^{-1}\mathcal{K}A'=\chi\lambda\left(\mathrm{I_u}+\frac{\lambda\gamma\Omega_\mu\mathcal{K}\mathrm{A}'\mathrm{A}\mathcal{K}'}{1-\lambda\gamma\mathrm{A}\mathcal{K}'\Omega_\mu\mathcal{K}\mathrm{A}'}\right)\Omega_\mu\mathcal{K}A'.$$

Therefore, we have that

$$\begin{aligned}
h_0 &= (1-\alpha)^{-1}\left[(1-\alpha)A+\alpha h'B\Lambda-h'B\right]\mathcal{K}'\tilde{\mu}\\
&= \left(A-h'B\right)\mathcal{K}'\tilde{\mu}\\
&= \left(A-h'B\right)\mathcal{K}'\chi\lambda\left(\mathrm{I_u}+\frac{\lambda\gamma\Omega_\mu\mathcal{K}\mathrm{A}'\mathrm{A}\mathcal{K}'}{1-\lambda\gamma\mathrm{A}\mathcal{K}'\Omega_\mu\mathcal{K}\mathrm{A}'}\right)\Omega_\mu\mathcal{K}A'\\
&= \chi\lambda\tau_\mu\left(A-h'B\right)\mathcal{K}'\Omega_\eta\mathcal{K}A'\left(1+\frac{\lambda\gamma\tau_\mu A\mathcal{K}'\Omega_\eta\mathcal{K}A'}{1-\lambda\gamma\tau_\mu A\mathcal{K}'\Omega_\eta\mathcal{K}A'}\right).
\end{aligned}$$

Taking the limit we get

$$\mathcal{B}=\lim_{T\to\infty}h_0=\chi\lambda\tau_\mu\mathbb{COV}(\xi_t-K_t,\xi_t)\left(1+\frac{\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)}{1-\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)}\right)=\chi\frac{\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\left(1-\mathcal{S}\right),$$

which completes the proof of part 3 of the proposition. □

Proof of Proposition 3.1. Using the equivalence result from Proposition 3.3, establishing existence of an equilibrium reduces to showing that there exists a (w,r) pair that satisfies equations (A.6) and (A.7).

We start by using the intermediate value theorem to prove that there exists $w\in[\tau_\mu,\infty)$ that satisfies equation (A.7). Define

$$F(w)\equiv w\left[1-\lambda\left(1-\alpha\right)\tau_\mu\left(\mathbb{V}\left(\xi_t-K_t\right)+\frac{\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)^2\left(1-\mathcal{S}\right)^2}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\right)\right]-\tau_\mu,$$

such that $F(w)=0$ implies equation (A.7). Next, notice that as $w\to\infty$, private information becomes infinitely precise and, therefore, $p\left(L;w,\alpha\right)\to a(L)$, or $K_t\to\xi_t$. It follows that $\mathcal{S}\to1$ and $\mathbb{V}\left(\xi_t-K_t\right)\to0$, so that $\lim_{w\to\infty}F(w)=\infty$ and there must exist some finite $\bar{w}\geq\tau_\mu$ large enough such that $F(\bar{w})>0$. Next, notice that when $w=\tau_\mu$,

$$F\left(\tau_\mu\right)=-\lambda\left(1-\alpha\right)\tau_\mu^2\left(\mathbb{V}\left(\xi_t-K_t\right)+\frac{\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)^2\left(1-\mathcal{S}\right)^2}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\right)<0.$$

Thus, since $F(\cdot)$ is continuous, $F(\tau_\mu)<0$, and $F(\bar{w})>0$, there must exist some finite $w\in[\tau_\mu,\bar{w}]$ such that $F(w)=0$.

Further, from the definition of $\mathcal{S}$ we have that (see footnote 28)

$$1-\mathcal{S}=\frac{\mathbb{COV}\left(\xi_t-K_t,\xi_t\right)}{\mathbb{V}\left(\xi_t\right)} \quad\Rightarrow\quad 1-\mathcal{S}=1-(1+r)\frac{\mathbb{COV}(\hat{K}_t,\xi_t)}{\mathbb{V}\left(\xi_t\right)}.$$

Therefore, equation (A.6) becomes

$$r=\frac{w}{1+w}\frac{\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}{1-\lambda\gamma\tau_\mu\mathbb{V}\left(\xi_t\right)}\left(1-(1+r)\frac{\mathbb{COV}(\hat{K}_t,\xi_t)}{\mathbb{V}\left(\xi_t\right)}\right).$$

Since $\mathbb{COV}(\hat{K}_t,\xi_t)$ does not depend on r, the existence of w directly implies the existence of r. □

Proof of Proposition 3.5. According to equation (3.17), α affects the bias, $\mathcal{B}$, only through $1-\mathcal{S}$. It is, then, sufficient to prove that the sensitivity, $\mathcal{S}$, is decreasing in α. Further, since $\gamma=0$ implies $r=0$, α affects $\mathcal{S}$ only through the endogenous scalar w. To facilitate the proof, define an alternative signal process such that

$$\xi_t=a(L)\eta_t, \qquad \text{with} \quad \eta_t\sim\mathcal{N}(0,\sigma_\eta^2), \tag{A.8}$$
$$\hat{x}_{it}=m(L)\eta_t+n(L)\hat{\epsilon}_{it}, \qquad \text{with} \quad \hat{\epsilon}_{it}\sim\mathcal{N}(0,(1-\alpha)^{-1}(1+w)^{-1}\Sigma), \tag{A.9}$$

and let the corresponding optimal Bayesian forecast be given by

$$\hat{\mathbb{E}}_{it}[\xi_t]=\hat{p}(L;w,\alpha)\hat{x}_{it}.$$

It is straightforward to show that this signal process is equivalent to the (w,α)-modified signal process for Definition 3.2, that is

$$\hat{p}(L;w,\alpha)=p(L;w,\alpha).$$

For the current proof, this signal process is more helpful. Notice that $\mathcal{S}$ is affected by α only through $p(L;w,\alpha)$, since it is defined on the basis of the objective signal process.

In what follows, we first show that

$$\lim_{\alpha\to1^-}\frac{\mathrm{d}\mathbb{COV}\left(\xi_t-K_t,\xi_t\right)}{\mathrm{d}\alpha}>0.$$

We then prove, by contradiction, that there does not exist $\alpha\in[0,1)$ such that

$$\frac{\mathrm{d}\mathbb{COV}\left(\xi_t-K_t,\xi_t\right)}{\mathrm{d}\alpha}<0.$$

Then, the result follows by continuity of $\mathbb{COV}\left(\xi_t-K_t,\xi_t\right)$ with respect to α.

Step 1: $\lim_{\alpha\to1^-}\frac{\mathrm{d}\mathbb{COV}(\xi_t-K_t,\xi_t)}{\mathrm{d}\alpha}>0$:

It follows from equation (3.15) that $\lim_{\alpha\to1^-}w=\tau_\mu$. So, as $\alpha\to1^-$, the signals $\hat{x}_{it}$ become useless and, as a result,

$$\mathbb{COV}\left(K_t, \xi_t\right) = \mathbb{V}\left(K_t\right) = 0.$$

Further, since $w \geq \tau_\mu$, we have that

$$\lim_{\alpha \to 1^-} \frac{\mathrm{d}w}{\mathrm{d}\alpha} \leq 0 \quad \Rightarrow \quad \lim_{\alpha \to 1^-} \frac{\mathrm{d}\left(1-\alpha\right)\left(1+w\right)}{\mathrm{d}\alpha} < 0.$$

Therefore, at the limit of $\alpha \to 1^-$, an increase in α is akin to an increase in the variance of every idiosyncratic noise, which implies that (see Lemma D.2 in the Online Appendix D of Huo and Pedroni (2020)),

$$\lim_{\alpha \to 1^-} \frac{\mathrm{d}\,\mathbb{COV}\left(\xi_t - K_t, \xi_t\right)}{\mathrm{d}\alpha} > 0.$$

Step 2: $\frac{\mathrm{d}\mathbb{COV}(\xi_t - K_t, \xi_t)}{\mathrm{d}\alpha} > 0$ for all $\alpha \in [0,1)$:

Suppose the there exists $\alpha \in [0,1)$ such that $\frac{\mathrm{d}\mathbb{COV}(\xi_t - K_t, \xi_t)}{\mathrm{d}\alpha} < 0$. Then, by the intermediate value theorem and continuity of $\frac{\mathrm{d}\mathbb{COV}(\xi_t - K_t, \xi_t)}{\mathrm{d}\alpha}$, there must exist some $\alpha_\dagger$ such that

$$\left.\frac{\mathrm{d}\mathbb{COV}\left(\xi_t - K_t, \xi_t\right)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = 0 \quad \Rightarrow \quad \left.\frac{\mathrm{d}\left(1-\alpha\right)\left(1+w\right)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = 0 \quad \Rightarrow \quad \left.\frac{\mathrm{d}\mathbb{V}\left(\xi_t - K_t\right)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = 0,$$

since, for $\mathbb{COV}\left(\xi_t - K_t, \xi_t\right)$ not to change with α, it must be that the variance of the noise, $\left(1-\alpha\right)\left(1+w\right)$, is unchanged. Since

$$\frac{\mathrm{d}\left(1-\alpha\right)\left(1+w\right)}{\mathrm{d}\alpha} = -\left(1 + w\tau_\mu\right) + \left(1-\alpha\right)\frac{\mathrm{d}w}{\mathrm{d}\alpha},$$

it follows that

$$\left.\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} > 0.$$

However, since $\mathbb{COV}\left(\xi_t - K_t, \xi_t\right)$ and $\mathbb{V}\left(\xi_t - K_t\right)$ do not vary with α, it follows from equation (3.15) that

$$\left.\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = -\frac{\lambda\tau_\mu\left(\mathbb{V}\left(\xi_t - K_t\right) + \frac{\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2}{1-\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)}\right)}{\left[1 - \lambda\left(1-\alpha_\dagger\right)\tau_\mu\left(\mathbb{V}\left(\xi_t - K_t\right) + \frac{\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2}{1-\lambda\gamma\tau_\mu\mathbb{V}(\xi_t)}\right)\right]^2} < 0.$$

Thus, we have a contradiction, and we can conclude that

$$\frac{\mathrm{d}\mathbb{COV}\left(\xi_t - K_t, \xi_t\right)}{\mathrm{d}\alpha} < 0 \quad \Rightarrow \quad \frac{\mathrm{d}\mathcal{S}}{\mathrm{d}\alpha} < 0.$$

□

B Extensions

In this section, we consider two extensions to the baseline model setup. The first extension allows a more general utility specification, which covers economies with different forms of inefficiencies. The second extension is to the information structure, allowing the fundamental to depend on multiple aggregate shocks.

B.1 Inefficient economies

The economy in our baseline setup is assumed to be efficient under both complete and incomplete information. We now consider a generalized utility in the vein of Angeletos and Pavan (2007),

$$u\left(k_{it}, k_t, \xi_t\right) = -\frac{1}{2}\left[(1-\alpha)\left(k_{it}-\xi_t\right)^2 + \alpha\left(k_{it}-K_t\right)^2\right] - \frac{1}{2}\gamma\xi_t^2 - \chi\xi_t - \frac{1}{2}\psi\left(K_t-\xi_t\right)^2 - \phi K_t\xi_t - \varphi K_t, \quad \text{(B.1)}$$

which allows inefficiencies under both complete and incomplete information. Specifically, it can be shown that:

- Under complete information, the equilibrium allocation is such that $k_{it} = K_t = \xi_t$, whereas the efficient allocation is such that $k_{it} = K_t = \kappa_1^*\xi_t + \kappa_0^*$ with (κ_1^*, κ_0^*) being given by

$$\kappa_1^* = \frac{1-(\alpha-\psi)-\phi}{1-(\alpha-\psi)}, \quad \text{, and} \quad \kappa_0^* = \frac{\varphi}{1-(\alpha-\psi)}.$$

- Under incomplete information, the equilibrium degree of coordination is α, while the efficient degree of coordination is $\alpha^* = \alpha - \psi$.

The following proposition generalizes our equivalence result to the utility function in equation (B.1). The equilibrium strategy still features the simple form, which results in additional sensitivity and bias.

Proposition B.1. *The linear strategy in equilibrium takes the following form*

$$g(x_i^t) = (1+r)p(L; w, \alpha)x_{it} + \mathcal{B}. \quad \text{(B.2)}$$

1. *The polynomial matrix $p(L; w, \alpha)$ is the Bayesian forecasting rule with the (w, α)-modified signal process and w satisfies*

$$w = \frac{\tau_\mu}{(1+\nu_1) - \lambda\left(1-\alpha+\psi\right)\tau_\mu\left(\mathbb{V}\left(\xi_t - K_t\right) + \frac{\lambda\gamma\tau_\mu(1+\nu_2)\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2}{1-\lambda\gamma\tau_\mu(1+\nu_3)\mathbb{V}(\xi_t)}\right)};$$

2. *The additional amplification, r, satisfies*

$$r = \frac{\gamma\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)\left(1+\nu_2\right)}{1-\gamma\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)\left(1+\nu_3\right)}\frac{w}{1+w}\left(1-\mathcal{S}\right);$$

3. *The level of bias, $\mathcal{B}$, satisfies*

$$\mathcal{B} = \frac{\chi\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)\left(1-\mathcal{S}\right)+\nu_4}{1-\gamma\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)+\nu_5};$$

4. Relative to Proposition 3.3, the inefficiencies imply the following correction terms

$$\begin{aligned}
\nu_1 &\equiv \frac{\lambda^2\phi^2\tau_\mu^2\left(\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2-\mathbb{V}(\xi_t)\mathbb{V}(\xi_t-K_t)\right)-\lambda\phi\tau_\mu\mathbb{V}(\xi_t)(1-\mathcal{S})}{1-\lambda\tau_\mu\mathbb{V}(\xi_t)(2\gamma-\phi(1+\mathcal{S}))},\\
\nu_2 &\equiv 1-\frac{\phi}{\gamma}\left(2-\frac{\mathbb{V}(\xi_t-K_t)}{\mathbb{V}(\xi_t)(1-\mathcal{S})}\right),\\
\nu_3 &\equiv 1-\frac{\phi}{\gamma}(1+\mathcal{S}),\\
\nu_4 &\equiv \lambda\varphi\tau_\mu\left(\mathbb{V}(\xi_t)(1-\mathcal{S})-\mathbb{V}(\xi_t-K_t)\right)\\
&\quad-\lambda^2\tau_\mu^2\left(\phi(\chi-\varphi)+2\gamma\varphi\right)\left(\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2-\mathbb{V}(\xi_t)\mathbb{V}(\xi_t-K_t)\right)\\
\nu_5 &\equiv \lambda\tau_\mu\mathbb{V}(\xi_t)(2\phi\mathcal{S}-\gamma)+\lambda^2\tau_\mu^2\phi^2\left(\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2-\mathbb{V}(\xi_t)\mathbb{V}(\xi_t-K_t)\right).
\end{aligned}$$

It is easy to see that without inefficiencies, that is if $\psi=\phi=\varphi=0$, we have that $\nu_1=\nu_2=\nu_3=\nu_4=\nu_5=0$, and the formulas reduce to the ones in Proposition 3.3.

Proof of Proposition B.1. Consider the same truncated version of the model described in the proof of Proposition 3.3. For the utility in equation (B.1), we have that

$$\hat{p}(\mu|x_i)\propto\exp\left(-\frac{1}{2}\mu'S^{-1}\mu+\frac{1}{2}\mu'S^{-1}(Mx_i+\pi)+\frac{1}{2}(Mx_i+\pi)'S^{-1}\mu\right),$$

where matrices M, π, and S are such that

$$M\equiv S\mathcal{K}(B\Omega B')^{-1},\quad \pi\equiv S\left[-\lambda(1-\alpha^*)h_0\mathcal{K}(A'-B'h)+\lambda\chi\mathcal{K}A'+\lambda\varphi\mathcal{K}B'h\right],$$

and

$$\begin{aligned}
S\equiv\Big(&\mathcal{K}B'(B\Omega B')^{-1}B\mathcal{K}'+\Omega_\mu^{-1}-\lambda\left[(1-\alpha^*)\mathcal{K}(A'-B'h)(A-h'B)\mathcal{K}'+\gamma\mathcal{K}A'A\mathcal{K}'\right]\\
&-\lambda\phi\mathcal{K}(\Lambda B'hA+A'h'B\Lambda)\mathcal{K}'\Big)^{-1},
\end{aligned}$$

which, using $\phi=(1-\alpha^*)(1-\kappa_1^*)$, can be rearranged into

$$S=\left(\mathcal{K}B'(B\Omega B')^{-1}B\mathcal{K}'+\Omega_\mu^{-1}-\lambda\gamma^*\mathcal{K}A'A\mathcal{K}'-\lambda(1-\alpha^*)\mathcal{K}(\kappa_1^*A'-B'h)(\kappa_1^*A-h'B)\mathcal{K}'\right)^{-1},$$

where

$$\gamma^*\equiv\gamma+(1-\alpha^*)(1-(\kappa_1^*)^2).$$

We have the same equilibrium conditions for h and h_0 as in Proposition 3.3, equations (A.3) and (A.4), and the proof proceeds analogously and we keep the same structure to facilitate comparison.

Lemma B.1. *Define*

$$\hat{\Omega}\equiv\Omega+\mathcal{K}'W\mathcal{K},\quad \hat{\mathrm{T}}\equiv A\hat{\Omega}B'\left(B\hat{\Omega}B'\right)^{-1},\quad \hat{\mathrm{H}}\equiv B\Lambda\hat{\Omega}B'\left(B\hat{\Omega}B'\right)^{-1},$$

and

$$W \equiv \left(\Omega_\mu^{-1} - \lambda\gamma^* \mathcal{K} A' A \mathcal{K}' - \lambda\left(1-\alpha^*\right) \mathcal{K}\left(\kappa_1^* A' - B'h\right)\left(\kappa_1^* A - h'B\right)\mathcal{K}'\right)^{-1}.$$

Then, the equilibrium h solves the following fixed-point problem

$$h' = \left(1-\alpha\right)\hat{\mathrm{T}} + \alpha h'\hat{\mathrm{H}}.$$

Proof. This proof is exactly analogous to the proof of Lemma A.1. In particular, notice that W and S are still such that

$$S = \left(\mathcal{K}B'\left(B\Omega B'\right)^{-1} B\mathcal{K}' + W^{-1}\right)^{-1}.$$

□

Lemma B.2. *Define*

$$\Omega_\Gamma \equiv \Gamma\,\hat{\Omega}, \quad \textit{with} \quad \Gamma \equiv \begin{bmatrix} \mathrm{I}_u & 0_{u,m-u} \\ 0_{m-u,u} & \frac{\mathrm{I}_{m-u}}{1-\alpha} \end{bmatrix}.$$

Then, the equilibrium h satisfies

$$h' = A\Omega_\Gamma B'\left(B\Omega_\Gamma B'\right)^{-1}.$$

Proof. This lemma is exactly the same as Lemma A.2, and is repeated here just for convenience. □

Lemma B.3. *Define*

$$\Delta \equiv \Gamma\Omega + \hat{w}\tau_\mu^{-1}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K},$$

and

$$\tilde{\Omega}_\mu \equiv \left(\Omega_\mu^{-1} - \lambda\gamma^*\mathcal{K}A'A\mathcal{K}' - \lambda\left(1-\alpha^*\right)\mathcal{K}\left[\left(\kappa_1^*A' - B'h\right)\left(\kappa_1^*A - h'B\right) - \left(A' - B'h\right)\left(A - h'B\right)\right]\mathcal{K}'\right)^{-1},$$

with the scalar $\hat{w}$ given by

$$\hat{w} \equiv \frac{\tau_\mu}{1 - \lambda(1-\alpha^*)\left(A - h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A' - B'h\right)}.$$

Then, the equilibrium h satisfies

$$h' = A\Delta B'\left(B\Delta B'\right)^{-1}.$$

Proof. It follows from Lemma B.2 that

$$\left(A - h'B\right)\Omega_\Gamma B' = 0,$$

and from the definition of Ω_Γ and $\tilde{\Omega}_\mu$ we have that

$$\Omega_\Gamma = \Gamma\Omega + \mathcal{K}'\left(\tilde{\Omega}_\mu^{-1} - \lambda(1-\alpha^*)\mathcal{K}\left(A' - B'h\right)\left(A - h'B\right)\mathcal{K}'\right)^{-1}\mathcal{K}.$$

It is then sufficient to show that

$$(A-h'B)\left(\Gamma\Omega+\hat{w}\tau_\mu^{-1}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\right)=(A-h'B)\left(\Gamma\Omega+\mathcal{K}'\left(\tilde{\Omega}_\mu^{-1}-\lambda\left(1-\alpha^*\right)\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\right)^{-1}\mathcal{K}\right),$$

or, equivalently,

$$\begin{aligned}\hat{w}\tau_\mu^{-1}\left(A-h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}&=\left(A-h'B\right)\mathcal{K}'\left(\tilde{\Omega}_\mu^{-1}-\lambda\left(1-\alpha^*\right)\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\right)^{-1}\mathcal{K}\\&=\left(A-h'B\right)\mathcal{K}'\left(\mathrm{I}_u-\lambda\left(1-\alpha^*\right)\tilde{\Omega}_\mu\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\right)^{-1}\tilde{\Omega}_\mu\mathcal{K}.\end{aligned}$$

Thus, it is sufficient to establish that

$$\hat{w}\tau_\mu^{-1}\left(A-h'B\right)\mathcal{K}'=\left(A-h'B\right)\mathcal{K}'\left(\mathrm{I}_u-\lambda\left(1-\alpha^*\right)\tilde{\Omega}_\mu\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\right)^{-1}.$$

It follows that

$$\hat{w}\tau_\mu^{-1}\left(A-h'B\right)\mathcal{K}'\left(\mathrm{I}_u-\lambda\left(1-\alpha^*\right)\tilde{\Omega}_\mu\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\right)=\left(A-h'B\right)\mathcal{K}',$$

which can be rewritten as

$$\hat{w}\tau_\mu^{-1}\left(1-\lambda\left(1-\alpha^*\right)\left(A-h'B\right)\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K}\left(A'-B'h\right)\right)\left(A-h'B\right)\mathcal{K}'=\left(A-h'B\right)\mathcal{K}'.$$

The definition of $\hat{w}$ then yields the result. □

Lemma B.4. *Let*

$$\omega\equiv-\frac{\gamma}{\left(1-\alpha^*\right)\left(1-\kappa_1^*\right)},\quad v_1\equiv-\frac{\gamma\lambda}{\omega}\mathcal{K}A',\quad v_2\equiv\mathcal{K}\left(\omega A'-B'h\right),$$

and

$$c_{ij}\equiv v_i'\Omega_\mu v_j,\quad \textit{for } i,j\in\{1,2\},\quad \textit{and}\quad s_i\equiv\left(A-h'B\right)\mathcal{K}'\Omega_\mu v_i,\quad \textit{for } i\in\{1,2\}.$$

Further, define

$$\tilde{\Delta}\equiv\Gamma\Omega+\tilde{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu\mathcal{K},$$

with the scalar $\tilde{w}$ given by

$$\tilde{w}=\left(1+\frac{c_{11}s_2-\left(1+c_{12}\right)s_1}{\left(1+c_{12}\right)\left(1+c_{21}\right)-c_{11}c_{22}}\right)\hat{w},$$

and let the scalar $\tilde{r}$ be given by

$$\tilde{r}=-\frac{\frac{\lambda\gamma}{\omega}\left(c_{22}s_1-\left(1+c_{21}\right)s_2\right)+\left(1-\omega\right)\left(c_{11}s_2-\left(1+c_{12}\right)s_1\right)}{\left(1+c_{12}\right)\left(1+c_{21}\right)-c_{11}c_{22}+c_{11}s_2-\left(1+c_{12}\right)s_1}\frac{\tilde{w}}{1+\tilde{w}}.$$

Then, the equilibrium h satisfies

$$h'=\left(1+\tilde{r}\right)A\tilde{\Delta}B'\left(B\tilde{\Delta}B'\right)^{-1}.$$

Proof. From the definition of $\tilde{\Omega}_\mu$ in Lemma B.3, we have that

$$\tilde{\Omega}_\mu = \left(\Omega_\mu^{-1} + \lambda\left(1-\alpha^*\right)\left(1-\kappa_1^*\right)\mathcal{K}\left[A'\left(\omega A - h'B\right) + \left(\omega A' - B'h\right)A\right]\mathcal{K}'\right)^{-1},$$

with

$$\omega \equiv \frac{\left(1-\alpha^*\right)\left(1-(\kappa_1^*)^2\right)-\gamma^*}{\left(1-\alpha^*\right)\left(1-\kappa_1^*\right)} = -\frac{\gamma}{\left(1-\alpha^*\right)\left(1-\kappa_1^*\right)}.$$

Thus, defining

$$v_1 \equiv -\frac{\gamma\lambda}{\omega}\mathcal{K}A', \quad \text{and} \quad v_2 \equiv \mathcal{K}\left(\omega A' - B'h\right),$$

we can write

$$\tilde{\Omega}_\mu = (\Omega_\mu^{-1} + v_1 v_2' + v_2 v_1')^{-1},$$

and applying the Sherman-Morrison formula twice, we obtain

$$\tilde{\Omega}_\mu = \Omega_\mu + \frac{c_{11}\Omega_\mu v_2 v_2'\Omega_\mu + c_{22}\Omega_\mu v_1 v_1'\Omega_\mu - \left(1+c_{12}\right)\Omega_\mu v_1 v_2'\Omega_\mu - \left(1+c_{21}\right)\Omega_\mu v_2 v_1'\Omega_\mu}{\left(1+c_{12}\right)\left(1+c_{21}\right) - c_{11}c_{22}},$$

with

$$c_{ij} \equiv v_i'\Omega_\mu v_j, \quad \text{for } i,j \in \{1,2\}.$$

Thus, from the definition of Δ in Lemma B.3 and defining

$$\bar{\Delta} \equiv \Gamma\Omega + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu\mathcal{K},$$

we have that

$$\Delta = \Gamma\Omega + \hat{w}\tau_\mu^{-1}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K} = \bar{\Delta} + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K},$$

with

$$V \equiv \frac{c_{11}v_2 v_2' + c_{22}v_1 v_1' - \left(1+c_{12}\right)v_1 v_2' - \left(1+c_{21}\right)v_2 v_1'}{\left(1+c_{12}\right)\left(1+c_{21}\right) - c_{11}c_{22}}.$$

Hence, it follows from the result in Lemma B.3 that

$$h' = A\left(\bar{\Delta} + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}\right)B'\left[B\left(\bar{\Delta} + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}\right)B'\right]^{-1},$$

and, therefore,

$$h'\left[B\left(\bar{\Delta} + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}\right)B'\right] = A\left(\bar{\Delta} + \hat{w}\tau_\mu^{-1}\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}\right)B'.$$

Rearranging, we get

$$h'B\bar{\Delta}B' + \hat{w}\tau_\mu^{-1}h'B\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}B' = A\bar{\Delta}B' + \hat{w}\tau_\mu^{-1}A\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}B',$$

and right-multiplying both side by $\left(B\bar{\Delta}B'\right)^{-1}$ yields

$$h' = A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \hat{w}\tau_\mu^{-1}\left(A - h'B\right)\mathcal{K}'\Omega_\mu V\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1}.$$

Defining

$$s_i \equiv (A - h'B)\,\mathcal{K}'\Omega_\mu v_i, \quad \text{for } i \in \{1,2\},$$

we obtain

$$\begin{aligned} h' &= A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \hat{w}\tau_\mu^{-1}\frac{\left(c_{22}s_1 - (1+c_{21})\,s_2\right)v_1' + \left(c_{11}s_2 - (1+c_{12})\,s_1\right)v_2'}{(1+c_{12})\,(1+c_{21}) - c_{11}c_{22}}\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1} \\ &= A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \hat{w}\tau_\mu^{-1}\alpha_1 A\mathcal{K}'\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1} + \hat{w}\tau_\mu^{-1}\alpha_2 h'B\mathcal{K}'\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1} \end{aligned}$$

with

$$\alpha_1 \equiv \frac{-\left(c_{22}s_1 - (1+c_{21})\,s_2\right)\frac{\gamma\lambda}{\omega} + \left(c_{11}s_2 - (1+c_{12})\,s_1\right)\omega}{(1+c_{12})\,(1+c_{21}) - c_{11}c_{22}}, \quad \text{and} \quad \alpha_2 \equiv -\frac{\left(c_{11}s_2 - (1+c_{12})\,s_1\right)}{(1+c_{12})\,(1+c_{21}) - c_{11}c_{22}}.$$

Next, notice that

$$\mathcal{K}'\Omega_\mu\mathcal{K} = \tau_\mu\mathcal{K}'\mathcal{K}\Omega,$$

and

$$\bar{\Delta} = \Big((1+\hat{w})\,\mathcal{K}'\mathcal{K} + (1-\alpha)^{-1}\left(\mathrm{I}_m - \mathcal{K}'\mathcal{K}\right)\Big)\,\Omega,$$

so we have

$$\mathcal{K}'\Omega_\mu\mathcal{K} = \frac{\tau_\mu}{1+\hat{w}}\mathcal{K}'\mathcal{K}\bar{\Delta}.$$

Thus, it follows that

$$h' = \left(1 + \alpha_1\frac{\hat{w}}{1+\hat{w}}\right)A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \alpha_2\frac{\hat{w}}{1+\hat{w}}h'B\mathcal{K}'\mathcal{K}\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1},$$

or

$$h' = \beta_1 A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \beta_2 h'B\mathcal{K}'\mathcal{K}\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1}$$

with

$$\beta_1 \equiv 1 + \alpha_1\frac{\hat{w}}{1+\hat{w}}, \quad \text{and} \quad \beta_2 = \alpha_2\frac{\hat{w}}{1+\hat{w}}$$

Define

$$\tilde{\Delta} \equiv \left(\mathrm{I}_m - \beta_2\mathcal{K}'\mathcal{K}\right)\bar{\Delta},$$

and guess that

$$h' = \frac{\beta_1}{1-\beta_2}A\tilde{\Delta}B'\left(B\tilde{\Delta}B'\right)^{-1}.$$

It follows that

$$\begin{aligned} \frac{\beta_1}{1-\beta_2}A\tilde{\Delta}B'\left(B\tilde{\Delta}B'\right)^{-1}B\bar{\Delta}B' &= \beta_1 A\bar{\Delta}B' + \frac{\beta_1}{1-\beta_2}\beta_2 A\tilde{\Delta}B'\left(B\tilde{\Delta}B'\right)^{-1}B\mathcal{K}'\mathcal{K}\bar{\Delta}B' \\ \frac{\beta_1}{1-\beta_2}A\tilde{\Delta}B'\left(B\tilde{\Delta}B'\right)^{-1}B\left(\mathrm{I}_m - \beta_2\mathcal{K}'\mathcal{K}\right)\bar{\Delta}B' &= \beta_1 A\bar{\Delta}B' \\ \beta_1 A\tilde{\Delta}B' &= \beta_1\left(1-\beta_2\right)A\bar{\Delta}B', \end{aligned}$$

which verifies the guess, since $A\tilde{\Delta} = (1-\beta_2)\, A\bar{\Delta}$. Finally, we have that

$$\begin{aligned}\tilde{\Delta} &= (\mathrm{I}_m - \beta_2 \mathcal{K}'\mathcal{K}) \left((1+\hat{w})\, \mathcal{K}'\mathcal{K} + (1-\alpha)^{-1} (\mathrm{I}_m - \mathcal{K}'\mathcal{K}) \right) \Omega \\ &= \left((1-\beta_2)\,(1+\hat{w})\, \mathcal{K}'\mathcal{K} + (1-\alpha)^{-1} (\mathrm{I}_m - \mathcal{K}'\mathcal{K}) \right) \Omega \\ &= \Gamma\Omega + \tilde{w}\tau_\mu^{-1} \mathcal{K}'\Omega_\mu \mathcal{K},\end{aligned}$$

with

$$\tilde{w} = (1-\beta_2)\,(1+\hat{w}) - 1 = \left(1 - \alpha_2 \frac{\hat{w}}{1+\hat{w}}\right)(1+\hat{w}) - 1 = (1-\alpha_2)\hat{w}.$$

and

$$\tilde{r} = \frac{\beta_1}{1-\beta_2} - 1 = \frac{1+\alpha_1 \frac{\hat{w}}{1+\hat{w}}}{1-\alpha_2 \frac{\hat{w}}{1+\hat{w}}} - 1 = \frac{\alpha_1+\alpha_2}{1-\alpha_2} \frac{\tilde{w}}{1+\tilde{w}}.$$

Substituting the definitions of α_1 and α_2 yields the result. □

Parts 1 and 2 of Proposition B.1. Given the result in Lemma B.4, we are left with taking the limit as $T \to \infty$ of the truncated problem. In particular, we have that

$$\begin{aligned}\lim_{T\to\infty} c_{11} &= \tau_\mu \left(\frac{\gamma\lambda}{\omega}\right)^2 \mathbb{V}(\xi_t)\,, \quad \lim_{T\to\infty} c_{12} = \lim_{T\to\infty} c_{21} = -\tau_\mu \frac{\gamma\lambda}{\omega} \mathbb{COV}(\omega\xi_t - K_t, \xi_t)\,, \\ \lim_{T\to\infty} c_{22} &= \tau_\mu \mathbb{V}(\omega\xi_t - K_t)\,, \quad \lim_{T\to\infty} s_1 = -\tau_\mu \frac{\gamma\lambda}{\omega} \mathbb{COV}(\xi_t - K_t, \xi_t)\,, \quad \text{and} \\ \lim_{T\to\infty} s_2 &= \tau_\mu \mathbb{COV}(\omega\xi_t - K_t, \xi_t - K_t)\,.\end{aligned}$$

Notice that

$$\hat{w} = \frac{\tau_\mu}{1 - \lambda(1-\alpha^*) \left\{ (A - h'B)\, \mathcal{K}'\Omega_\mu \mathcal{K}\, (A' - B'h) + \frac{c_{11}s_2^2 + c_{22}s_1^2 - (2+c_{12}+c_{21})s_1 s_2}{(1+c_{12})(1+c_{21}) - c_{11}c_{22}} \right\}}.$$

Let $w \equiv \lim_{T\to\infty} \tilde{w}$, and $r \equiv \lim_{T\to\infty} \tilde{r}$. Using equations $\omega = \frac{\gamma}{(1-\alpha^*)(1-\kappa_1^*)}$, $\alpha^* = \alpha - \psi$, and $(1-\alpha^*)(1-\kappa_1^*) = \phi$, in order to return to primitive parameters, it follows that

$$w = \frac{\tau_\mu \left(1 + \lambda\,(1-\alpha+\psi)\, r\mathbb{V}(\xi_t)\,(1-\mathcal{S})\right)}{1 - \lambda\,(1-\alpha+\psi)\,\tau_\mu \left(\mathbb{V}(\xi_t - K_t) + r\mathbb{V}(\xi_t)\,(1-\mathcal{S})\right) + \nu_1},$$

and

$$r = \frac{\gamma\lambda\tau_\mu \mathbb{V}(\xi_t)\,(1+\nu_2)}{1 - \gamma\lambda\tau_\mu \mathbb{V}(\xi_t)\,(1+\nu_3)} \frac{w}{1+w}\,(1-\mathcal{S})\,,$$

with

$$
\begin{aligned}
\nu_1 &\equiv \frac{\lambda^2\phi^2\tau_\mu^2\left(\mathbb{V}\left(\xi_t\right)^2\left(1-\mathcal{S}\right)^2-\mathbb{V}\left(\xi_t\right)\mathbb{V}\left(\xi_t-K_t\right)\right)-\lambda\phi\tau_\mu\mathbb{V}\left(\xi_t\right)\left(1-\mathcal{S}\right)}{1-\lambda\tau_\mu\mathbb{V}\left(\xi_t\right)\left(2\gamma-\phi\left(1+\mathcal{S}\right)\right)},\\
\nu_2 &\equiv 1-\frac{\phi}{\gamma}\left(2-\frac{\mathbb{V}\left(\xi_t-K_t\right)}{\mathbb{V}\left(\xi_t\right)\left(1-\mathcal{S}\right)}\right),\\
\nu_3 &\equiv 1-\frac{\phi}{\gamma}\left(1+\mathcal{S}\right).
\end{aligned}
$$

This completes the proof of parts 1 and 2 of Proposition B.1.

Part 3 of Proposition B.1. Next, we switch focus to the level of the $\mathcal{B}\equiv\lim_{T\to\infty}h_0$. From equation (A.4) and the definition of π, we have that

$$
\begin{aligned}
\left(1-\alpha\right)h_0 = &\left[\left(1-\alpha\right)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\\
&\times\mathcal{K}'S\left[-\lambda\left(1-\alpha^*\right)h_0\mathcal{K}\left(A'-B'h\right)+\lambda\chi\mathcal{K}A'+\lambda\varphi\mathcal{K}B'h\right],
\end{aligned}
$$

which, using $\varphi=(1-\alpha^*)\kappa_0^*$ and defining $\chi^*\equiv\chi+(1-\alpha^*)\kappa_0^*$, can be rewritten as

$$
\begin{aligned}
\left(1-\alpha\right)h_0 = &\left[\left(1-\alpha\right)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\\
&\times\mathcal{K}'S\left[-\lambda\left(1-\alpha^*\right)\left(h_0+\kappa_0^*\right)\mathcal{K}\left(A'-B'h\right)+\lambda\chi^*\mathcal{K}A'\right].
\end{aligned}
\tag{B.3}
$$

It is straightforward to see there exists a unique h_0 that satisfies this equation. We postulate that there exists $\tilde{\mu}$ such that

$$
\left(1-\alpha\right)h_0=\left[(1-\alpha)A+\alpha h'B\Lambda-h'B\right]\mathcal{K}'\tilde{\mu},
$$

so that solving for $\tilde{\mu}$ pins down the unique h_0. To proceed, first replace the guess for h_0 on the RHS of equation (B.3),

$$
\begin{aligned}
\mathrm{RHS}\equiv &\left[\left(1-\alpha\right)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\mathcal{K}'S\left[-\lambda\left(1-\alpha^*\right)\left(h_0+\kappa_0^*\right)\mathcal{K}\left(A'-B'h\right)+\lambda\chi^*\mathcal{K}A'\right]\\
= &\left[\left(1-\alpha\right)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\mathcal{K}'S\\
&\times\left\{-\lambda\mathcal{K}\left(A'-B'h\right)\left(1-\alpha^*\right)\left\{\frac{\left[\left(1-\alpha\right)\left(A-h'B\right)+\alpha h'B\left(\Lambda-\mathrm{I}_m\right)\right]\mathcal{K}'\tilde{\mu}}{1-\alpha}+\kappa_0^*\right\}+\lambda\chi^*\mathcal{K}A'\right\}
\end{aligned}
$$

Next, for the LHS of the equation, we have that

$$
\mathrm{LHS}\equiv\left(1-\alpha\right)h_0=\left[(1-\alpha)A+\alpha h'B\Lambda-h'B\right]\mathcal{K}'\tilde{\mu},
$$

and, substituting the last h using equation (A.3), it follows that

$$\begin{aligned}
\mathrm{LHS} =& \ \left[(1-\alpha)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\left[\mathrm{I_m}-\mathcal{K}'\mathrm{S}\mathcal{K}\mathrm{B}'\left(\mathrm{B}\Omega\mathrm{B}'\right)^{-1}\mathrm{B}\right]\mathcal{K}'\tilde{\mu} \\
=& \ \left[(1-\alpha)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\mathcal{K}'S\left[S^{-1}-\mathcal{K}B'\left(B\Omega B'\right)^{-1}B\mathcal{K}'\right]\tilde{\mu} \\
=& \ \left[(1-\alpha)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\mathcal{K}'S \\
& \times\left\{\Omega_\mu^{-1}-\lambda\left[(1-\alpha^*)\,\mathcal{K}\left(\kappa_1^*A'-B'h\right)\left(\kappa_1^*A-h'B\right)\mathcal{K}'+\gamma^*\mathcal{K}A'A\mathcal{K}'\right]\right\}\tilde{\mu},
\end{aligned}$$

where the last equality uses the definition of S. Putting these results together, we have that

$$\begin{aligned}
\mathrm{LHS}-\mathrm{RHS} =& \ \left[(1-\alpha)\left(A-\mathrm{T}B\right)+\alpha h'\left(B\Lambda-\mathrm{H}B\right)\right]\mathcal{K}'S \\
& \times\Big\{\left\{\Omega_\mu^{-1}-\lambda\left[(1-\alpha^*)\,\mathcal{K}\left(\kappa_1^*A'-B'h\right)\left(\kappa_1^*A-h'B\right)\mathcal{K}'+\gamma^*\mathcal{K}A'A\mathcal{K}'\right]\right\}\tilde{\mu} \\
& \quad +\lambda(1-\alpha^*)\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\tilde{\mu}+\lambda\left(1-\alpha^*\right)\kappa_0^*\mathcal{K}\left(A'-B'h\right)-\lambda\chi^*\mathcal{K}A'\Big\},
\end{aligned}$$

where we used the fact that $\mathcal{K}\left(A'-B'h\right)h'B\left(\Lambda-\mathrm{I_m}\right)\mathcal{K}'=0$. Thus, a sufficient condition for $\mathrm{LHS}-\mathrm{RHS}=0$ is

$$\begin{aligned}
&\left\{\Omega_\mu^{-1}-\lambda\left[(1-\alpha^*)\,\mathcal{K}\left(\kappa_1^*A'-B'h\right)\left(\kappa_1^*A-h'B\right)\mathcal{K}'+\gamma^*\mathcal{K}A'A\mathcal{K}'\right]\right\}\tilde{\mu} \\
&\quad +\lambda(1-\alpha^*)\mathcal{K}\left(A'-B'h\right)\left(A-h'B\right)\mathcal{K}'\tilde{\mu}+\lambda\left(1-\alpha^*\right)\kappa_0^*\mathcal{K}\left(A'-B'h\right)-\lambda\chi^*\mathcal{K}A' = \ 0.
\end{aligned}$$

Notice that, using the definitions from Lemma B.4, this equation can be rewritten as

$$\left\{\Omega_\mu^{-1}+v_1v_1'+v_2v_2'\right\}\tilde{\mu}=-\lambda\left(1-\alpha^*\right)\kappa_0^*\mathcal{K}\left(A'-B'h\right)+\lambda\chi^*\mathcal{K}A'.$$

It follows that

$$\tilde{\mu}=\left\{\Omega_\mu+\Omega_\mu V\Omega_\mu\right\}\left\{-\lambda\left(1-\alpha^*\right)\kappa_0^*\mathcal{K}\left(A'-B'h\right)+\lambda\chi^*\mathcal{K}A'\right\},$$

and, therefore,

$$\begin{aligned}
h_0 &= (1-\alpha)^{-1}\left[(1-\alpha)\,A+\alpha h'B\Lambda-h'B\right]\mathcal{K}'\tilde{\mu} \\
&= \left(A-h'B\right)\mathcal{K}'\tilde{\mu} \\
&= \left(A-h'B\right)\mathcal{K}'\left\{\Omega_\mu+\Omega_\mu V\Omega_\mu\right\}\left\{-\lambda\left(1-\alpha^*\right)\kappa_0^*\mathcal{K}\left(A'-B'h\right)+\lambda\chi^*\mathcal{K}A'\right\} \\
&= -\lambda\left(1-\alpha^*\right)\kappa_0^*\left\{\left(A-h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}\left(A'-B'h\right)+\frac{c_{11}s_2^2+c_{22}s_1^2-\left(2+c_{12}+c_{21}\right)s_1s_2}{\left(1+c_{12}\right)\left(1+c_{21}\right)-c_{11}c_{22}}\right\} \\
&+\lambda\chi^*\left\{\left(A-h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A'+\frac{c_{11}s_2z_2+c_{22}s_1z_1-\left(1+c_{12}\right)s_1z_2-\left(1+c_{21}\right)s_2z_1}{\left(1+c_{12}\right)\left(1+c_{21}\right)-c_{11}c_{22}}\right\},
\end{aligned}$$

with

$$z_i\equiv A\mathcal{K}'\Omega_\mu v_i, \quad \text{for } i\in\{1,2\}.$$

Notice that we have the following limits

$$\lim_{T\to\infty} z_1=-\tau_\mu\frac{\gamma\lambda}{\omega}\mathbb{V}\left(\xi_t\right), \quad \text{and} \quad \lim_{T\to\infty} z_2=\tau_\mu\mathbb{COV}\left(\omega\xi_t-K_t,\xi_t\right).$$

Therefore, using $\chi^* = \chi + \varphi$, and $(1-\alpha^*)\,\kappa_0^* = \varphi$, we obtain the bias as a function of primitive parameters,

$$\mathcal{B} = \frac{\chi\lambda\tau_\mu \mathbb{V}(\xi_t)(1-\mathcal{S}) + \nu_4}{1 - \gamma\lambda\tau_\mu \mathbb{V}(\xi_t) + \nu_5},$$

with

$$\begin{aligned}
\nu_4 &\equiv \lambda\varphi\tau_\mu\left(\mathbb{V}(\xi_t)(1-\mathcal{S}) - \mathbb{V}(\xi_t - K_t)\right) \\
&\quad - \lambda^2\tau_\mu^2\left(\phi(\chi-\varphi) + 2\gamma\varphi\right)\left(\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2 - \mathbb{V}(\xi_t)\mathbb{V}(\xi_t - K_t)\right) \\
\nu_5 &\equiv \lambda\tau_\mu\mathbb{V}(\xi_t)(2\phi\mathcal{S} - \gamma) + \lambda^2\tau_\mu^2\phi^2\left(\mathbb{V}(\xi_t)^2(1-\mathcal{S})^2 - \mathbb{V}(\xi_t)\mathbb{V}(\xi_t - K_t)\right),
\end{aligned}$$

which completes the proof of part 3 of the proposition. □

B.2 Multiple aggregate shocks

Consider the same setup described in Section 3, but suppose that the common fundamental, ξ_t, is now driven by a $Z \times 1$ vector of shocks, η_t according to the following stochastic process:

$$\xi_t = a(L)\eta_t, \quad \text{with} \quad \eta_t \sim \mathcal{N}(0, \Sigma_\eta),$$

where $a(L)$ is a polynomial in the lag operator L. In the objective environment, η_t is normally distributed with mean zero: $\mu_t = 0$. Subjectively, agents believe that η_t is drawn from a Gaussian distribution with the same covariance matrix, Σ_η, but there is uncertainty about its prior mean, denoted by the $Z \times 1$ vector μ_t. Ambiguity about ξ_t is then captured by the perception that

$$\eta_t \sim \mathcal{N}(\mu_t, \Sigma_\eta), \quad \text{and} \quad \mu_t \sim \mathcal{N}(0, \Sigma_\mu).$$

In Section 3, the degree of ambiguity is captured by the σ_μ^2. Here, the covariance matrix Σ_μ plays this role. Without loss of generality, we assume that Σ_η and Σ_μ are diagonal matrices, that is $\Sigma_\eta = \text{diag}(\sigma_{\eta,1}^2, \ldots, \sigma_{\eta,Z}^2)$ and $\Sigma_\mu = \text{diag}(\sigma_{\mu,1}^2, \ldots, \sigma_{\mu,Z}^2)$.

Auxiliary forecasting problem Consider the following pure forecasting problem, which we later link back to the economy with ambiguity.

Definition B.1. *The $(w, \alpha, \{r_i\}_{i=1}^Z)$-modified signal process is given by*

$$\begin{aligned}
\tilde{\xi}_t &= a(L)\,\text{diag}(1+r_1, \ldots, 1+r_q)\tilde{\eta}_t, \quad &&\textit{with} \quad \tilde{\eta}_t \sim \mathcal{N}(0, (1+w)\Sigma_\eta + w\Sigma_\mu), \\
\tilde{x}_{it} &= m(L)\tilde{\eta}_t + n(L)\tilde{\epsilon}_{it}, &&\textit{with} \quad \tilde{\epsilon}_{it} \sim \mathcal{N}(0, (1-\alpha)^{-1}\Sigma),
\end{aligned}$$

where w is a non-negative scalar and α is the degree of complementarity. Let the optimal Bayesian forecast be given by

$$\tilde{\mathbb{E}}_{it}[\tilde{\xi}_t] = p(L; w, \alpha, \{r_i\}_{i=1}^Z)\tilde{x}_{it}.$$

This modified signal process is analogous to the baseline. The adjustment w to the volatility of η_t is the counterpart to $\tilde{w} = w\tau_\mu^{-1}$ in the univariate baseline, that is $\Sigma_\eta + w\Sigma_\mu$ is the counterpart of $(1+w)\sigma_\eta^2 = (1+\tilde{w}\tau_\mu)\sigma_\eta^2 = \sigma_\eta^2 + \tilde{w}\sigma_\mu^2$. Further, the amplification factor, $(1+r)$ in the univariate case, has now been incorporated into this modified signal process since in the multivariate case each shock requires a potentially different adjustment, before being put together into a modified fundamental. So, $p(L; w, \alpha, \{r_i\}_{i=1}^Z)$ here is the counterpart of $(1+r)p(L; w, \alpha)$ in the univariate case. To proceed we need the additional following definitions.

Definition B.2. *Define the μ-modified fundamental and (unbiased) aggregate action as*

$$\xi_t^\mu = a(L)\mu_t, \quad \textit{and} \quad K_t^\mu = p(L; w, \alpha, \{r_i\}_{i=1}^Z)\mu_t,$$

and the μ-modified aggregate sensitivity to signals as

$$\mathcal{S}^\mu \equiv 1 - \frac{\mathbb{COV}\left(\xi_t^\mu - K_t^\mu, \xi_t^\mu\right)}{\mathbb{V}\left(\xi_t^\mu\right)}.$$

We can then prove the following proposition.

Proposition B.2. *The linear strategy in equilibrium takes the following form*

$$g(x_i^t) = p(L; w, \alpha, \{r_i\}_{i=1}^Z)x_{it} + \mathcal{B}.$$

1. *The polynomial matrix $p(L; w, \alpha, \{r_i\}_{i=1}^Z)$ is the Bayesian forecasting rule with the $(w, \alpha, \{r_i\}_{i=1}^Z)$-modified signal process and w satisfies*

$$w = \frac{1}{1 - \lambda(1-\alpha)\left(\mathbb{V}\left(\xi_t^\mu - K_t^\mu\right) + \frac{\lambda\gamma\mathbb{V}(\xi_t^\mu)^2(1-\mathcal{S}^\mu)^2}{1-\lambda\gamma\mathbb{V}(\xi_t^\mu)}\right)};$$

2. *For all $i \in \{1, \ldots, Z\}$, the additional amplification, r_i, satisfies*

$$r_i = \gamma\frac{\lambda\mathbb{V}\left(\xi_t^\mu\right)}{1 - \lambda\gamma\mathbb{V}\left(\xi_t^\mu\right)}\frac{w\tau_{\mu,i}}{1 + w\tau_{\mu,i}}(1 - \mathcal{S}^\mu);$$

3. *The level of bias, $\mathcal{B}$, satisfies*

$$\mathcal{B} = \chi\frac{\lambda\mathbb{V}\left(\xi_t^\mu\right)}{1 - \lambda\gamma\mathbb{V}\left(\xi_t^\mu\right)}\left(1 - \mathcal{S}^\mu\right).$$

Proof of Proposition B.2. The truncated version of the problem is analogous to the case with one common shock, with the following adjustments: (1) the size of the vector of aggregate common shocks must be set to $u \equiv Z(T+1)$; (2) the size of the vector of all shocks becomes $m \equiv (Z+N)(T+1)$; (3) instead of $\Omega_\eta = \mathrm{I}_u\,\sigma_\eta^2$ and $\Omega_\mu = \mathrm{I}_u\,\sigma_\mu^2$, we now have $\Omega_\eta = \mathrm{I}_{T+1} \otimes \Sigma_\eta$ and $\Omega_\mu = \mathrm{I}_{T+1} \otimes \Sigma_\mu$. These modifications do not affect in any way the results in Lemmas A.1, A.2, and A.3. However, Lemma A.4 relies on the fact that $\Omega_\eta = \mathrm{I}_u\,\sigma_\eta^2$ and $\Omega_\mu = \mathrm{I}_u\,\sigma_\mu^2$. The following lemma provides the relevant analogous result.

Lemma B.5. *Define*

$$\bar{\Delta} \equiv \Gamma\Omega + \hat{w}\mathcal{K}'\Omega_\mu\mathcal{K},$$

with the scalar $\hat{w}$ given by

$$\hat{w} = \frac{1}{1 = 2\lambda\left(1-\alpha\right)\left(A - h'B\right)\mathcal{K}'\left(\Omega_\mu + \frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\mathcal{K}\left(A' - B'h\right)}.$$

Also, let the diagonal matrix $\hat{R}$ be given by

$$\hat{R} \equiv \mathrm{I}_{T+1} \otimes \mathrm{diag}\left(\hat{r}_1, \ldots, \hat{r}_Z\right),$$

with the scalars $\hat{r}_i$, for $i \in \{1, \ldots, Z\}$, given by

$$\hat{r}_i \equiv \frac{\hat{w}\tau_{\mu,i}}{1 + \hat{w}\tau_{\mu,i}}\left(\frac{\lambda\gamma}{1 - \lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\left(A - h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A', \quad \textit{with} \quad \tau_{\mu,i} \equiv \frac{\sigma^2_{\mu,i}}{\sigma^2_{\eta,i}}.$$

Then, the equilibrium h satisfies

$$h' = A\left(\mathrm{I}_m + \hat{R}\right)\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1}.$$

Proof. From the definition of $\tilde{\Omega}_\mu$ and Δ in Lemma A.3, we have that

$$\tilde{\Omega}_\mu \equiv \left(\Omega_\mu^{-1} - \lambda\gamma\mathcal{K}A'A\mathcal{K}'\right)^{-1} = \Omega_\mu + \frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1 - \lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}.$$

Thus, from the definition of Δ in Lemma A.3 and defining

$$\bar{\Delta} \equiv \Gamma\Omega + \hat{w}\mathcal{K}'\Omega_\mu\mathcal{K},$$

we have that

$$\Delta \equiv \Gamma\Omega + \hat{w}\mathcal{K}'\tilde{\Omega}_\mu\mathcal{K} = \bar{\Delta} + \hat{w}\mathcal{K}'\left(\frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1 - \lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\mathcal{K} = \bar{\Delta} + s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K},$$

with $s \equiv \lambda\gamma/(1 - \lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A')$. Hence, it follows from the result in Lemma A.3 that

$$h' = A\left(\bar{\Delta} + s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'\left[B\left(\bar{\Delta} + s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'\right]^{-1},$$

and, therefore,

$$h'\left[B\left(\bar{\Delta} + s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'\right] = A\left(\bar{\Delta} + s\hat{w}\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}\right)B'.$$

Rearranging, we get

$$h'B\bar{\Delta}B' + s\hat{w}h'B\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}B' = A\bar{\Delta}B' + s\hat{w}A\mathcal{K}'\left(\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\right)\mathcal{K}B',$$

and right-multiplying both sides by $\left(B\bar{\Delta}B'\right)^{-1}$ yields

$$\begin{aligned} h' &= A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + s\hat{w}\left(A - h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1} \\ &= A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \hat{z}\hat{w}A\mathcal{K}'\Omega_\mu\mathcal{K}B'\left(B\bar{\Delta}B'\right)^{-1}. \end{aligned}$$

Next, notice that

$$\mathcal{K}'\Omega_\mu\mathcal{K} = \mathcal{K}'\Omega_\mu\Omega_\eta^{-1}\mathcal{K}\Omega,$$

and

$$\bar{\Delta} = \Big(\mathcal{K}'(\mathrm{I}_u + \hat{w}\Omega_\mu\Omega_\eta^{-1})\mathcal{K} + (1-\alpha)^{-1}\left(\mathrm{I}_m - \mathcal{K}'\mathcal{K}\right)\Big)\,\Omega,$$

so we have

$$\mathcal{K}'\Omega_\mu\mathcal{K} = \mathcal{K}'\left(\Omega_\mu(\Omega_\eta + \hat{w}\Omega_\mu)^{-1}\right)\mathcal{K}\bar{\Delta}.$$

Thus, it follows that

$$h' = A\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} + \hat{z}\hat{w}A\mathcal{K}'\left(\Omega_\mu(\Omega_\eta + \hat{w}\Omega_\mu)^{-1}\right)\mathcal{K}\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1},$$

with the scalar $\hat{z}$ given by

$$\hat{z} \equiv \left(\frac{\lambda\gamma}{1 - \lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\left(A - h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A'.$$

Further, we can write

$$h' = A\left(\mathrm{I}_m + \hat{R}\right)\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1},$$

with

$$\begin{aligned} \hat{R} &= \mathcal{K}'\left(\hat{z}\hat{w}\Omega_\mu(\Omega_\eta + \hat{w}\Omega_\mu)^{-1}\right)\mathcal{K} \\ &= \mathcal{K}'\left(\hat{z}\hat{w}(\mathrm{I}_{T+1}\otimes\Sigma_\mu)((\mathrm{I}_{T+1}\otimes\Sigma_\eta) + \hat{w}(\mathrm{I}_{T+1}\otimes\Sigma_\mu))^{-1}\right)\mathcal{K} \\ &= \mathcal{K}'\left(\mathrm{I}_{T+1}\otimes(\hat{z}\hat{w}\Sigma_\mu(\Sigma_\eta + \hat{w}\Sigma_\mu)^{-1})\right)\mathcal{K} \\ &= \mathcal{K}'\left(\mathrm{I}_{T+1}\otimes\mathrm{diag}\left(\hat{z}\hat{w}\sigma^2_{\mu,1}(\sigma^2_{\eta,1} + \hat{w}\sigma^2_{\mu,1})^{-1}, \ldots, \hat{w}\sigma^2_{\mu,Z}(\sigma^2_{\eta,Z} + \hat{w}\sigma^2_{\mu,Z})^{-1}\right)\right)\mathcal{K} \\ &= \mathcal{K}'\left(\mathrm{I}_{T+1}\otimes\mathrm{diag}\left(\hat{r}_1, \ldots, \hat{r}_Z\right)\right)\mathcal{K}, \end{aligned}$$

which concludes the proof. □

Parts 1 and 2 of Proposition B.2. Given the result in Lemma B.5, we are left with taking the limit as $T \to \infty$ of the truncated problem. In particular, we have that

$$\lim_{T\to\infty} A\left(\mathrm{I}_m + \hat{R}\right)\bar{\Delta}B'\left(B\bar{\Delta}B'\right)^{-1} = p\left(L; w, \alpha, \{r_i\}_{i=1}^Z\right), \quad \lim_{T\to\infty} A\mathcal{K}'\Omega_\mu\mathcal{K}A' = \mathbb{V}(\xi_t^\mu),$$

$$\lim_{T\to\infty}\left(A - h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}(A' - Bh') = \mathbb{V}(\xi_t^\mu - K_t^\mu), \quad \lim_{T\to\infty}\left(A - h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A' = \mathbb{COV}\left(\xi_t^\mu - K_t^\mu, \xi_t^\mu\right),$$

$$\lim_{T\to\infty}\frac{\left(A - h'B\right)\mathcal{K}'\Omega_\mu\mathcal{K}A'}{A\mathcal{K}'\Omega_\mu\mathcal{K}A'} = 1 - \mathcal{S}^\mu.$$

Let $w \equiv \lim_{T\to\infty} \hat{w}$, and $r_i \equiv \lim_{T\to\infty} \hat{r}_i$, for $i \in \{1, \ldots, Z\}$. Then, we can show that

$$\begin{aligned} r_i &= \lim_{T\to\infty} \frac{\hat{w}\tau_{\mu,i}}{1+\hat{w}\tau_{\mu,i}} \frac{\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'} \frac{(A-h'B)\,\mathcal{K}'\Omega_\mu\mathcal{K}A'}{A\mathcal{K}'\Omega_\mu\mathcal{K}A'} \\ &= \frac{w\tau_{\mu,i}}{1+w\tau_{\mu,i}} \frac{\lambda\gamma\mathbb{V}\left(\xi_t^\mu\right)}{1-\lambda\gamma\mathbb{V}\left(\xi_t^\mu\right)} \left(1-\mathcal{S}^\mu\right), \end{aligned}$$

and

$$\begin{aligned} w &= \lim_{T\to\infty} \frac{1}{1-\lambda\,(1-\alpha)\,(A-h'B)\,\mathcal{K}'\left(\Omega_\mu + \frac{\lambda\gamma\Omega_\mu\mathcal{K}A'A\mathcal{K}'\Omega_\mu}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)\mathcal{K}\,(A'-B'h)} \\ &= \lim_{T\to\infty} \frac{1}{1-\lambda\,(1-\alpha)\left((A-h'B)\,\mathcal{K}'\Omega_\mu\mathcal{K}\,(A'-B'h) + \frac{\lambda\gamma((A-h'B)\mathcal{K}'\Omega_\mu\mathcal{K}A')(A\mathcal{K}'\Omega_\mu\mathcal{K}(A'-B'h))}{1-\lambda\gamma A\mathcal{K}'\Omega_\mu\mathcal{K}A'}\right)} \\ &= \frac{1}{1-\lambda(1-\alpha)\left(\mathbb{V}\left(\xi_t^\mu - K_t^\mu\right) + \frac{\lambda\gamma\mathbb{V}(\xi_t^\mu)^2(1-\mathcal{S}^\mu)^2}{1-\lambda\gamma\mathbb{V}(\xi_t^\mu)}\right)}. \end{aligned}$$

Part 3 of Proposition B.2. All the steps used in the proof of part 3 of Proposition 3.3 hold without change except for the last step. From those derivations we have that

$$h_0 = \chi\lambda\,(A-h'B)\,\mathcal{K}'\Omega_\mu\mathcal{K}A'\left(\mathrm{I_u} + \frac{\lambda\gamma\mathrm{A}\mathcal{K}'\Omega_\mu\mathcal{K}\mathrm{A}'}{1-\lambda\gamma\mathrm{A}\mathcal{K}'\Omega_\mu\mathcal{K}\mathrm{A}'}\right)$$

Taking the limit we get

$$\mathcal{B} = \lim_{T\to\infty} h_0 = \chi\lambda\tau_\mu\mathbb{COV}(\xi_t^\mu - K_t^\mu, \xi_t^\mu)\left(1+\frac{\lambda\gamma\mathbb{V}(\xi_t^\mu)}{1-\lambda\gamma\mathbb{V}(\xi_t^\mu)}\right) = \frac{\chi\lambda\mathbb{V}\left(\xi_t^\mu\right)}{1-\lambda\gamma\mathbb{V}\left(\xi_t^\mu\right)}\left(1-\mathcal{S}^\mu\right),$$

which completes the proof of part 3 of the proposition.

□

C Proofs of Other Results

Proof of Proposition 2.1. Following the same arguments used in the proof of Proposition 3.2, the optimal linear strategy, $g(x_i) \equiv s^* x_i + \mathcal{B}$, solves the following fixed point problem

$$s^* x_i + \mathcal{B} = \int_\mu \mathbb{E}^\mu [\xi | x_i] \, \hat{p}(\mu | x_i) \, \mathrm{d}\mu = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\epsilon^2} x_i + \frac{\sigma_\epsilon^2}{\sigma_\xi^2 + \sigma_\epsilon^2} \int_\mu \mu \, \hat{p}(\mu | x_i) \, \mathrm{d}\mu,$$

with

$$\begin{aligned}
\hat{p}(\mu | x_i) &\propto \exp\left(\lambda \mathbb{E}^\mu \left[(s^* x_i + \mathcal{B} - \xi)^2 - \chi \xi\right]\right) p(x_i | \mu) \, p(\mu) \\
&\propto \exp\left(\lambda (1 - s^*)^2 \mu^2 + 2\lambda (s^* - 1) \mathcal{B} \mu - \chi \mu - \frac{(x_i - \mu)^2}{2\left(\sigma_\xi^2 + \sigma_\epsilon^2\right)} - \frac{1}{2\sigma_\mu^2} \mu^2\right).
\end{aligned}$$

Mapping it into the kernel of a normal distribution yields

$$\mu \sim \mathcal{N}\left(\frac{\frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} x_i + 2\lambda (s^* - 1) \mathcal{B} - \lambda \chi}{\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} - 2\lambda (1 - s^*)^2}, \frac{1}{\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} - 2\lambda (1 - s^*)^2}\right),$$

which implies that

$$\int_\mu \mu \, \hat{p}(\mu | x_i) \, \mathrm{d}\mu = \frac{\frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} x_i + 2\lambda (s^* - 1) \mathcal{B} - \lambda \chi}{\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} - 2\lambda (1 - s^*)^2}.$$

Matching coefficients leads to the following conditions

$$s^* = \frac{\sigma_\xi^2}{\sigma_\xi^2 + \sigma_\epsilon^2} + \frac{\sigma_\epsilon^2}{\sigma_\xi^2 + \sigma_\epsilon^2} \frac{\frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2}}{\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} - 2\lambda (1 - s^*)^2},$$

and

$$\mathcal{B} = \frac{\sigma_\epsilon^2}{\sigma_\xi^2 + \sigma_\epsilon^2} \frac{2\lambda (s^* - 1) \mathcal{B} - \lambda \chi}{\frac{1}{\sigma_\mu^2} + \frac{1}{\sigma_\xi^2 + \sigma_\epsilon^2} - 2\lambda (1 - s^*)^2}.$$

Solving for s^* and $\mathcal{B}$ leads to the expressions stated in the proposition. $\square$

Proof of Corollary 3.1. Aggregating the individual best response in equation (3.6) leads to

$$K_t = (1 - \alpha) \overline{\mathcal{F}}_t^1 [\xi_t] + \alpha \overline{\mathcal{F}}_t^1 [K_t].$$

Iterating forward using the definitions of subjective higher-order expectations, it follows that

$$\begin{aligned}
K_t &= (1-\alpha)\,\overline{\mathcal{F}}_t^{1}\left[\xi_t\right] + \alpha\,(1-\alpha)\,\overline{\mathcal{F}}_t^{2}\left[\xi_t\right] + \alpha^2\overline{\mathcal{F}}_t^{2}\left[K_t\right] \\
&= \cdots \\
&= (1-\alpha)\sum_{j=0}^{N}\alpha^j\overline{\mathcal{F}}_t^{j+1}\left[\xi_t\right] + \alpha^{N+1}\overline{\mathcal{F}}_t^{N+1}\left[K_t\right] \\
&= \cdots \\
&= (1-\alpha)\sum_{j=0}^{\infty}\alpha^j\overline{\mathcal{F}}_t^{j+1}\left[\xi_t\right],
\end{aligned}$$

which completes the proof. □

Proof of Corollary 3.2. This result follows directly from the fact that $p\,(L; w, \alpha)$ permits a finite state representation. □

Proof of Proposition 3.4. Applying Proposition 3.3, we obtain

$$\mathcal{F}_i\left[\xi\right] = \varsigma x_i - (1-\varsigma)\,\lambda\chi\sigma_\mu^2, \quad \text{with} \quad \varsigma \equiv \frac{(1+w)\,\sigma_\xi^2}{(1+w)\,\sigma_\xi^2 + \sigma_\epsilon^2}.$$

Aggregating over i, it follows that

$$\overline{\mathcal{F}}\left[\xi\right] = \varsigma\xi - (1-\varsigma)\,\lambda\chi\sigma_\mu^2.$$

Applying the operator $\mathcal{F}_i$ to both sides and aggregating again yields,

$$\overline{\mathcal{F}}^2\left[\xi\right] = \varsigma^2\xi - (1-\varsigma)\,(1+\varsigma)\,\lambda\chi\sigma_\mu^2.$$

Iterating forward, it follows that

$$\overline{\mathcal{F}}^m\left[\xi\right] = \varsigma^m\xi - (1-\varsigma)\sum_{k=0}^{m-1}\varsigma^k\lambda\chi\sigma_\mu^2 = \kappa_m\xi + \beta_m,$$

with

$$\kappa_m \equiv \varsigma^m, \quad \text{and} \quad \beta_m \equiv -\,(1-\varsigma)\sum_{k=0}^{m-1}\varsigma^k\lambda\chi\sigma_\mu^2.$$

Therefore, we have that

$$\beta_m = \beta_{m-1} - (1-\varsigma)\,\kappa_{m-1}\lambda\chi\sigma_\mu^2 = \beta_{m-1} + (\kappa_m - \kappa_{m-1})\,\lambda\chi\sigma_\mu^2,$$

which completes the proof of Part 1. Moreover, combing equation (3.7) with the fact that $\overline{\mathcal{F}}^m\left[\xi\right] = \kappa_m\xi + \beta_m$ leads to

$$K = (1-\alpha)\sum_{m=0}^{\infty}\alpha^m\kappa_m\,\xi + (1-\alpha)\sum_{m=0}^{\infty}\alpha^m\beta_m,$$

which completes the proof of Part 3.

To establish Part 2 notice that, from Proposition 3.3, we have

$$w = \left[\frac{1}{\tau_\mu} - \lambda\left(1-\alpha\right)\mathbb{V}\left(\xi - K\right)\right]^{-1},$$

which, differentiating with respect to α, yields

$$\frac{\mathrm{d}w}{\mathrm{d}\alpha} = \lambda w^2\left[\left(1-\alpha\right)\frac{\mathrm{d}\mathbb{V}\left(\xi-K\right)}{\mathrm{d}\alpha} - \mathbb{V}\left(\xi-K\right)\right].$$

Since

$$\mathbb{V}\left(\xi - K\right) = \left(\frac{\sigma_\epsilon^2}{\left(1+w\right)\left(1-\alpha\right)\sigma_\xi^2+\sigma_\epsilon^2}\right)^2\sigma_\xi^2,$$

it follows that

$$\frac{\mathrm{d}\mathbb{V}\left(\xi-K\right)}{\alpha} = 2\left(\frac{\sigma_\xi^2}{\left(1+w\right)\left(1-\alpha\right)\sigma_\xi^2+\sigma_\epsilon^2}\right)\mathbb{V}\left(\xi-K\right)\left(w - \left(1-\alpha\right)\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right),$$

and, therefore,

$$\begin{aligned}\frac{\mathrm{d}w}{\mathrm{d}\alpha} &= \lambda w^2\mathbb{V}\left(\xi-K\right)\left[2\left(1-\alpha\right)\left(\frac{\sigma_\xi^2}{\left(1+w\right)\left(1-\alpha\right)\sigma_\xi^2+\sigma_\epsilon^2}\right)\left(w-\left(1-\alpha\right)\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right)-1\right]\\ &= \frac{\lambda w^2\mathbb{V}\left(\xi-K\right)\left(\left(w-1\right)\left(1-\alpha\right)\sigma_\xi^2-\sigma_\epsilon^2\right)}{\left[1+w+2\lambda\left(1-\alpha\right)w^2\mathbb{V}\left(\xi-K\right)\right]\left(1-\alpha\right)\sigma_\xi^2+\sigma_\epsilon^2}.\end{aligned}$$

Then, since, in the limit as α increases to 1, we have that $w \to \tau_\mu$, and $\mathbb{V}(\xi - K) \to \mathbb{V}(\xi)$, it follows that

$$\lim_{\alpha\to 1^-}\frac{\mathrm{d}w}{\mathrm{d}\alpha} = -\lambda\tau_\mu^2\mathbb{V}(\xi) < 0.$$

On the other hand, notice that

$$\mathrm{sgn}\left[\lim_{\alpha\to 0^+}\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right] = \mathrm{sgn}\left[\left(w-1\right)\sigma_\xi^2 - \sigma_\epsilon^2\right],$$

so that, since $w \geq \tau_\mu$, we have that

$$\tau_\mu > \frac{\sigma_\xi^2+\sigma_\epsilon^2}{\sigma_\xi^2} \quad\Rightarrow\quad \lim_{\alpha\to 0^+}\frac{\mathrm{d}w}{\mathrm{d}\alpha} > 0.$$

Hence, w is non-monotonic in α if τ_μ is large enough. □

Proof of Proposition 3.5. It follows from Proposition 3.3 that, when $\gamma = 0$,

$$\mathcal{B} = \chi\lambda\tau_\mu \mathbb{V}\left(\xi_t\right)\left(1 - \mathcal{S}\right).$$

Therefore, to prove that $|\mathcal{B}|$ is increasing in α, it is sufficient to prove that the sensitivity $\mathcal{S}$ is decreasing in α. Since, by definition

$$\mathcal{S} = \frac{\mathbb{COV}\left(K_t, \xi_t\right)}{\mathbb{V}\left(\xi_t\right)},$$

with $\mathbb{V}\left(\xi_t\right)$ independent of α, it is sufficient to show that

$$\frac{\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)}{\mathrm{d}\alpha} < 0.$$

Following the notation of the truncated economy introduced in the proof of Proposition 3.3, we have that

$$\mathbb{COV}\left(K_t, \xi_t\right) = h' B \Lambda \Omega A',$$

with h denoting the optimal forecasting rule

$$h = A\bar{\Omega}B'\left(B\bar{\Omega}B'\right)^{-1}, \quad \text{with} \quad \bar{\Omega} = \left(1 + w\right)\Lambda\Omega + \left(1 - \alpha\right)^{-1}\left(\mathrm{I}_m - \Lambda\right)\Omega.$$

Since Ω is diagonal, we can rewrite h as

$$h = A\hat{\Omega}B'\left(B\hat{\Omega}B'\right)^{-1}, \quad \text{with} \quad \hat{\Omega} = \Lambda\Omega + m_\alpha\left(\mathrm{I}_m - \Lambda\right)\Omega, \quad \textit{and} \quad m_\alpha \equiv \left[\left(1 - \alpha\right)\left(1 + w\right)\right]^{-1}.$$

It follows that

$$\begin{aligned}
\frac{\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)}{\mathrm{d}\alpha} &= A\Omega\Lambda B' \frac{\mathrm{d}\left(B\hat{\Omega}B'\right)^{-1}}{\mathrm{d}\alpha} B\Lambda\Omega A' \\
&= -A\Omega\Lambda B'\left(B\hat{\Omega}B'\right)^{-1} B\frac{\mathrm{d}\hat{\Omega}}{\mathrm{d}\alpha}B'\left(B\hat{\Omega}B'\right)^{-1} B\Lambda\Omega A' \\
&= -\left(z'\left(\mathrm{I}_m - \Lambda\right)\Omega\, z\right) m_\alpha^2 \left[\left(1 + w\right) - \left(1 - \alpha\right)\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right],
\end{aligned}$$

where z is a column vector,

$$z \equiv B'\left(B\hat{\Omega}B'\right)^{-1} B\Lambda\Omega A'.$$

Since $\left(\mathrm{I}_m - \Lambda\right)\Omega$ is positive semi-definite, it follows that

$$\operatorname{sgn}\left[\frac{\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)}{\mathrm{d}\alpha}\right] = -\operatorname{sgn}\left[\left(1 + w\right) - \left(1 - \alpha\right)\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right].$$

Further, notice that since $w \geq \tau_\mu$ and $\lim_{\alpha \to 1^-} w = \tau_\mu$, we have that the $\lim_{\alpha \to 1^-} \mathrm{d}w/\mathrm{d}\alpha$ is bounded and, therefore,

$$\lim_{\alpha \to 1^-} \frac{\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)}{\mathrm{d}\alpha} < 0.$$

Finally, for a contradiction, suppose there exists some $\alpha \in [0,1)$ such that $\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)/\mathrm{d}\alpha > 0$. It follows from the intermediate value theorem and the continuity of $\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)/\mathrm{d}\alpha$ that there must exist some $\alpha_\dagger$ such that

$$\left.\frac{\mathrm{d}\mathbb{COV}\left(K_t, \xi_t\right)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = 0 \quad \Rightarrow \quad \left.\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = \frac{1+w_\dagger}{1-\alpha_\dagger} > 0,$$

where $w_\dagger$ denotes w evaluated at $\alpha_\dagger$. With $\gamma = 0$, Proposition 3.3 implies that

$$w = \left[\frac{1}{\tau_\mu} - \lambda\left(1-\alpha\right)\mathbb{V}\left(\xi_t - K_t\right)\right]^{-1},$$

and it follows that

$$\frac{\mathrm{d}w}{\mathrm{d}\alpha} = -\lambda w^2 \left[\mathbb{V}\left(\xi_t - K_t\right) - \left(1-\alpha\right)\frac{\mathrm{d}\mathbb{V}\left(\xi_t - K_t\right)}{\mathrm{d}\alpha}\right].$$

Using the fact that, similarly to $\mathbb{COV}\left(\xi_t, K_t\right)$, $\mathbb{V}\left(\xi_t - K_t\right)$ depends on α only through m_α, we have that

$$\left.\frac{\mathrm{d}\mathbb{V}\left(\xi_t - K_t\right)}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = \left.\frac{\mathrm{d}\mathbb{V}\left(\xi_t - K_t\right)}{\mathrm{d}m_\alpha}\frac{\mathrm{d}m_\alpha}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = \left.\frac{\mathrm{d}\mathbb{V}\left(\xi_t - K_t\right)}{\mathrm{d}m_\alpha} m_\alpha^2 \left[\left(1+w\right) - \left(1-\alpha\right)\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right]\right|_{\alpha=\alpha_\dagger} = 0,$$

and, therefore,

$$\left.\frac{\mathrm{d}w}{\mathrm{d}\alpha}\right|_{\alpha=\alpha_\dagger} = -\lambda w^2 \mathbb{V}\left(\xi_t - K_t\right) < 0,$$

which yields the desired contradiction. □

Proof of Lemma 4.1. We start by characterizing the zero-inflation steady state. From the budget constraint of household i, we have that

$$C_{i,g,t+1} = \frac{Y_g - C_{i,g,t}}{1+\pi_{t+1}}.$$

Substituting $C_{i,g,t+1}$ into the utility function $U\left(C_{i,g,t}, C_{i,g,t+1}\right)$ yields

$$U\left(C_{i,g,t}, \pi_{t+1}\right) = \frac{C_{i,g,t}^{1-\nu}}{1-\nu} + \beta\frac{\left(\frac{Y_g - C_{i,g,t}}{1+\pi_{t+1}}\right)^{1-\nu}}{1-\nu}.$$

The Euler equation in the zero-inflation steady state implies that

$$\bar{C}_g^{-\nu} - \beta\left(Y_g - \bar{C}_g\right)^{-\nu} = 0. \tag{C.1}$$

Let $c_{i,g,t}$ be the log-deviation from the zero-inflation steady state, that is

$$c_{i,g,t} \equiv \log C_{i,g,t} - \log \bar{C}_g.$$

The quadratic approximation of $U(C_{i,g,t}, \pi_{t+1})$ around the zero-inflation steady state leads to

$$\begin{aligned}
U(C_{i,g,t}, \pi_{t+1}) \approx\ & \mathcal{Q}(\hat{c}_t, \pi_{t+1}) \\
& \equiv const - \bar{C}_g^{1-\nu}\left(\frac{Y-\bar{C}_g}{\bar{C}_g}\right)\pi_{t+1} + (1-\nu)\,\bar{C}_g^{1-\nu} c_{i,g,t}\pi_{t+1} \\
& + \frac{1}{2}(1-\nu)\,\bar{C}_g^{1-\nu}\left(\frac{Y_g-\bar{C}_g}{\bar{C}_g}\right)\pi_{t+1}^2 - \frac{1}{2}\nu\bar{C}_g^{1-\nu}\left[1+\left(\frac{Y_g-\bar{C}_g}{\bar{C}_g}\right)^{-1}\right]c_{i,g,t}^2 \\
& = const - \bar{C}_g^{1-\nu}\left(\frac{Y_g-\bar{C}_g}{\bar{C}_g}\right)\pi_{t+1} + \frac{1}{2}(1-\nu)\,\bar{C}_g^{1-\nu}\left(\frac{Y_g-\bar{C}_g}{\bar{C}_g} + \frac{1-\nu}{\nu}\frac{Y_g-\bar{C}_g}{Y_g}\right)\pi_{t+1}^2 \\
& - \frac{1}{2}\nu\bar{C}_g^{1-\nu}\frac{Y_g}{Y_g-\bar{C}_g}\left(c_{i,g,t} - \frac{1-\nu}{\nu}\frac{Y_g-\bar{C}_g}{Y_g}\pi_{t+1}\right)^2.
\end{aligned}$$

Given subjective beliefs $\mathcal{F}_{i,g,t}[\cdot]$, the optimal consumption must be proportional to the households subjective expectation about inflation:

$$\begin{aligned}
c_{i,g,t} =& \frac{1-\nu}{\nu}\frac{Y_g-\bar{C}_g}{Y_g}\mathcal{F}_{i,g,t}[\pi_{t+1}] \\
=& \frac{\beta^{1/\nu}}{1+\beta^{1/\nu}}\mathcal{F}_{i,g,t}[\pi_{t+1}],
\end{aligned}$$

where the last equality directly follows equation (C.1).

In the smooth model of ambiguity, similarly to the proof of Proposition 3.2, it can be shown that

$$c_{i,g,t} = \frac{1-\nu}{\nu}\frac{Y_g-\bar{C}_g}{Y_g}\int_{\mu^t}\mathbb{E}^{\mu^t}[\pi_{t+1}|\mathcal{I}_{i,g,t}]\,\hat{p}\left(\mu^t|\mathcal{I}_{i,g,t}\right)\mathrm{d}\mu^t,$$

where the distorted posterior $\hat{p}(\mu^t|\mathcal{I}_{i,g,t})$ is such that

$$\hat{p}\left(\mu^t|\mathcal{I}_{i,g,t}\right) \propto \exp\left(-\lambda\mathbb{E}^{\mu^t}[\mathcal{Q}(\hat{c}_t, \pi_{t+1})]\right).$$

Let the subjective belief of the household be such that

$$\mathcal{F}_{i,g,t}[\cdot] \equiv \int_{\mu^t}\mathbb{E}^{\mu^t}[\cdot|\mathcal{I}_{i,g,t}]\,\hat{p}\left(\mu^t|\mathcal{I}_{i,g,t}\right)\mathrm{d}\mu^t,$$

then, it follows that

$$c_{i,g,t} = \frac{\beta^{1/\nu}}{1+\beta^{1/\nu}}\mathcal{F}_{i,g,t}[\pi_{t+1}],$$

which yields equation (4.2). Substituting $c_{i,g,t}$ into $\mathcal{Q}(\hat{c}_t, \pi_{t+1})$, leads to equation (4.3) with

$$\chi_g \equiv Y_g^{1-\nu} \frac{\beta^{1/\nu}}{\left(1+\beta^{1/\nu}\right)^{1-\nu}}$$
$$\gamma_g \equiv \frac{1}{2} Y_g^{1-\nu} \frac{(\nu-1)\,\beta^{1/\nu}}{\left(1+\beta^{1/\nu}\right)^{1-\nu}} \frac{1+\nu\beta^{1/\nu}}{\nu\left(1+\beta^{1/\nu}\right)}$$
$$\delta_g \equiv \frac{1}{2} Y_g^{1-\nu} \frac{(1-\nu)^2\,\beta^{1/\nu}}{\nu\left(1+\beta^{1/\nu}\right)^{2-\nu}}.$$

Notice that δ_g, χ_g, and γ_g are all proportional to $Y_g^{\nu-1}$. Moreover, when $\nu > 1$, they are all positive and decreasing in Y_g. $\square$

The following lemma is used in the proof of the next propositions.

Lemma C.1 (Kalman filter for AR(1))**.** *Given a state equation*

$$\xi_t = \rho\xi_{t-1} + \nu_t, \quad \textit{with} \quad \nu_t \sim \mathcal{N}\left(0, \sigma_\nu^2\right),$$

and an observation equation

$$x_t = \xi_t + u_t, \quad \textit{with} \quad u_t \sim \mathcal{N}\left(0, \sigma_u^2\right),$$

the steady-state Kalman gain is given by

$$\kappa = \frac{1}{2\rho}\left(\rho - \frac{\sigma_u^2 + \sigma_\nu^2}{\rho\sigma_u^2} - \sqrt{\left(\rho - \frac{\sigma_u^2+\sigma_\nu^2}{\rho\sigma_u^2}\right)^2 + 4\frac{\sigma_\nu^2}{\sigma_u^2}}\right),$$

and the updating rule for the Bayesian forecast follows

$$\mathbb{E}_t\left[\xi_{t+1}\right] = \rho\,(1-\kappa)\,\mathbb{E}_{t-1}\left[\xi_t\right] + \rho\kappa x_t.$$

Proof of Proposition 4.1. Cosider Lemma C.1 with $\xi_t = \pi_t$, $\sigma_\nu^2 = \sigma_\eta^2$, and $\sigma_u^2 = \sigma_\varepsilon^2$, and define $\omega \equiv \rho\,(1-\kappa)$. Since every agent i in every group g has the same information structure with signals given by

$$x_{i,g,t} = \pi_t + \varepsilon_{i,g,t}, \quad \text{with} \quad \varepsilon_{i,g,t} \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right),$$

it immediately follows from Lemma C.1 that

$$\mathbb{E}_{i,g,t}\left[\pi_{t+1}\right] = \omega\mathbb{E}_{i,g,t-1}\left[\pi_t\right] + (\rho-\omega)\,x_{i,g,t},$$

and

$$\omega = \frac{1}{2}\left(\rho + \frac{\sigma_\varepsilon^2+\sigma_\eta^2}{\rho\sigma_\varepsilon^2} - \sqrt{\left(\rho + \frac{\sigma_\varepsilon^2+\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 - 4}\right).$$

Integrating the updating rule for the forecast, we have that

$$\int \mathbb{E}_{i,g,t}\left[\pi_{t+1}\right] = \omega \int \mathbb{E}_{i,g,t-1}\left[\pi_t\right] + (\rho - \omega) \int x_{i,g,t}$$

and, therefore,

$$\overline{\mathbb{E}}_{g,t}\left[\pi_{t+1}\right] = \omega \overline{\mathbb{E}}_{g,t-1}\left[\pi_t\right] + (\rho - \omega)\,\pi_t,$$

which can be rewritten as

$$\overline{\mathbb{E}}_{g,t}\left[\pi_{t+1}\right] = \frac{\rho - \omega}{1 - \omega L}\pi_t.$$

The average forecast error is, then, given by

$$\begin{aligned} \pi_{t+1} - \overline{\mathbb{E}}_{g,t}\left[\pi_{t+1}\right] &= \pi_{t+1} - \frac{\rho - \omega}{1 - \omega L}\pi_t \\ &= \frac{\eta_{t+1}}{1 - \rho L} - \frac{\rho - \omega}{1 - \omega L}\frac{L\eta_{t+1}}{1 - \rho L} \\ &= \frac{\eta_{t+1}}{1 - \omega L}, \end{aligned}$$

which concludes the proof. □

Proof of Proposition 4.2. It follows from Proposition 3.3 that

$$\mathcal{F}_{i,g,t}\left[\pi_{t+1}\right] = (1 + r_g)\,\mathbb{E}_{i,g,t}\left[\pi_{t+1}\right] + \mathcal{B}_g$$

where $\mathbb{E}_{i,g,t}\left[\pi_{t+1}\right]$ denotes the periot-t Bayesian forecast of π_{t+1} of agent i in group g given the $(w_g, 0)$-modified information structure (notice that here $\alpha = 0$). Thus, setting $\xi_t = \pi_t$, $\sigma_\nu^2 = (1 + w)\,\sigma_\eta^2$, and $\sigma_u^2 = \sigma_\varepsilon^2$, it follows from Lemma C.1 that

$$\mathbb{E}_{i,g,t}\left[\pi_{t+1}\right] = \rho\,(1 - \kappa_g)\,\mathbb{E}_{i,g,t-1}\left[\pi_t\right] + \rho\kappa_g x_{i,g,t},$$

with

$$\kappa_g = \frac{1}{2\rho}\left(\left(\rho - \frac{\sigma_\varepsilon^2 + (1 + w_g)\,\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right) - \sqrt{\left(\rho - \frac{\sigma_\varepsilon^2 + (1 + w_g)\,\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 + 4\frac{(1 + w_g)\,\sigma_\eta^2}{\sigma_\varepsilon^2}}\right).$$

It follows that

$$(1 + r_g)\,\mathbb{E}_{i,g,t}\left[\pi_{t+1}\right] + \mathcal{B}_g = \rho\,(1 - \kappa_g)\left((1 + r_g)\,\mathbb{E}_{i,g,t-1}\left[\pi_t\right] + \mathcal{B}_g\right) + (1 + r_g)\,\rho\kappa_g x_{i,g,t} - \rho\,(1 - \kappa_g)\,\mathcal{B}_g + \mathcal{B}_g$$

and, therefore,

$$\mathcal{F}_{i,g,t}\left[\pi_{t+1}\right] = \rho\,(1 - \kappa_g)\,\mathcal{F}_{i,g,t-1}\left[\pi_t\right] + (1 + r_g)\,\rho\kappa_g x_{i,g,t} + (1 - \rho\,(1 - \kappa_g))\,\mathcal{B}_g.$$

Defining $\vartheta_g \equiv \rho\,(1 - \kappa_g)$, we obtain

$$\mathcal{F}_{i,g,t}\left[\pi_{t+1}\right] = \vartheta_g \mathcal{F}_{i,g,t-1}\left[\pi_t\right] + (1 + r_g)\,(\rho - \vartheta_g)\,x_{i,g,t} + (1 - \vartheta_g)\,\mathcal{B}_g,$$

with

$$\vartheta_g = \frac{1}{2}\left(\rho + \frac{\sigma_\varepsilon^2 + (1+w_g)\,\sigma_\eta^2}{\rho\sigma_\varepsilon^2} - \sqrt{\left(\rho + \frac{\sigma_\varepsilon^2 + (1+w_g)\,\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 - 4}\right).$$

Integrating the updating rule for the forecast, we have that

$$\int \mathcal{F}_{i,g,t}\left[\pi_{t+1}\right] = \vartheta_g \int \mathcal{F}_{i,g,t-1}\left[\pi_t\right] + (1+r_g)\left(\rho - \vartheta_g\right) \int x_{i,g,t} + \left(1 - \vartheta_g\right) \mathcal{B}_g$$

and, therefore,

$$\overline{\mathcal{F}}_{g,t}\left[\pi_{t+1}\right] = \vartheta_g \overline{\mathcal{F}}_{g,t-1}\left[\pi_t\right] + (1+r_g)\left(\rho - \vartheta_g\right)\pi_t + \left(1 - \vartheta_g\right)\mathcal{B}_g,$$

which can be rewritten as

$$\overline{\mathcal{F}}_{g,t}\left[\pi_{t+1}\right] = \frac{(1+r_g)\left(\rho - \vartheta_g\right)\pi_t}{1 - \vartheta_g L} + \mathcal{B}_g.$$

The average forecast error is, then, given by

$$\begin{aligned}
\pi_{t+1} - \overline{\mathcal{F}}_{g,t}\left[\pi_{t+1}\right] &= \pi_{t+1} - \frac{(1+r_g)\left(\rho - \vartheta_g\right)\pi_t}{1 - \vartheta_g L} - \mathcal{B}_g \\
&= \frac{(1+r_g)\,\eta_{t+1}}{1 - \vartheta_g L} - \frac{r_g}{1 - \rho L}\eta_{t+1} - \mathcal{B}_g.
\end{aligned}$$

The fact that $r_g > 0$, $w_g > 0$, and $\mathcal{B}_g > 0$ follows immediately from Proposition 3.3 together with the fact that $\delta_g > 0$, $\chi_g > 0$, and $\gamma_g > 0$ established in Lemma 4.1 and that, by assumption, $\lambda > 0$ and $\sigma_\mu^2 > 0$. Finally, to see that $\vartheta_g < \omega$ notice that, from the triangle inequality, we have that

$$\sqrt{\left(\rho + \frac{\sigma_\varepsilon^2 + \sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 - 4} + \sqrt{\left(\frac{w_g\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2} < \sqrt{\left(\rho + \frac{\sigma_\varepsilon^2 + (1+w_g)\,\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 - 4},$$

so that

$$\frac{w_g\sigma_\eta^2}{\rho\sigma_\varepsilon^2} - \sqrt{\left(\rho + \frac{\sigma_\varepsilon^2 + (1+w_g)\,\sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 - 4} < -\sqrt{\left(\rho + \frac{\sigma_\varepsilon^2 + \sigma_\eta^2}{\rho\sigma_\varepsilon^2}\right)^2 - 4}.$$

Adding $\rho + \frac{\sigma_\varepsilon^2 + \sigma_\eta^2}{\rho\sigma_\varepsilon^2}$ and dividing by 2 yields the result. □

Proof of Proposition 5.1. Under rational expectations, the optimal inflation forecast is such that

$$\mathcal{F}_i\left[\pi\right] = \mathbb{E}_i\left[(1-\alpha)\,\pi^* + \alpha\overline{\mathcal{F}}\left[\pi\right]\right].$$

It follows from the the equivalence result in Huo and Pedroni (2020), that the optimal forecast is given by

$$\mathcal{F}_i\left[\pi\right] = \frac{\sigma_\pi^2}{\sigma_\pi^2 + (1-\alpha)^{-1}\sigma_\epsilon^2}\, x_i.$$

Aggregating, we obtain

$$\overline{\mathcal{F}}\left[\pi\right] = \frac{\sigma_\pi^2}{\sigma_\pi^2 + \left(1-\alpha\right)^{-1}\sigma_\epsilon^2}\,\pi^*.$$

Plugging this into the time-invariant inflation policy rule (5.1) completes the proof. □

Proof of Proposition 5.2. To ease notation, let

$$k_i \equiv \mathcal{F}_i\left[\pi\right], \quad \text{and} \quad K = \overline{\mathcal{F}}\left[\pi\right].$$

Plugging (5.1) into the utility function of the agent results in

$$\begin{aligned} u\left(k_i, K, \pi^*\right) &= -\left(k_i - \left(1-\alpha\right)\pi^* - \alpha K\right)^2 - \chi\left(\left(1-\alpha\right)\pi^* + \alpha K\right) \\ &= -\left[\left(1-\alpha\right)\left(k_i - \pi^*\right)^2 + \alpha\left(k_i - K\right)^2\right] - \left(1-\alpha\right)\chi\pi^* + \alpha\left(1-\alpha\right)\left(K-\pi^*\right)^2 - \alpha\chi K. \end{aligned}$$

This is an inefficient economy, so we use Proposition B.1 to characterize the optimal forecasts. Let

$$\begin{aligned} &\lambda_{\text{ineff.}} \equiv 2\lambda, \quad \alpha_{\text{ineff.}} \equiv \alpha, \quad \gamma_{\text{ineff.}} \equiv 0, \quad \chi_{\text{ineff.}} \equiv \frac{1}{2}\left(1-\alpha\right)\chi, \\ &\psi_{\text{ineff.}} \equiv -\alpha\left(1-\alpha\right), \quad \phi_{\text{ineff.}} \equiv 0, \quad \text{and} \quad \varphi_{\text{ineff.}} \equiv \frac{1}{2}\alpha\chi, \end{aligned}$$

where parameters with a subscript "ineff." correspond to the ones in the setup of Proposition B.1. It follows that

$$w = \frac{\tau_\mu}{1 - 2\lambda\left(1-\alpha\right)^2\tau_\mu\mathbb{V}\left(\pi - K\right)}, \quad \text{and} \quad r = 0,$$

where $\tau_\mu \equiv \sigma_\mu^2/\sigma_\pi^2$ is the normalized amount of ambiguity. Moreover, the bias is given by

$$\mathcal{B} = \lambda\left(1-\alpha\right)\chi\tau_\mu\mathbb{V}\left(\pi\right)\left(1-\mathcal{S}\right) + \lambda\alpha\chi\tau_\mu\left[\mathbb{V}\left(\pi\right)\left(1-\mathcal{S}\right) - \mathbb{V}\left(\pi - K\right)\right].$$

Using $\mathbb{V}\left(\pi\right) = \sigma_\pi^2$ and $\mathbb{V}\left(\pi - K\right) = \left(1-\mathcal{S}\right)^2\sigma_\pi^2$, we obtain the desired expressions for sensitivity $\mathcal{S}$ and bias $\mathcal{B}$. Finally, the implied inflation policy directly follows from equation (5.1), which completes the proof. □

Proof of Proposition 5.3. Since the loss function is continuous in σ_μ^2, it is sufficient to show that

$$\left.\frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\sigma_\mu^2}\right|_{\sigma_\mu^2=0} < 0.$$

Fist notice that

$$\mathcal{L} = \frac{\omega}{\alpha}\left[\left(1-\mathcal{R}\right)^2\sigma_\pi^2 + \mathcal{C}^2\right] \quad \Rightarrow \quad \frac{\mathrm{d}\mathcal{L}}{\mathrm{d}\sigma_\mu^2} = \frac{2\omega}{\alpha}\left[-\left(1-\mathcal{R}\right)\sigma_\pi^2\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\sigma_\mu^2} + \mathcal{C}\frac{\mathrm{d}\mathcal{C}}{\mathrm{d}\sigma_\mu^2}\right].$$

If $\sigma_\mu^2 = 0$, it is optimal to set $\mathcal{R} < 1$ and $\mathcal{C} = 0$, so that it is sufficient to show that

$$\left.\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\sigma_\mu^2}\right|_{\sigma_\mu^2=0} > 0,$$

or, equivalently,

$$\left.\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}\sigma_\mu^2}\right|_{\sigma_\mu^2=0} > 0,$$

since $\mathcal{R} = 1 - \alpha + \alpha\mathcal{S}$. Further, notice that sensitivity $\mathcal{S}$ depends on σ_μ^2 only through w and is monotonically increasing in w, it is then sufficient to show that

$$\left.\frac{\mathrm{d}w}{\mathrm{d}\sigma_\mu^2}\right|_{\sigma_\mu^2=0} > 0.$$

This, in turn, follows from the fact that $w = 0$ if $\sigma_\mu^2 = 0$, and $w > 0$ for any $\sigma_\mu^2 > 0$. □

Proof of Proposition 5.4. The optimal inflation forecast must satisfy

$$\mathcal{F}_i[\pi] = (1-\alpha)\,\mathcal{F}_i[\pi^*] + \alpha\mathcal{F}_i\left[\overline{\mathcal{F}}[\pi]\right].$$

With heterogeneous priors, the belief system of agent i is such that

$$\begin{aligned}\mathcal{F}_i[\pi^*] &= \mathbb{E}_i[\pi^*] = \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right) x_i, \quad \text{and}\\ \mathcal{F}_i[\mathcal{F}_j[\pi^*]] &= \mathcal{F}_i\left[\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\,x_j + \frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\,\mathcal{B}\right] = \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^2 x_i + \left(\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\mathcal{B}.\end{aligned}$$

It follows that

$$\begin{aligned}\mathcal{F}_i\left[\overline{\mathcal{F}}[\pi^*]\right] &= \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^2 x_i + \left(\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\mathcal{B}, \quad \text{and}\\ \mathcal{F}_i\left[\mathcal{F}_j\left[\overline{\mathcal{F}}[\pi^*]\right]\right] &= \mathcal{F}_i\left[\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^2 x_j + \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\left(\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\mathcal{B} + \left(\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\mathcal{B}\right],\end{aligned}$$

and, therefore,

$$\mathcal{F}_i\left[\overline{\mathcal{F}}^2[\pi^*]\right] = \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^3 x_i + \left(\sum_{s=0}^{1}\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^s\right)\left(\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\mathcal{B}.$$

Continuing to iterate forwards, we obtain that, for all $k \geq 1$,

$$\begin{aligned}
\mathcal{F}_i\left[\overline{\mathcal{F}}^k\left[\pi^*\right]\right] &= \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^k\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)x_i + \left(\sum_{s=0}^{k-1}\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^s\right)\left(\frac{\sigma_\epsilon^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)\mathcal{B} \\
&= \left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^k\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)x_i + \left(1-\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^k\right)\mathcal{B}.
\end{aligned}$$

Notice that the optimal forecast of agent i can be expressed as a weighted sum of higher-order beliefs,

$$\begin{aligned}
\mathcal{F}_i\left[\pi\right] &= (1-\alpha)\,\mathcal{F}_i\left[\pi^*\right] + (1-\alpha)\sum_{k=1}^{\infty}\alpha^k\mathcal{F}_i\left[\overline{\mathcal{F}}^k\left[\pi^*\right]\right] \\
&= (1-\alpha)\left(\sum_{k=0}^{\infty}\alpha^k\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^k\right)\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)x_i + (1-\alpha)\sum_{k=1}^{\infty}\alpha^k\left(1-\left(\frac{\sigma_\pi^2}{\sigma_\pi^2+\sigma_\epsilon^2}\right)^k\right)\mathcal{B} \\
&= \mathcal{S}^{\mathrm{RE}}\,x_i + \alpha\left(1-\mathcal{S}^{\mathrm{RE}}\right)\mathcal{B},
\end{aligned}$$

where $\mathcal{S}^{\mathrm{RE}} \equiv \frac{\sigma_\pi^2}{\sigma_\pi^2+(1-\alpha)^{-1}\sigma_\epsilon^2}$ denotes the sensitivity under rational expectations.

From the inflation policy in equation (5.1), it follows that

$$\mathcal{R} = 1-\alpha+\alpha\mathcal{S}^{\mathrm{RE}} = \mathcal{R}^{\mathrm{RE}}, \quad \text{and} \quad \mathcal{C} = \alpha\left(\alpha-\alpha\mathcal{S}^{\mathrm{RE}}\right)\mathcal{B} = \alpha\left(1-\mathcal{R}^{\mathrm{RE}}\right)\mathcal{B}.$$

Finally, the social loss function is given by

$$\mathcal{L} = \frac{\omega}{\alpha}\left[\left(1-\mathcal{R}\right)^2\sigma_\pi^2 + \mathcal{C}^2\right],$$

which is increasing in $\mathcal{B}$ since $\mathcal{C} = \alpha\left(1-\mathcal{R}^{\mathrm{RE}}\right)\mathcal{B}$. $\square$

D Robust Preferences: Derivations and Proofs

Lemma D.1. *Taking the law of motion of K_t as given, individual i's best response satisfies*

$$k_{it} = (1-\alpha)\,\mathcal{F}_{it}[\xi_t] + \alpha\mathcal{F}_{it}[K_t],$$

where $\mathcal{F}_{it}[\cdot]$ denotes agent i's subjective expectation, such that for any random variable X,

$$\mathcal{F}_{it}[X] \equiv \int X\, \tilde{\mathrm{p}}_{it}(X)\mathrm{d}X, \quad \textit{with} \quad \tilde{\mathrm{p}}_{it}(X) \propto \exp\left(-\varpi u(k_{it}, K_t, \xi_t)\right) \mathrm{p}(X \mid x_i^t).$$

Proof of Lemma D.1. The first-order-condition for the minimization with respect to m_{it} is given by

$$u\left(k_{it}, K_t, \xi_t\right) + \frac{1}{\varpi}\log m_{it} + \frac{1}{\varpi} = 0.$$

Together with the fact that $\mathbb{E}_{it}[m_{it}] = 1$, it follows that

$$m_{it} = \frac{\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)}{\mathbb{E}_{it}\left[\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)\right]}.$$

Thus, problem (6.1) can be rewritten as the following problem with risk sensitivity:

$$\max_{k_{it}} -\frac{1}{\varpi}\log\left(\mathbb{E}_{it}\left[\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)\right]\right).$$

The first-order-condition for this problem with respect to k_{it} is given by

$$\frac{\mathbb{E}_{it}\left[\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)\frac{\partial u(k_{it}, K_t, \xi_t)}{\partial k_{it}}\right]}{\mathbb{E}_{it}\left[\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)\right]} = 0.$$

Since

$$\frac{\partial u(k_{it}, K_t, \xi_t)}{\partial k_{it}} = k_{it} - (1-\alpha)\,\xi_t - \alpha K_t,$$

it follows that

$$k_{it} = (1-\alpha)\,\mathbb{E}_{it}\left[\xi_t \frac{\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)}{\mathbb{E}_{it}\left[\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)\right]}\right] + \alpha\mathbb{E}_{it}\left[K_t \frac{\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)}{\mathbb{E}_{it}\left[\exp\left(-\varpi u\left(k_{it}, K_t, \xi_t\right)\right)\right]}\right].$$

Letting $\frac{\exp(-\varpi u(k_{it}, K_t, \xi_t))}{\mathbb{E}_{it}[\exp(-\varpi u(k_{it}, K_t, \xi_t))]}$ be the Radon-Nikodym derivative completes the proof. □

Proof of Proposition 6.1. Consider the same truncated version of the model used in the proof of Proposition 3.3. From Lemma D.1 we have that the optimal strategy then satisfies that

$$k_i = (1-\alpha)\,\mathcal{F}[\theta \mid x_i] + \alpha\mathcal{F}[K \mid x_i], \tag{D.1}$$

with the distorted posterior given by

$$\tilde{p}(\eta|x_i) \propto \exp(-\varpi u(k_i, K, \theta))\, p(\eta \mid x_i).$$

We proceed with a guess-and-verify strategy. First we guess that

$$k_i = h' B\nu_i + h_0.$$

Substituting this into equation (D.1), it follows that

$$k_i = ((1-\alpha) A\mathcal{K}' + \alpha h' B\mathcal{K}')\, \mathcal{F}[\eta \mid B\nu_i] + \alpha h_0.$$

Thus, we need to determine the subjective conditional expectation $\mathcal{F}[\eta \mid B\nu_i]$. We proceed to characterize the distorted posterior $\tilde{p}(\eta \mid B\nu_i)$ by the following three steps:

1. First, the Bayesian posterior $p(\eta \mid B\nu_i)$ is such that

$$p(\eta \mid B\nu_i) \propto \exp\left(-\frac{1}{2}\left(\eta - \mu_{\eta|B\nu_i}\right)' \Sigma^{-1}_{\eta|B\nu_i}\left(\eta - \mu_{\eta|B\nu_i}\right)\right),$$

   with the conditional mean and variance of given by

$$\mu_{\eta|B\nu_i} = \mathcal{K}\Omega B' \left(B\Omega B'\right)^{-1} B\nu_i, \quad \text{and} \quad \Sigma_{\eta|B\nu_i} = \mathcal{K}\Omega\mathcal{K}' - \mathcal{K}\Omega B' \left(B\Omega B'\right)^{-1} B\Omega\mathcal{K}'.$$

2. Second, notice that

$$\begin{aligned}
u(k, K, \theta) = &-\frac{1}{2}\left[(1-\alpha)\left(h' B\nu_i + h_0 - A\mathcal{K}'\eta\right)^2 + \alpha\left(h' B\nu_i - h' B\mathcal{K}'\eta\right)^2\right] - \chi A\mathcal{K}'\eta - \frac{1}{2}\gamma\eta'\mathcal{K}A'A\mathcal{K}'\eta \\
&= \text{constant} - \frac{1}{2}\gamma\eta'\mathcal{K}A'A\mathcal{K}'\eta - \frac{1}{2}\left[(1-\alpha)\,\eta'\mathcal{K}A'A\mathcal{K}'\eta + \alpha\eta'\mathcal{K}B'hh'B\mathcal{K}'\eta\right] \\
&\quad + \frac{1}{2}\left[(1-\alpha)\left(h_0 + \nu_i' B' h\right) A + \alpha\nu_i' B' hh' B - \chi A\right]\mathcal{K}'\eta \\
&\quad + \eta'\mathcal{K}\frac{1}{2}\left[(1-\alpha)\, A'\left(h_0 + h' B\nu_i\right) + \alpha B' hh' B\nu_i - \chi A'\right],
\end{aligned}$$

   with the constant independent of η.

3. Finally, putting these results together, the distorted posterior must be such that

$$\tilde{p}(\eta \mid B\nu_i) \propto \exp\left(-\frac{1}{2}\eta'\tilde{\Sigma}^{-1}_{\eta|B\nu_i}\eta + \frac{1}{2}\tilde{\mu}'_{\eta|B\nu_i}\tilde{\Sigma}^{-1}_{\eta|B\nu_i}\eta + \frac{1}{2}\eta'\tilde{\Sigma}^{-1}_{\eta|B\nu_i}\tilde{\mu}_{\eta|B\nu_i}\right)$$

   where the distorted posterior variance and mean are given by

$$\tilde{\Sigma}^{-1}_{\eta|B\nu_i} \equiv \Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q} \quad \text{and} \quad \tilde{\mu}_{\eta|B\nu_i} \equiv \tilde{\Sigma}_{\eta|B\nu_i}\left(\Sigma^{-1}_{\eta|B\nu_i}\mu_{\eta|B\nu_i} + \mathrm{R}B\nu_i\right) + \pi_\mu,$$

with the matrices Q and R and the vector π_μ given by

$$\mathrm{Q} \equiv -\varpi\gamma\mathcal{K}A'A\mathcal{K}' - \varpi\left[(1-\alpha)\mathcal{K}A'A\mathcal{K}' + \alpha\mathcal{K}B'hh'B\mathcal{K}'\right], \tag{D.2}$$

$$\mathrm{R} \equiv -\varpi\mathcal{K}\left[(1-\alpha)A' + \alpha B'h\right]h', \tag{D.3}$$

$$\pi_\mu \equiv -\varpi\left(\Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q}\right)^{-1}\mathcal{K}\left[(1-\alpha)A'h_0 - \chi A'\right]. \tag{D.4}$$

The distorted expectation under robust preferences can, then, be written as

$$\tilde{\mathbb{E}}\left[\eta \mid B\nu_i\right] = \tilde{\mu}_{\eta|B\nu_i} = \mathrm{M}B\nu_i + \pi_\mu,$$

with the matrix M given by

$$\mathrm{M} \equiv \left(\Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q}\right)^{-1}\left(\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega B'\left(B\Omega B'\right)^{-1} + \mathrm{R}\right). \tag{D.5}$$

Thus, we that that

$$k_i = \left((1-\alpha)A\mathcal{K}' + \alpha h'B\mathcal{K}'\right)\left(\mathrm{M}B\nu_i + \pi_\mu\right) + \alpha h_0.$$

and for the initial guess to be correct the following fixed-point conditions must be satisfied:

$$h' = \left[(1-\alpha)A + \alpha h'B\right]\mathcal{K}'\mathrm{M}, \tag{D.6}$$

$$h_0 = \left[(1-\alpha)A + \alpha h'B\right]\mathcal{K}'\pi_\mu + \alpha h_0. \tag{D.7}$$

In what follows, we first characterize the responsiveness to signals h that solves equation (D.6) and then characterize the bias h_0 that solves equation (D.7).

Characterization of the responsiveness, h. We start by rewriting the equation for the matrix M. Substituting h' from equation (D.6) into equation (D.3), we obtain

$$\mathrm{R} = -\varpi\mathcal{K}\left((1-\alpha)A' + \alpha B'h\right)\left((1-\alpha)A + \alpha h'B\right)\mathcal{K}'\mathrm{M}$$

Plugging this expression for R into the definition of M, equation (D.5), it follows that

$$\left(\Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q}\right)\mathrm{M} = \left(\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega B'\left(B\Omega B'\right)^{-1} - \varpi\mathcal{K}\left((1-\alpha)A' + \alpha B'h\right)\left((1-\alpha)A + \alpha h'B\right)\mathcal{K}'\mathrm{M}\right).$$

Solving for M we get

$$\mathrm{M} = \left(\mathrm{I}_u + \Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\right)^{-1}\mathcal{K}\Omega B'\left(B\Omega B'\right)^{-1},$$

where the I_u is the identity matrix of dimension u and the matrix $\tilde{\mathrm{Q}}$ is given by

$$\begin{aligned}\tilde{\mathrm{Q}} &\equiv \mathrm{Q} + \varpi\mathcal{K}\left((1-\alpha)A' + \alpha B'h\right)\left((1-\alpha)A + \alpha h'B\right)\mathcal{K}' \\ &= -\varpi\gamma\mathcal{K}A'A\mathcal{K}' - \varpi\alpha(1-\alpha)\mathcal{K}\left(B'h - A'\right)\left(h'B - A\right)\mathcal{K}'.\end{aligned} \tag{D.8}$$

To ease notation, we define matrices

$$Z_1 \equiv -\varpi\gamma\mathcal{K}A' - \varpi\alpha\left(1-\alpha\right)\mathcal{K}\left(A' - B'h\right), \quad \text{and} \quad Z_2 \equiv -\varpi\alpha\left(1-\alpha\right)\mathcal{K}\left(B'h - A'\right),$$

so that

$$\tilde{Q} = \mathrm{Z}_1 A\mathcal{K}' + \mathrm{Z}_2 h' B\mathcal{K}'.$$

The fixed-point condition (D.6) can, then, be rewritten as

$$h' = \left[\left(1-\alpha\right)A + \alpha h' B\right]\mathcal{K}'\left(\mathrm{I}_u + \Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\right)^{-1}\Omega_\eta\mathcal{K}B'\left(B\Omega B'\right)^{-1},$$

where we used the fact that $\mathcal{K}\Omega = \Omega_\eta\mathcal{K}$. Using the Woodbury matrix identity, we obtain

$$\left(\mathrm{I}_u + \Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\right)^{-1}\Omega_\eta = \Omega_\eta - \left(\mathrm{I}_u + \Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\right)^{-1}\Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\Omega_\eta,$$

so, we can further rewrite the fixed-point condition as

$$\begin{aligned}
h' =& \left[\left(1-\alpha\right)A + \alpha h' B\right]\mathcal{K}'\left(\Omega_\eta - \left(\mathrm{I}_u + \Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\right)^{-1}\Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\Omega_\eta\right)\mathcal{K}B'\left(B\Omega B'\right)^{-1} \\
&= \left[\left(1-\alpha\right)A + \alpha h' B\right]\mathcal{K}'\left(\Omega_\eta - \left(\mathrm{I}_u + \Sigma_{\eta|B\nu_i}\tilde{\mathrm{Q}}\right)^{-1}\Sigma_{\eta|B\nu_i}\left(\mathrm{Z}_1 A\mathcal{K}' + \mathrm{Z}_2 h' B\mathcal{K}'\right)\Omega_\eta\right)\mathcal{K}B'\left(B\Omega B'\right)^{-1} \\
&= \left(1-\alpha+\varkappa_1\right)A\Lambda\Omega B' + \left(\alpha-\varkappa_2\right)h' B\Lambda\Omega B',
\end{aligned}$$

where $\Lambda = \mathcal{K}'\mathcal{K}$ and the endogenous scalars $\varkappa_1$ and $\varkappa_2$ are given by

$$\begin{aligned}
\varkappa_1 &\equiv -\left[\left(1-\alpha\right)A + \alpha h' B\right]\mathcal{K}'\left(I_u + \Sigma_{\eta|B\nu_i}\tilde{Q}\right)^{-1}\Sigma_{\eta|B\nu_i}Z_1, \\
\varkappa_2 &\equiv \left[\left(1-\alpha\right)A + \alpha h' B\right]\mathcal{K}'\left(I_u + \Sigma_{\eta|B\nu_i}\tilde{Q}\right)^{-1}\Sigma_{\eta|B\nu_i}Z_2.
\end{aligned}$$

Solving for h' we obtain

$$h' = \frac{1-\alpha+\varkappa_1}{1-\alpha+\varkappa_2}A\Lambda\hat{\Omega}B'\left(B\hat{\Omega}B'\right)^{-1}, \tag{D.9}$$

where the transformed variance-covariance matrix $\hat{\Omega}$ is given by

$$\hat{\Omega} \equiv \frac{1-\alpha+\varkappa_2}{1-\alpha}\Lambda\Omega + \frac{1}{1-\alpha}\left(\mathrm{I}_m - \Lambda\right)\Omega, \tag{D.10}$$

with I_m denoting the identity matrix of dimension m.

In what follows, we provide expressions for the two endogenous scalars $(\varkappa_1, \varkappa_2)$ such that we can take the limit as $T \to \infty$ and obtain the formulas in Proposition 6.1. For this purpose, it is useful to define

$$\mathrm{X} \equiv \left[\left(1-\alpha\right)A + \alpha h' B\right]\mathcal{K}'\left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1}.$$

Notice that $(\varkappa_1, \varkappa_2)$ can then be written as

$$\varkappa_1 = -\mathrm{X}Z_1 = \varpi\gamma\mathrm{X}\mathcal{K}A' + \varkappa_2, \quad \text{and} \quad \varkappa_2 = -\mathrm{X}Z_2 = \varpi\alpha\,(1-\alpha)\,\mathrm{X}\mathcal{K}\,(A' - B'h)\,.$$

Therefore, it follows that

$$\begin{aligned}
\mathrm{X} &= [(1-\alpha)\,A + \alpha h'B]\,\mathcal{K}'\left(\Sigma_{\eta|B\nu_i} - \left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1}\tilde{\mathrm{Q}}\Sigma_{\eta|B\nu_i}\right) \\
&= [(1-\alpha)\,A + \alpha h'B]\,\mathcal{K}'\Sigma_{\eta|B\nu_i} - \mathrm{X}\tilde{\mathrm{Q}}\Sigma_{\eta|B\nu_i} \\
&= [(1-\alpha)\,A + \alpha h'B]\,\mathcal{K}'\Sigma_{\eta|B\nu_i} - \mathrm{X}\,(\mathrm{Z}_1 A\mathcal{K}' + \mathrm{Z}_2 h'B\mathcal{K}')\,\Sigma_{\eta|B\nu_i} \\
&= [(1-\alpha)\,A + \alpha h'B]\,\mathcal{K}'\Sigma_{\eta|B\nu_i} + (\varkappa_1 A\mathcal{K}' - \varkappa_2 h'B\mathcal{K}')\,\Sigma_{\eta|B\nu_i} \\
&= (1-\alpha+\varkappa_1)\,A\mathcal{K}'\Sigma_{\eta|B\nu_i} + (\alpha - \varkappa_2)\,h'B\mathcal{K}'\Sigma_{\eta|B\nu_i}.
\end{aligned}$$

Thus, since $\varkappa_1 - \varkappa_2 = \varpi\gamma\mathrm{X}\mathcal{K}A'$, we have that,

$$\varkappa_1 - \varkappa_2 = \varpi\gamma\,(1-\alpha+\varkappa_1)\,A\mathcal{K}'\Sigma_{\eta|B\nu_i}\mathcal{K}A' + \varpi\gamma\,(\alpha-\varkappa_2)\,h'B\mathcal{K}'\Sigma_{\eta|B\nu_i}\mathcal{K}A'. \tag{D.11}$$

Next, notice that

$$\mathrm{X} = \mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\Sigma_{\eta|B\nu_i} = \mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega\mathcal{K}' - \mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega B'\,(B\Omega B')^{-1}\,B\Omega\mathcal{K}' = \mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega\mathcal{K}' - h'B\Omega\mathcal{K}',$$

where the second equality uses the definition of $\Sigma_{\eta|B\nu_i}$ and the last equality uses the fact that

$$h' = \mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega B'\,(B\Omega B')^{-1}\,.$$

Rearranging terms and right-multiplying $(\mathcal{K}\Omega\mathcal{K}')^{-1}\,\mathcal{K}\Omega B'$ to both sides of the equation, we obtain

$$\mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega B' = \mathrm{X}\,(\mathcal{K}\Omega\mathcal{K}')^{-1}\,\mathcal{K}\Omega B' + h'B\Omega\mathcal{K}'\,(\mathcal{K}\Omega\mathcal{K}')^{-1}\,\mathcal{K}\Omega B' = \mathrm{X}\mathcal{K}B' + h'B\Lambda\Omega B'.$$

Further, since $\mathrm{X}\Sigma^{-1}_{\eta|B\nu_i}\mathcal{K}\Omega B' = h'B\Omega B'$, it follows that

$$\mathrm{X}\mathcal{K}B'h = h'B\,(\mathrm{I}_m - \Lambda)\,\Omega B'h.$$

Hence, we have that

$$\varkappa_2 = \varpi\alpha\,(1-\alpha)\,\mathrm{X}\mathcal{K}A' - \varpi\alpha\,(1-\alpha)\,\mathrm{X}\mathcal{K}B'h \tag{D.12}$$

$$= \frac{\alpha\,(1-\alpha)}{\gamma}\,(\varkappa_1 - \varkappa_2) - \varpi\alpha\,(1-\alpha)\,h'B\,(\mathrm{I}_m - \Lambda)\,\Omega B'h, \tag{D.13}$$

where we use the fact that $\varpi\gamma\mathrm{X}\mathcal{K}A' = \varkappa_1 - \varkappa_2$.

Given the above results, we are left with taking the limit as $T \to \infty$ of the truncated problem. In particular, we have that

$$\lim_{T\to\infty} A\Lambda\hat{\Omega}B'\left(B\hat{\Omega}B'\right)^{-1} = p\left(L; w, \alpha\right) \qquad \lim_{T\to\infty} A\mathcal{K}'\Sigma_{\eta|B\nu_i}\mathcal{K}A' = \mathbb{V}_{it}\left(\xi_t\right)$$

$$\lim_{T\to\infty} h'B\mathcal{K}'\Sigma_{\eta|B\nu_i}\mathcal{K}A' = \mathbb{COV}_{it}\left(K_t, \xi_t\right) \qquad \lim_{T\to\infty} h'B\left(\mathrm{I}_m - \Lambda\right)\Omega B'h = \mathbb{DISP}\left(k_{it}\right)$$

which, together with equations (D.9), (D.10), (D.11), and (D.13), completes the characterization of the responsiveness to signals.

Characterization of the bias, h_0. From the fixed-point condition (D.7) and the definition of π_μ in equation (D.4), it follows that

$$\left(1-\alpha\right)h_0 = \varpi\left[\left(1-\alpha\right)A + \alpha h'B\right]\mathcal{K}'\left(\Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q}\right)^{-1}\mathcal{K}\left[\chi A' - \left(1-\alpha\right)A'h_0\right],$$

which can be solved for h_0 implying

$$h_0 = \frac{\chi\varpi\mathrm{Y}}{\left(1-\alpha\right)\left(1+\varpi\mathrm{Y}\right)},$$

with Y given by

$$\mathrm{Y} \equiv \left[\left(1-\alpha\right)A + \alpha h'B\right]\mathcal{K}'\left(\Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q}\right)^{-1}\mathcal{K}A'.$$

Using the definition of $\tilde{\mathrm{Q}}$ in equation (D.8) and the Woodbury matrix identity, it follows that

$$\begin{aligned}
\left(\Sigma^{-1}_{\eta|B\nu_i} + \mathrm{Q}\right)^{-1} &= \left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}} - \varpi\mathcal{K}\left(\left(1-\alpha\right)A' + \alpha B'h\right)\left(\left(1-\alpha\right)A + \alpha h'B\right)\mathcal{K}'\right)^{-1} \\
&= \left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1} + \\
&\quad \frac{\varpi\left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1}\mathcal{K}\left(\left(1-\alpha\right)A' + \alpha B'h\right)\left(\left(1-\alpha\right)A + \alpha h'B\right)\mathcal{K}'\left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1}}{1 - \varpi\left(\left(1-\alpha\right)A + \alpha h'B\right)\mathcal{K}'\left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1}\mathcal{K}\left(\left(1-\alpha\right)A' + \alpha B'h\right)} \\
&= \left(\Sigma^{-1}_{\eta|B\nu_i} + \tilde{\mathrm{Q}}\right)^{-1} + \frac{\varpi\mathrm{X}'\mathrm{X}}{1 - \varpi\mathrm{X}\mathcal{K}\left(\left(1-\alpha\right)A' + \alpha B'h\right)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathrm{Y} &= \mathrm{X}\mathcal{K}A' + \frac{\varpi\left[\left(1-\alpha\right)A + \alpha h'B\right]\mathcal{K}'\mathrm{X}'\mathrm{X}\mathcal{K}A'}{1 - \varpi\mathrm{X}\mathcal{K}\left(\left(1-\alpha\right)A' + \alpha B'h\right)} \\
&= \frac{\mathrm{X}\mathcal{K}A'}{1 - \varpi\mathrm{X}\mathcal{K}\left(\left(1-\alpha\right)A' + \alpha B'h\right)} \\
&= \frac{\frac{\varkappa_1 - \varkappa_2}{\varpi\gamma}}{1 - \left(\frac{1-\alpha}{\gamma}\right)\left(\varkappa_1 - \varkappa_2\right) - \varpi\alpha h'B\left(\mathrm{I}_m - \Lambda\right)\Omega B'h},
\end{aligned}$$

where the last equality uses the fact that

$$\varkappa_1 - \varkappa_2 = \varpi\gamma\mathrm{X}\mathcal{K}A', \quad \text{and} \quad \mathrm{X}\mathcal{K}B'h = h'B\left(\mathrm{I}_m - \Lambda\right)\Omega B'h.$$

Therefore, we have that

$$h_0 = \frac{\chi(\varkappa_1 - \varkappa_2)}{(1-\alpha)\left(\gamma + \alpha\left(\varkappa_1 - \varkappa_2\right) - \gamma\varpi\alpha h' B\left(\mathrm{I}_m - \Lambda\right)\Omega B' h\right)}.$$

Finally, taking the limit as $T \to \infty$ leads to

$$\mathcal{B} = \lim_{T\to\infty} h_0 = \frac{\chi(\varkappa_1 - \varkappa_2)}{(1-\alpha)\left(\gamma + \alpha\left(\varkappa_1 - \varkappa_2\right) - \gamma\varpi\alpha\mathbb{DISP}\left(k_{it}\right)\right)}.$$

$\square$

Proof of Corollary 6.1. Observe that, by using (3.16), the expression of w under smooth model (3.15) can be transformed into

$$w = \left[\frac{1}{\tau_\mu} - \lambda\left(1-\alpha\right)\left(\mathbb{V}\left(\xi_t - K_t\right) + r\frac{1+w}{w}\left(1-\mathcal{S}\right)\mathbb{V}\left(\xi_t\right)\right)\right]^{-1} \tag{D.14}$$

Take any pair (w, r) and the associated sensitivity $\mathcal{S}$ that would arise from robust preferences. We may solve $\left(\lambda, \sigma_\mu^2\right)$ from (3.16) and (D.14). Note that the first condition $w \geq 0, r \geq 0, \mathcal{S} \leq 1$ ensures that Assumption 2 can be satisfied and the second condition$(1-\mathcal{S})\left(\frac{\gamma w}{(1+w)r} - \frac{(1-\alpha)(1+w)r}{w}\right) + \gamma > (1-\alpha)\frac{\mathbb{V}(\xi_t - K_t)}{\mathbb{V}(\xi_t)}$ ensures that the resulted $\tau_\mu > 0$.

$\square$

E Ambiguity about Variance

In this section, we explore the case in which the ambiguity is about the variance of the fundamental. Specifically, we assume that the fundamental ξ follows a normal distribution with mean 0 and variance $\sigma^2_{\xi,*}$: $\xi \sim \mathcal{N}\left(0, \sigma^2_{\xi,*}\right)$. Agents exhibit ambiguity regarding the true variance of the fundamental, $\sigma^2_{\xi,*}$. We let Γ_ξ be the the range of possible values for the variance of the fundamental, σ^2_ξ. Analysts believe that $\sigma^2_\xi \in \Gamma_\xi$ and have some prior belief about Γ_ξ with density distribution given by $p\left(\sigma^2_\xi\right)$. To ensure that strategies based on Bayesian inference and ambiguity neutrality coincide, we impose the following assumption on the agents' prior belief:

Assumption 3. *The prior belief of the agent is such that*

$$\int_{\Gamma_\xi} \sigma^2_\xi \, p\left(\sigma^2_\xi\right) \mathrm{d}\sigma^2_\xi = \sigma^2_{\xi,*} \, .$$

Similar to the setup of ambiguity about the mean of the fundamental, each agent receives a private signal

$$x_i = \xi + \varepsilon_i, \quad \text{with} \quad \varepsilon_i \sim \mathcal{N}\left(0, \sigma^2_\epsilon\right).$$

Agents are ambiguity averse and select a strategy $g(x_i)$ to minimize the following objective:

$$\mathcal{L}(g) = \phi^{-1}\left(\int_{\Gamma_\xi} \phi\left(\mathbb{E}^{\tau_\xi}\left[(g(x_i) - \xi)^2 - \chi\xi\right]\right) p(\tau_\xi)\mathrm{d}\tau_\xi\right),$$

where $\phi(x) = \frac{1}{\lambda}\exp(\lambda x)$ takes the CAAA form with λ representing the degree of ambiguity aversion. Finally, we restrict our analysis to linear strategies such that

$$g(x_i) = s x_i + b,$$

which facilitates a direct comparison with our baseline setup, where ambiguity pertains to the mean of the fundamental.

The following proposition suggests that ambiguity has a more limited effect, leading to an optimal linear strategy that exhibits higher sensitivity compared to the rational RE benchmark, but no bias.

Proposition E.1. *When agents are ambiguity averse, $\lambda > 0$, the optimal linear strategy exhibits higher sensitivity than the RE benchmark and features no bias:*

$$s^* > s^{RE} \equiv \frac{\sigma^2_{\xi,*}}{\sigma^2_{\xi,*} + \sigma^2_\epsilon}, \quad \textit{and} \quad b^* = 0.$$

Proof. Given the restriction to linear strategies, the objective function of the agents can be written as a function of the sensitivity, s, and bias, b, as follows

$$\mathcal{L}(s,b) = \frac{1}{\lambda}\ln\left(\int_{\Gamma_\xi} \exp\left(\lambda\left((s-1)^2\sigma^2_\xi + s^2\sigma^2_\epsilon\right)\right) p\left(\sigma^2_\xi\right)\mathrm{d}\sigma^2_\xi\right) + \frac{1}{2}b^2.$$

The zero-bias result is straight-forward: the FOC with respect to bias b is such that

$$\frac{\partial \mathcal{L}\left(s,b\right)}{\partial b}=b=0.$$

To characterize the optimal of sensitivity, s, we consider the corresponding FOC,

$$\frac{\partial \mathcal{L}\left(s,b\right)}{\partial s}=\frac{\int_{\Gamma_\xi}\exp\left(\lambda\left(\left(s-1\right)^2\sigma_\xi^2+s^2\sigma_\epsilon^2\right)\right)\left[\left(s-1\right)\sigma_\xi^2+s\sigma_\varepsilon^2\right]p\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2}{\int_{\Gamma_\xi}\exp\left(\lambda\left(\left(s-1\right)^2\sigma_\xi^2+s^2\sigma_\epsilon^2\right)\right)p\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2}=0,$$

which is equivalent to

$$s\,\sigma_\varepsilon^2=\left(1-s\right)\int_{\Gamma_\xi}\sigma_\xi^2\,\hat{p}\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2,$$

where the distorted belief $\hat{p}(\sigma_\xi^2)$ is such that

$$\hat{p}\left(\tau_\xi\right)\propto\exp\left(\lambda\left(s-1\right)^2\sigma_\xi^2\right)p\left(\sigma_\xi^2\right).$$

Notice that, relative to the agents' prior $p(\sigma_\xi^2)$, the distorted belief $\hat{p}(\sigma_\xi^2)$ puts higher weights on the larger σ_ξ^2 in Γ_ξ: $\hat{p}(\sigma_\xi^2)$ first-order stochastically dominates $p(\sigma_\xi^2)$. It follows that

$$\int_{\Gamma_\xi}\sigma_\xi^2\,\hat{p}\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2\geq\int_{\Gamma_\xi}\sigma_\xi^2\,p\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2=\sigma_{\xi,*}^2\,,$$

and, therefore,

$$s^*=\frac{\int_{\Gamma_\xi}\sigma_\xi^2\,\hat{p}\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2}{\int_{\Gamma_\xi}\sigma_\xi^2\,\hat{p}\left(\sigma_\xi^2\right)\mathrm{d}\sigma_\xi^2+\sigma_\epsilon^2}>\frac{\sigma_{\xi,*}^2}{\sigma_{\xi,*}^2+\sigma_\epsilon^2}=s^{\mathrm{RE}}.$$

□

F Evidence on Inflation Expectations by Income Group

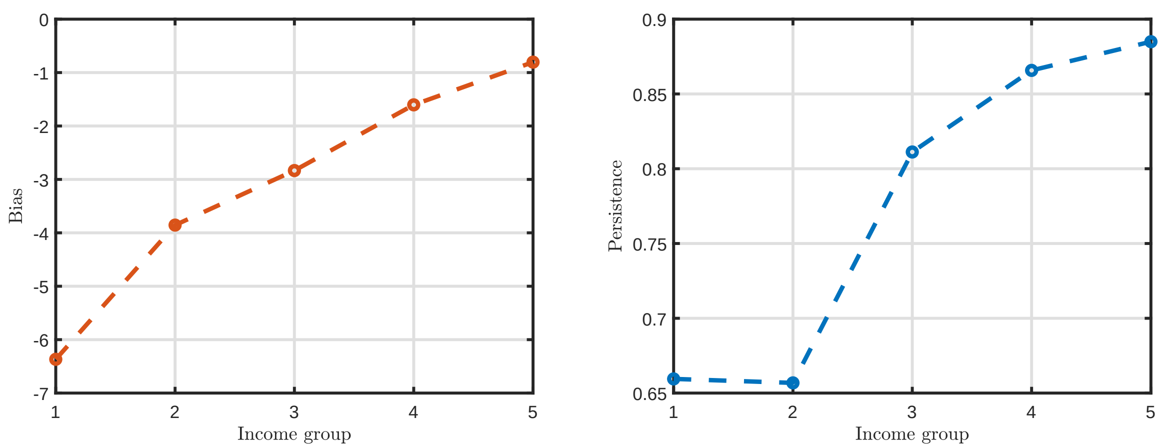
We investigate the joint behaviors of bias and persistence in forecast errors using both the Michigan Survey of Consumers (MSC) and the Survey of Consumer Expectations (SCE). We examine two regression equations:

$$\overline{\mathrm{FE}}_{g,t} = \sum_{g=1}^{N} \beta_g \mathcal{I}_g + \omega_{g,t},$$

$$\overline{\mathrm{FE}}_{g,t} = \sum_{g=1}^{N} \beta_g \mathcal{I}_g + \sum_{g=1}^{N} \alpha_g \overline{\mathrm{FE}}_{g,t-1} + \omega_{g,t},$$

where $\overline{\mathrm{FE}}_{g,t}$ represents the average forecast errors for group g at year-quarter t and $\mathcal{I}_g$ is the group dummy. For the MSC dataset, we divide individuals into $N = 7$ income groups, while for the SCE dataset, we divide individuals into $N = 5$ income groups. Table F.1 provides the results of our analysis. We use the poorest group (Group 1) as the reference group when reporting the results. The overall patterns of bias and persistence are similar in both the MSC and SCE datasets: as the income level increases, the amount of bias decreases, while the persistence of forecast errors increases. Similar to Figure 4.1 that displays the empirical patterns in MSC, Figure F.1 plot the point estimates of the biases and the persistence across different income groups in SCE.

FIGURE F.1: Bias and Persistence of Forecast Error in the Survey Data (NYSCE)



(A) Bias in forecast (B) Persistence of forecast error

Note: This figure reports bias (Panel A) and persistence (Panel B) of households' inflation forecasts in the cross-section of the income distribution. Bias and persistence of each income percentile are calculated by the mean and serial correlation of forecast errors of households' inflation expectations for the next 12 months. Data are obtained from the Survey of Consumer Expectations, NY Fed between 2013:II and 2022:II.

To address the concern that bias may be influenced by other observed individual characteristics, such as age and resident state, we introduce the following empirical specification at the individual level for both the MSC

Table F.1: Bias and Persistence of Forecast Errors: MSC and SCE

	MSC		SCE	
	Bias	Persistence	Bias	Persistence
Constant	-2.280***	-0.971***	-6.368***	-2.196***
	(0.085)	(0.122)	(0.178)	(0.223)
Group 2	0.235**	0.329**	2.512**	0.852**
	(0.068)	(0.097)	(0.119)	(0.161)
Group 3	0.766***	0.538***	3.533***	1.682***
	(0.059)	(0.060)	(0.153)	(0.139)
Group 4	1.102***	0.634***	4.764***	2.022***
	(0.068)	(0.036)	(0.151)	(0.032)
Group 5	1.258***	0.720***	5.563***	2.147***
	(0.062)	(0.040)	(0.171)	(0.063)
Group 6	1.535***	0.798***		
	(0.059)	(0.038)		
Group 7	1.924***	0.887***		
	(0.052)	(0.075)		
FE_{t-1}		0.565***		0.659***
		(0.058)		(0.029)
$FE_{t-1} \times$ Group 2		0.111**		-0.003
		(0.041)		(0.070)
$FE_{t-1} \times$ Group 3		0.140*		0.152
		(0.054)		(0.079)
$FE_{t-1} \times$ Group 4		0.133*		0.206**
		(0.059)		(0.064)
$FE_{t-1} \times$ Group 5		0.168*		0.225**
		(0.073)		(0.063)
$FE_{t-1} \times$ Group 6		0.178*		
		(0.078)		
$FE_{t-1} \times$ Group 7		0.153*		
		(0.077)		
Time fixed effects	Yes	Yes	Yes	Yes
Obs.	952	945	180	175

* p<0.1, ** p<0.05, *** p<0.001.

and the SCE:

$$\text{FE}_{i,t} = \sum_{g=1}^{N} \beta_g \mathcal{I}_{i,g,t} + \gamma' \text{X}_{i,t} + \delta_t + \omega_{i,t},$$

where $\mathcal{I}_{i,g,t}$ is a dummy variable that equals to 1 if individual i belongs to income group g at year-month t, and $\text{X}_{i,t}$ is a vector of observed individual characteristics. For the MSC dataset, we control for age, gender, education, birth cohort, marital status, region, and the number of kids and adults in the household. It is worth noting that controlling for the birth cohort helps address concerns regarding the impact of inflation experiences on households' inflation expectations (Malmendier and Nagel, 2016). For the SCE dataset, we control for age group, numeracy, education, and region. Table F.2 reports the results. Again, we use the poorest group (Group 1) as the base group for both the MSC and SCE datasets. Even after controlling for additional individual characteristics, the biases in forecasts persist and exhibit a negative correlation with households' income levels.

Table F.2: Bias of Forecast Errors Controlling Individual Characteristics: MSC and SCE

	MSC	SCE
Constant	-2.602***	-5.169***
	(0.322)	(0.239)
Group 2	0.178***	1.801**
	(0.034)	(0.110)
Group 3	0.620***	2.414***
	(0.032)	(0.126)
Group 4	0.904***	3.184***
	(0.035)	(0.185)
Group 5	1.049***	3.713***
	(0.039)	(0.245)
Group 6	1.261***	
	(0.027)	
Group 7	1.554***	
	(0.033)	
Demographics	Yes	Yes
Birth Cohort	Yes	No
Age	Yes	Yes
Region	Yes	Yes
Time fixed effects	Yes	Yes
Obs.	140,340	133,007

* p<0.1, ** p<0.05, *** p<0.001.