NBER WORKING PAPER SERIES

INCOMPLETE INFORMATION AND SELF-FULFILLING INFLATION DYNAMICS —LUCAS MEETS KEYNES

Jess Benhabib
Pengfei Wang
Yi Wen

Working Paper 35563
http://www.nber.org/papers/w35563

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2026

We would like to thank Magali Marx, Bob King, and participants at the Informational Frictions in Macroeconomics Conference in honor of Robert E. Lucas Jr. at Gerzensee, Switzerland, for comments. Pengfei Wang acknowledges financial support from National Science Foundation of China (project No.72125007) and from Shenzhen Municipal Key Research Base for Humanities and Social Sciences. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w35563

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Jess Benhabib, Pengfei Wang, and Yi Wen. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Incomplete Information and Self-Fulfilling Inflation Dynamics-Lucas meets Keynes
Jess Benhabib, Pengfei Wang, and Yi Wen
NBER Working Paper No. 35563
July 2026
JEL No. D8, D84, E03, E32

ABSTRACT

This paper incorporates the Lucas (1973) island model into a standard DSGE framework. It demonstrates that self-fulfilling stochastic inflation equilibria, which are driven by intrinsic uncertainty or pure sentiments, can exist under rational expectations. Furthermore, it shows that these sentiment-driven stochastic equilibria exhibit monetary non-neutrality, even when the fundamental equilibrium is unique and monetarily neutral. As the aggregate price or inflation rate can appear to fluctuate independently of the money supply, our model explains the complex relationship between inflation and the money supply in the real world, where the basic quantity theory of money often seems invalid.

Jess Benhabib
New York University
Department of Economics
and NBER
jess.benhabib@nyu.edu

Pengfei Wang
Peking University
pfwang@phbs.pku.edu.cn

Yi Wen
Shanghai Jiaotong University
Antai College of Economics
and Management
Department of Economics
wenyi001@sjtu.edu.cn

A online appendix is available at http://www.nber.org/data-appendix/w35563

1 Introduction

Lucas (1972, 1973) is a landmark in the macroeconomic literature for understanding the economic mechanism of output-inflation trade-offs within a rational expectations general equilibrium framework. While Keynes (1936) emphasized market coordination failures or animal spirits in the process of expectation formation (via forecasting the forecasts of others) as a fundamental source of monetary non-neutrality, and argued that firms' expectations of aggregate demand could become self-fulfilling when reinforced by iterative speculation, Lucas (1973) emphasized signal extraction under imperfect information as the source of the output–inflation trade-offs behind the Phillips curve.

In the Lucas Island model, the driving force is money shocks. The transmission mechanism consists of competitive markets, rational expectations and imperfect information. This means that firms have better information about their own market than they do about the economy as a whole.

This paper integrates the Lucas island model into a standard DSGE framework with imperfect competition (as described by Dixit and Stiglitz). Our analysis yields three key findings. Firstly, we demonstrate that the Lucas model itself can exhibit multiple fundamental equilibria when the variance of the aggregate price is treated as endogenous. Secondly, in our DSGE extension, we show that self-fulfilling stochastic aggregate price equilibria can arise under rational expectations, even when the fundamental equilibrium is unique, driven by intrinsic uncertainty or pure sentiments rather than randomization over multiple fundamental equilibria. Thirdly, these sentiment-driven equilibria exhibit monetary non-neutrality; furthermore, they permit the aggregate price to fluctuate independently of the money supply. Importantly, these equilibria represent a new category: unlike sunspots in the indeterminacy literature (Benhabib and Farmer, 1994), which depend on randomization over multiple certainty equilibria, our sentiment equilibria emerge endogenously, even when the certainty equilibrium is unique, thanks to the Lucas mechanism of signal extraction under imperfect information.

The mechanism behind self-fulfilling price (or inflation) dynamics lies in market decentralization and information frictions. In such a system, firms produce goods and services that they do not consume, while households consume goods and services that they do not produce. Consequently, firms must form expectations regarding aggregate demand and real labor costs, while households must form expectations regarding aggregate income. Forecasting the aggregate price level (or the inflation rate) is therefore important for all decision-makers in the economy, including the monetary authority. The separation of supply and demand in a decentralized market, along with the mutual expectations involved in decision-making, leads to a self-fulfilling equilibrium of rational expectations under imperfect information.

While the literature often discusses the microfoundations of macroeconomics, it rarely addresses

the macrofoundations of microeconomics, as emphasized by Keynes (1936). In reality, the outcome of individual agents' optimal plans often depends crucially on macroeconomic conditions (such as the rate of inflation or the aggregate price level), which agents can influence collectively, but not individually. If agents' optimal decisions are based on their expectations of such aggregate conditions, these expectations may become self-fulfilling if they are synchronized. Our analysis in this paper therefore provides a natural basis for reconciling the insights of Keynes and Lucas.[1]

While our model is a Lucas (1973)-type period by period static setting and thus very stylized, it yields several predictions consistent with existing literature. First, inflation expectations play a significant role as a driver of inflation dynamics in our framework. This aligns with recent empirical work on sentiment shocks—shocks to expectations. For example, Lagerborg, Pappa, and Ravn (2023) exploit data on mass shootings to demonstrate that sentiment shocks induce significant fluctuations in consumption and real output. Similarly, Benhabib and Spiegel (2019) show that state-level consumer sentiment shocks, instrumented with political outcomes, have measurable effects on state output and consumption.[2]

With regard to inflation expectations specifically, recent literature — as summarized by D'Acunto, Malmendier and Weber (2022) — finds that inflation expectations are an important determinant of actual inflation and are highly dispersed among different market participants. Empirical studies using survey data provide supportive evidence. For example, Leduc, Sill, and Stark (2007) document self-fulfilling inflationary expectations prior to 1979, while Ascari, Fasani, Grazzini, and Rossi (2023) demonstrate that an increase in short-term inflationary expectations leads to higher inflation and lower output, which is consistent with the predictions of our model.

The remainder of the paper is structured as follows: Section 2 revisits Lucas's (1973) island model, demonstrating that uncertainty in the aggregate money supply results in multiple fundamental equilibria. Section 3 provides a microfoundation for the Lucas model by extending it to a standard DSGE framework with monopolistic competition (à la Dixit–Stiglitz). We show that the model exhibits multiple fundamental and stochastic self-fulfilling equilibria in the aggregate price level. Unlike the Lucas model, these self-fulfilling equilibria are driven by intrinsic uncertainty alone, independent of policy uncertainty and without relying on randomization over multiple fun-

[1] This paper is related to our previous paper (Benhabib, Wang, and Wen, BWW 2015) with one important difference. While BWW (2015) uses a real model to focus on self-fulfilling fluctuations in aggregate output, this paper uses a Lucas-type monetary model to study self-fulfilling fluctuations in the aggregate price level (or inflation rate).

[2] Throughout this paper, the terms 'aggregate price level' and 'inflation' are used interchangeably. Our benchmark model is static and describes equilibrium prices in a single period. By 'inflation dynamics', we refer to the fluctuations in the aggregate price level from one period to the next in an environment where each period's equilibrium is determined by the same static signal-extraction problem. To keep the paper concise, we do not model intertemporal price-setting frictions or forward-looking Phillips curve dynamics. Instead, we focus on the coordination mechanism that can cause aggregate price fluctuations to become self-fulfilling, even in the absence of fundamental shocks. This static framework is sufficient to demonstrate the existence of sentiment-driven equilibria while contrasting with the Lucas (1973) model, and remains tractable for analytical characterisation.

damental equilibria. Section 4 discusses the robustness of our results by studying variations of our benchmark model and briefly reviews related work. Section 5 concludes.

2 Lucas (1973) Model Revisited

This section re-examines the Lucas (1973) model to demonstrate that when information is imperfect, extracting signals under policy uncertainty regarding the money supply can result in monetary non-neutrality and multiple fundamental equilibria in the mappings between aggregate output and the total money supply. Lucas (1973) did not identify these multiple equilibrium mappings.

In the Lucas (1973) model, where all variables are expressed in log terms, the output supply in island $j \in [0,1]$ is determined by its relative price $p_{jt} - \mathbf{p}_t = \varepsilon_{jt}$:

$$y_{jt} = \gamma \mathbb{E}[(p_{jt} - \mathbf{p}_t) \,|s_{jt})], \tag{1}$$

where $\gamma > 0$ measures the elasticity of supply, $\mathbf{p}_t = \int p_{jt} dj$ is the aggregate price in the entire economy, ε_{jt} is an island-specific supply shock, and s_{jt} is a signal received by firms in island j about the aggregate price $\mathbf{p}_t$. As price-takers, firms in each island produce output according to the supply equation (1).

Lucas (1973) assumes that firms on each island j cannot perfectly observe the aggregate price $\mathbf{p}_t$ and the island-specific shock ε_{jt} (with zero mean and variance σ_ε^2), so the signal s_j is a mixture of the aggregate price and idiosyncratic noise:

$$s_{jt} (= p_{jt}) = \mathbf{p}_t + \varepsilon_{jt}. \tag{2}$$

Therefore, optimal output supply y_{jt} involves signal extraction under imperfect information.

According to the quantity theory of money, the aggregate price and aggregate output have the following relationship:[3]

$$\mathbf{p}_t + \mathbf{y}_t = \mathbf{m}_t, \tag{3}$$

where $\mathbf{m}_t$ is the money stock.

If $\mathbf{m}_t$ is known, then signal extraction implies

$$y_{jt} = \gamma \mathbb{E}[\varepsilon_{jt} | (\mathbf{m}_t - \mathbf{y}_t + \varepsilon_{jt})] = \gamma \frac{\sigma_\varepsilon^2}{\sigma_{\mathbf{y}}^2 + \sigma_\varepsilon^2} (-\mathbf{y}_t + \varepsilon_{jt}), \tag{4}$$

where σ_y^2 denotes the variance of $\mathbf{y}_t$ and the second equation holds because $\mathbf{m}_t$ is uncorrelated with ε_{jt}. Since $\mathbf{y}_t = \int y_{jt} dj$, summing over j on both sides of the equation yields

$$\mathbf{y}_t = -\gamma \frac{\sigma_\varepsilon^2}{\sigma_y^2 + \sigma_\varepsilon^2} \mathbf{y}_t. \tag{5}$$

[3] This means that money is introduced into the model as a unit of account.

Since $\gamma > 0$, the only possible equilibrium is $\mathbf{y}_t = 0$, so $\sigma_y^2 = 0$. In this equilibrium, money is neutral and $\mathbf{p}_t = \mathbf{m}_t$. Even if information is perfect (or $\varepsilon_{jt} = 0$), the equilibrium is also trivially given by $y_{jt} = 0$ and $\mathbf{y}_t = \int y_{jt} dj = 0$.

However, if there is aggregate uncertainty in the money supply ($\sigma_m^2 > 0$), then money is no longer neutral. Following Lucas (1973), we conjecture the equilibrium solution: $\mathbf{y}_t = \alpha \mathbf{m}_t$, where the coefficient α is yet to be determined in equilibrium. The signal then becomes $s_{jt} = \mathbf{m}_t - \mathbf{y}_t + \varepsilon_{jt} = (1-\alpha)\mathbf{m}_t + \varepsilon_{jt}$; therefore, firm's output supply becomes

$$y_{jt} = \gamma \mathbb{E}[\varepsilon_{jt} | s_{jt}] = \gamma \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (1-\alpha)^2 \sigma_m^2} [(1-\alpha)\mathbf{m}_t + \varepsilon_{jt})]. \tag{6}$$

Integration over j leads to $\mathbf{y}_t = \gamma \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (1-\alpha)^2 \sigma_m^2} (1-\alpha)\mathbf{m}_t$. Since $\mathbf{y}_t = \alpha \mathbf{m}_t$, the above equation holds true if and only if

$$\alpha = \frac{\gamma \sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (1-\alpha)^2 \sigma_m^2} (1-\alpha). \tag{7}$$

Lucas (1973) assumes that $(1-\alpha)^2 \sigma_m^2$ is exogenously given. Under this assumption and with the notation $\omega \equiv \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (1-\alpha)^2 \sigma_m^2}$, we can solve for α and obtain the well-known result from Lucas (1973): $\mathbf{y}_t = \frac{\gamma\omega}{1+\gamma\omega} \mathbf{m}_t$ and $\mathbf{p}_t = \frac{1}{1+\gamma\omega} \mathbf{m}_t$. Money is thus non-neutral: a 1 percent increase in $\mathbf{m}_t$ can lead to a $\alpha = \frac{\gamma\omega}{1+\gamma\omega}$ percent increase in aggregate real output $\mathbf{y}_t$ and a less than proportional change in aggregate price $\mathbf{p}_t$. Therefore, aggregate price appears to be "sticky".

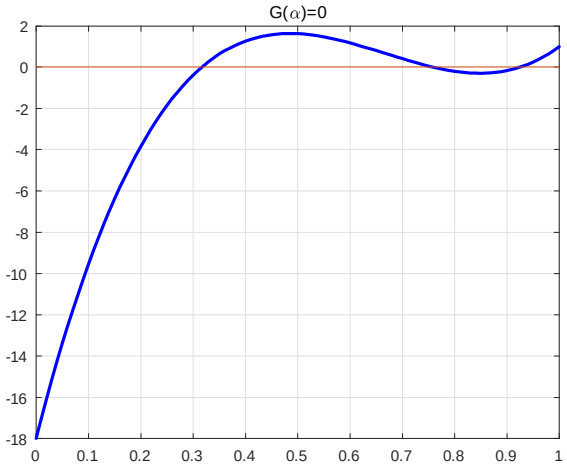


Figure 1. $G(\alpha) = 0$ has three roots.

However, Lucas (1973) treats the variance of $\mathbf{p}_t$, and therefore the parameter ω, as a given. Given this, the mapping between the money supply and real output is uniquely determined. However equation (7) is a cubic equation because the variance of price $\mathbf{p}_t$ $(= (1-\alpha)\mathbf{m}_t)$ is given by $(1-\alpha)^2\sigma_m^2$, which also depends on the undetermined coefficient, α. In general, there is no guarantee that the mapping α is unique.

To see this, denote $\tau = \sigma_m^2/\sigma_\varepsilon^2$, equation (7) then becomes

$$G(\alpha) = \left(1 + (1-\alpha)^2\tau\right)\alpha - \gamma(1-\alpha) = 0. \tag{8}$$

Notice that $G(0) = -\gamma < 0$ and $G(1) = 1 > 0$, so by the intermediate value theorem, there exists at least one solution for α. Differentiating the above equation with respect to α gives $G'(\alpha) = 1 + \gamma + \tau(1-\alpha)(1-3\alpha)$. Depending on the values of α and τ, $G'(\alpha)$ may not be monotonic. For example, if $\tau > 3(1+\gamma)$, then G is initially increasing and then decreasing in α. This means that it is possible to have more than one solution to $G(\alpha) = 0$. For instance, if $\gamma = 18$, $\sigma_m = 1$, and $\sigma_\varepsilon = 1/9$, then there are three solutions for α: $\alpha_1 < \frac{1}{3}$, $\frac{1}{3} < \alpha_2 < \frac{9}{10}$, and $\frac{9}{10} < \alpha_3 < 1$ (see Figure 1).

In summary, the Lucas (1973) model has the potential to exhibit three fundamental equilibria in the presence of aggregate uncertainty in the money supply. This suggests that monetary policy uncertainty itself may be a source of aggregate fluctuations, even within the simple Lucas (1973) model.[4]

3 A Micro-founded General Equilibrium Model

In this section, we apply the insights of the Lucas model to a monopolistic competition setting, in which firms set nominal prices rather than accepting them. We demonstrate that, when faced with imperfect information, the best response function of firms to the aggregate state exhibits a positive coefficient. In contrast, firms' best response function in the Lucas model features a negative coefficient (see equation 4). Consequently, in our model, signal extraction can give rise not only to multiple fundamental equilibria (as in the Lucas model), but also to equilibria driven by sentiment or intrinsic uncertainty. In a sentiment-driven equilibrium, the aggregate price (or inflation rate) can fluctuate in the absence of fundamental shocks, such as monetary shocks, even though the fundamental equilibrium is unique.

A representative household consists of a continuum of members $j \in [0,1]$, each member j supplies labor only to island j. Each Island has a monopolistic intermediate goods producer owned

[4] We can randomize over the three equilibria with sunspots. However, our aim in this paper is to demonstrate that sunspot-driven inflation dynamics can arise endogenously, even when the fundamental equilibrium is unique (as in Cass and Shell, 1983). Additionally, we can study (see Section 4 below) which of the three equilibria and the value of α can be learnt under adaptive constant gains learning. Such learning can induce additional persistent dynamics.

by the representative household, which produces a single differentiated good, Y_{jt}. The price of Y_{jt} is denoted by P_{jt}. These intermediate goods are used as inputs to produce a final good according to the technology:

$$Y_t = \left(\int_0^1 Y_{jt}^{\frac{\sigma-1}{\sigma}}(i)di \right)^{\frac{\sigma}{\sigma-1}}, \tag{9}$$

where $\sigma > 1$ measures the elasticity of substitution among intermediate goods. The final goods industry and the labor markets are assumed to be perfectly competitive.

The representative household gains utility from aggregate consumption, but experiences disutility for each household member who works on a different island. This is represented as[5]

$$u(C_t, N_{jt}) = \frac{C_t^{1-\gamma}}{1-\gamma} - \int_0^1 \epsilon_{jt} N_{jt} dj. \tag{10}$$

The household budget constraint is given by $C_t \leq \int \left(\frac{W_{jt}}{P_t} N_{jt} \right) dj + D_t$, where C_t denotes aggregate consumption, W_{jt} is the nominal wage on island j, N_{jt} is the labor supply to island j, ϵ_{jt} is an idiosyncratic cost shock to the labor supply on island j, P_t is the aggregate price of the final good, and D_t denotes aggregate profit income. For simplicity and analytical tractability, we assume that ϵ_{jt} follows a log-normal distribution.

The first-order condition for labor supply is

$$W_{jt} = \epsilon_{jt} P_t C_t^{\gamma}. \tag{11}$$

Note that the local nominal wage on each island is affected by both the aggregate price P_t and the aggregate demand C_t, as well as the island-specific disutility shock ϵ_{jt}. We follow Lucas (1973) by assuming that the simple quantity equation of money holds: $P_t C_t = M_t$, where M represents the aggregate monetary supply.

Now consider firms' price setting behavior. The final goods industry is assumed to be perfectly competitive. Profit maximization by the final goods producer yields the demand function for intermediate goods:

$$Y_{jt} = \left(\frac{P_{jt}}{P_t} \right)^{-\sigma} Y_t. \tag{12}$$

[5] We can relax the linearity assumption in leisure by making it strictly concave, or by adding an upper bound (kink) to labor supply that does not constrain our solution. This would rule out unbounded labor and output.

Substituting the demand function into the final goods production function (9) gives the aggregate price index,

$$P_t = \left(\int_0^1 {P_{jt}}^{1-\sigma} di \right)^{1/(1-\sigma)}. \tag{13}$$

For simplicity, the production technology for intermediate goods is given by

$$Y_{jt} = N_{jt}. \tag{14}$$

An intermediate goods producer has monopoly power in the output market but faces perfect competition in the factor market. Given the production technology, the cost function of an intermediate goods firm is simply given by $\frac{W_{jt}}{P_t} Y_{jt}$, meaning that the firm's nominal price is set to maximize its expected real profit:

$$\max_{P_{jt}} \mathbb{E} \left\{ \left[\frac{P_{jt}}{P_t} Y_{jt} - \frac{W_{jt}}{P_t} Y_{jt} \right] \middle| \, s_{jt} \right\}, \tag{15}$$

conditioned on the information set s_{jt} (to be specified shortly), while the demand function (12) and the joint distribution of $\{W_{jt},\, P_t,\, M_t,\, Y_t\}$ are taken as given.

The first-order condition for optimal price setting yields

$$P_{jt} = \frac{\sigma}{\sigma - 1} \frac{\mathbb{E} \left[W_{jt} \left(\frac{1}{P_t} \right)^{1-\sigma} Y_t \middle| \, s_{jt} \right]}{\mathbb{E} \left[\left(\frac{1}{P_t} \right)^{1-\sigma} Y_t \middle| \, s_{jt} \right]}. \tag{16}$$

A key feature of the model is that intermediate goods firms set prices simultaneously, based on the prices set by other firms. The equilibrium quantities (including the aggregate demand and price) are then determined at these prices.

We will demonstrate that, under imperfect information, this price-setting game allows for self-fulfilling expectations because intermediate goods firms must set profit-maximizing prices without knowing the aggregate price or the resulting equilibrium real wage W_{jt}/P_t that may prevail in the market. However, the aggregate economic variables (such as P_t or C_t) that determine the marginal cost $W_{jt} = \epsilon_{jt} P_t C_t^{\gamma}$ (or the real wage W_{jt}/P_t) depend crucially on the actions of households, other firms, and the monetary authority, none of which an individual firm can control. The possibility of a self-fulfilling rational expectations equilibrium in this class of models has largely been overlooked by existing literature, as this literature implicitly assumes that firms can perfectly anticipate the equilibrium real wage W_{jt}/P_t and aggregate demand ($C_t = \frac{M_t}{P_t}$) when setting prices at the firm level.

3.1 Perfect Information

First, we demonstrate the conventional result in the existing literature by assuming that the factors affecting the local marginal cost W_{jt} are perfectly observable for each firm j. Then, the price equation (16) becomes

$$P_{jt} = \frac{\sigma}{\sigma - 1} W_{jt} = \frac{\sigma}{\sigma - 1} P_t C_t^{\gamma} \epsilon_{jt}. \tag{17}$$

Equivalently, this can be written as $\frac{P_{jt}}{P_t} = \frac{\sigma}{\sigma-1} C_t^{\gamma} \epsilon_{jt}$, which suggests that firms set nominal (or real) prices as a constant markup over nominal (or real) marginal costs. Applying the aggregate price equation (13), aggregate consumption can be solved as

$$C = \left[\frac{\sigma}{\sigma - 1} \left(\int \epsilon_{jt}^{1-\sigma} dj \right)^{\frac{1}{1-\sigma}} \right]^{\frac{1}{\gamma}} = \left[\frac{\sigma}{\sigma - 1} \exp\left(\frac{1}{2} (1 - \sigma) \sigma_{\varepsilon}^2 \right) \right]^{\frac{1}{\gamma}}, \tag{18}$$

where the second equality follows from the log-normal distribution assumption of ϵ_{jt} under the notation $\varepsilon_{jt} \equiv \log \epsilon_{jt}$.

Note that, in equilibrium, the aggregate consumption C_t is constant and depends only on the variance σ_{ε}^2 of the idiosyncratic cost shock ε_{jt}. The aggregate price level can then be solved as $P_t = \frac{M_t}{C}$. This suggests that money is neutral because consumption is constant.

Under the assumption of perfect information, the log price of firm j is given by

$$p_{jt} = \log P_{jt} = \log \frac{\sigma}{\sigma - 1} + \log P_t + \gamma \log C_t + \varepsilon_{jt}. \tag{19}$$

Since ε_{jt} is assumed to be normally distributed, p_{jt} is also normally distributed. Using the properties of the log-normal distribution, we can express the aggregate price in log terms as

$$\mathbf{p}_t \equiv \log P_t = \frac{1}{1 - \sigma} \log \left(\int_0^1 {P_{jt}}^{1-\sigma} di \right) = \int_0^1 p_{jt} dj + \frac{(1 - \sigma) \sigma_{\varepsilon}^2}{2}, \tag{20}$$

where $p_{jt} \equiv \log P_{jt}$.

To make the full-fledged model comparable to the Lucas model, we define a lowercase variable $\hat{x}_t \equiv \log (X_t / X)$ as the deviation of a variable X_t form its deterministic steady state X in logarithmic terms. Under this notation, equation (19) becomes $\hat{p}_{jt} = \hat{\mathbf{p}}_t + \gamma \hat{\mathbf{c}}_t + \varepsilon_{jt}$, and the aggregate price in equation (20) becomes $\hat{\mathbf{p}}_t = \int_0^1 \hat{p}_{jt} dj$. Finally, following Lucas (1973), the quantity equation of money is given by $\hat{\mathbf{p}}_t + \hat{\mathbf{c}}_t = \hat{\mathbf{m}}_t$. Therefore, equation (19) becomes

$$\hat{p}_{jt} = (1 - \gamma) \hat{\mathbf{p}}_t + \gamma \hat{\mathbf{m}}_t + \varepsilon_{jt}. \tag{21}$$

It is also straightforward to verify that $\hat{\mathbf{p}}_t = \hat{\mathbf{m}}_t$ and $\hat{\mathbf{c}} = 0$, and that this fundamental equilibrium is unique.

3.2 Imperfect Information

However, there is no a priori guarantee that all firms can perfectly anticipate each other's actions and the general equilibrium outcome of the economy, as emphasized by Keynes (1936). If firms must form expectations about the aggregate equilibrium outcome or try to forecast the forecasts of others, multiple self-fulfilling rational expectations equilibria are possible when firms' information about the aggregate economy is incomplete or imperfect. This means that market participants' beliefs about the state of the economy may reinforce each other and become self-fulfilling.

For simplicity, let us assume that the money supply is constant for the time being, meaning that aggregate demand and aggregate price satisfy the quantity equation: $\hat{\mathbf{p}}_t + \hat{\mathbf{c}}_t = 0$. Therefore, there are only two components that determine the marginal cost W_{jt}: the aggregate price $\hat{\mathbf{p}}_t$ and the idiosyncratic labor supply shock ε_{jt}.

Furthermore, we assume that, before the local labor market clears and before the aggregate price and aggregate demand is realized, firms set nominal prices to maximize expected profits, committing to produce according to expected demand at those prices. In retrospect, the prices may be set too low or too high, but on average, they are correct in the sense of rational expectations, analogous to Calvo's sticky price model.[6]

In other words, firms cannot perfectly observe their marginal cost W_{jt}, nor can they distinguish the aggregate price $\hat{\mathbf{p}}_t$ and the idiosyncratic labor cost shock ε_{jt} that underlie the marginal cost. Instead, each firm j receives a noisy signal,

$$s_{jt} = \mu \hat{\mathbf{p}}_t + (1-\mu)\,\varepsilon_{jt}, \tag{22}$$

about the components in its marginal cost, where $\mu \in [0,1]$ is a parameter determining the relative weight of the components.[7]

What would be the source of the noise ε_{jt} that affect the firm's observation or inference of the marginal cost or the aggregate and idiosyncratic components behind it? These noise signals could come from firms' market surveys of customers regarding their willingness of labor supply or aggregate consumption demand. They could also come from people's expectations regarding the Fed's inflation target or the money supply. Of course, such surveys are also subject to sampling noise themselves.

[6] In the Lucas (1973) model, firms observe their market price, which may be high either because aggregate demand is high or because demand for their own product is high. This leads to a signal extraction process through which firms obtain an estimate of the aggregate price. In our Calvo-type setup, firms set nominal prices and commit to delivering at any subsequent realization of marginal costs. These costs may be high because the money supply or aggregate demand is high, or because the local labor market is tight. This leads to a signal extraction process, enabling firms to obtain their estimate of real marginal costs. Firms adhere to their commitment on a period-by-period basis. For example, in a similar spirit and by invoking menu costs in adjusting prices, Obstfeld and Rogoff (1996, page 674, section 10.1.7.1) also argue that under monopolistic and marginal cost pricing, it will still be profitable to meet demand if the shocks are small, since profits will still be positive. They state that "if a small unanticipated shock raises demand, it will be profitable for a producer to raise output, even if they cannot change their price."

[7] The signal structure is made more general by the introduction of this weight parameter μ into the signal specified by Lucas (1973). In what follows the meaning of μ will be discussed.

Then, under a log-normal distribution of ϵ_{jt}, the price-setting equation (16) in log terms becomes[8]

$$\begin{aligned} \hat{p}_{jt} &= \mathbb{E}\left[(\hat{\mathbf{p}}_t + \gamma\hat{\mathbf{c}}_t + \varepsilon_{jt}) \,|\, (\mu\hat{\mathbf{p}}_t + (1-\mu)\varepsilon_{jt})\right] \\ &= \mathbb{E}\left[(\beta\hat{\mathbf{p}}_t + \varepsilon_{jt}) \,|\, (\mu\hat{\mathbf{p}}_t + (1-\mu)\varepsilon_{jt})\right], \end{aligned} \tag{23}$$

where $\beta = 1 - \gamma \in (-\infty, 1)$. We also have that $\hat{\mathbf{p}}_t = \int_0^1 \hat{p}_{jt} dj$.

Sentiment-driven Equilibria. In the absence of monetary uncertainty, suppose that the aggregate price is affected by a self-fulfilling belief originating from the household side, unrelated to any fundamentals. We refer to this belief as "consumer sentiment" and denote it as $\mathbf{z}_t$. So

$$\hat{\mathbf{p}}_t = \mathbf{z}_t. \tag{24}$$

The variable z_t captures the common belief among households regarding the aggregate price level — a 'consumer sentiment' which influences labor supply decisions and, via the labor market, firms' marginal costs. Unlike fundamental (e.g. productivity) or policy (e.g. monetary) shocks, z_t has no independent existence outside of agents' beliefs. It is similar to Keynes's 'animal spirits' or the 'forecasts of others' that agents must make in a decentralized economy. The idea that beliefs can influence aggregate outcomes is empirically relevant: recent evidence summarized by D'Acunto, Malmendier and Weber (2022) suggests that inflation expectations are an important factor in determining actual inflation.

Crucially, z_t differs from a conventional sunspot variable in the indeterminacy literature following Benhabib and Farmer (1994) in that sunspots typically randomize over multiple fundamental equilibria, whereas our sentiment equilibrium exists even when the fundamental equilibrium is unique. Instead, z_t emerges as intrinsic uncertainty—a pure belief—that becomes self-fulfilling because firms' pricing decisions complement their beliefs about the aggregate price. As shown below, the variance σ_z^2 is not imposed externally, but rather determined internally by the equilibrium conditions, ensuring consistency between beliefs and realizations of the aggregate price.

Consumers have perfect information about $\mathbf{z}_t$, but firms do not. Firms' signals are imperfect, as shown in equation (22). We will prove that this sentiment variable $\mathbf{z}_t$ not only constitutes a self-fulfilling rational expectations equilibrium, but is also endogenously stochastic with a dynamic life of its own.

Given equation (24), firm's price setting equation becomes

$$\hat{p}_{jt} = \Omega_1 \times [\mu\mathbf{z}_t + (1-\mu)\varepsilon_{jt}], \tag{25}$$

[8] Taking log on both sides of equation (16) and applying the rule,

$$\log \mathbb{E}x_t = \log e^{\mathbb{E}\log x_t + \frac{1}{2}\sigma_x^2} = \mathbb{E}\log x_t + \frac{1}{2}\sigma_x^2,$$

gives the result (note $W_{jt} = \frac{\sigma}{\sigma-1} P_t C_t^{\gamma} \epsilon_{jt}$).

where $\Omega_1 \equiv \frac{\beta\mu\sigma_z^2+(1-\mu)\sigma_\varepsilon^2}{\mu^2\sigma_z^2+(1-\mu)^2\sigma_\varepsilon^2}$. Aggregation on both sides of the equation yields

$$\hat{\mathbf{p}}_t = \Omega_1\mu\mathbf{z}_t. \tag{26}$$

For this equation to be consistent with equation (24), it requires $\Omega_1\mu = 1$, or

$$\sigma_z^2 = \frac{(1-\mu)(2\mu-1)}{(1-\beta)\mu^2}\sigma_\varepsilon^2. \tag{27}$$

Note that when $\mu \in (1/2, 1)$, we have $\sigma_z^2 > 0$, meaning that the aggregate price $\hat{\mathbf{p}}_t = \mathbf{z}_t$ is self-fulfilling and has a life of its own, completely dictated by consumer sentiment $\mathbf{z}_t$. In other words, a sentiment equilibrium driven by intrinsic uncertainty alone can exist even if the fundamental equilibrium is unique when $\sigma_m^2 = 0$.

The intuition is as follows. If μ is too small ($\mu \leq 1/2$ or $\mu = 0$), firms' signals across islands are insufficiently complementary with respect to the influence of consumer sentiment. As a result, a sentiment-driven equilibrium does not exist ($\sigma_z^2 \leq 0$). Conversely, if there is no noise ($\mu = 1$), the information about the aggregate price is perfect and a sentiment-driven equilibrium cannot exist either. Therefore, μ must lie within a certain range to ensure a sentiment-driven equilibrium. Of course, when $\mu = 1/2$, it corresponds to the signal structure originally specified in Lucas (1973), and there is no sentiment-driven equilibrium, but this is just a very special knife-edge case.[9]

Multiple Fundamental Equilibria. Now, suppose that the money supply is a stochastic variable, as described by Lucas (1973): $\hat{\mathbf{m}}_t = \log M_t/M \sim \mathcal{N}(0, \sigma_m^2)$. For simplicity, assume that $\hat{\mathbf{m}}_t$ is an independently and identically distributed (iid) process. According to the quantity equation, we then have:

$$\hat{\mathbf{c}}_t = \hat{\mathbf{m}}_t - \hat{\mathbf{p}}_t. \tag{28}$$

Under imperfect information, equation (23) becomes

$$\hat{p}_{jt} = \mathbb{E}\left[(\beta\hat{\mathbf{p}}_t + (1-\beta)\,\hat{\mathbf{m}}_t + \varepsilon_{jt}) \,|\, (\mu\hat{\mathbf{p}}_t + (1-\mu)\varepsilon_{jt})\right]. \tag{29}$$

First, let us focus on the equilibrium without sentiment or intrinsic uncertainty. Following Lucas (1973), we can solve the model's equilibrium by hypothesizing that the aggregate price is a function of the money supply only: $\hat{\mathbf{p}}_t = \phi\hat{\mathbf{m}}_t$. Under this conjecture, we have

$$\hat{p}_{jt} = \Omega_2 \times [\mu\phi\hat{\mathbf{m}}_t + (1-\mu)\varepsilon_{jt}], \tag{30}$$

where $\Omega_2 \equiv \frac{(\beta\phi+1-\beta)\mu\phi\sigma_m^2+\sigma_\varepsilon^2(1-\mu)}{\mu^2\phi^2\sigma_m^2+(1-\mu)^2\sigma_\varepsilon^2}$. Aggregation on both sides of the above equation yields

$$\hat{\mathbf{p}}_t (= \phi\hat{\mathbf{m}}_t) = \Omega_2\mu\phi\hat{\mathbf{m}}_t. \tag{31}$$

[9] See the next section for discussions on why the Lucas model does not have a sentiment-driven equilibrium, regardless of μ.

Note that $\phi = 0$ is always a consistent solution to the above equation. This case corresponds to $\mu = 0$ in the signal structure $s_{jt} = \varepsilon_{jt}$. In this scenario, perfect price stickiness emerges as the equilibrium outcome—when all firms believe that the aggregate price does not respond to changes in money, this belief can become self-fulfilling because the signal is completely uninformative about the aggregate price and uncorrelated across firms. Consequently, each firm sets its price as $\hat{p}_{jt} = \varepsilon_{jt}$. The aggregate price is therefore $\hat{\mathbf{p}}_t = \int \varepsilon_{jt} dj = 0$. In this scenario, money is not neutral, since aggregate output is given by $\hat{\mathbf{y}}_t = \hat{\mathbf{m}}_t$.[10]

However, other possibilities arise as long as the following condition holds: $\Omega_2 \mu = 1$, or

$$\phi^2 - \phi = \frac{(1-\mu)(2\mu-1)}{(1-\beta)\,\mu^2} \frac{\sigma_\varepsilon^2}{\sigma_m^2}, \tag{32}$$

which yields the solution:

$$\phi = \frac{1}{2} \pm \frac{1}{2} \sqrt{1 - 4 \frac{(1-\mu)(1-2\mu)}{(1-\beta)\,\mu^2} \frac{\sigma_\varepsilon^2}{\sigma_m^2}}. \tag{33}$$

If $\mu < \frac{1}{2}$ and $4 \frac{(1-\mu)(1-2\mu)}{(1-\beta)\mu^2} \frac{\sigma_\varepsilon^2}{\sigma_m^2} < 1$, then ϕ has two positive solutions with $0 < \phi_1 < \phi_2 < 1$, both of which imply monetary non-neutrality. Therefore, as in the Lucas (1973) model, there potentially exist three fundamental equilibria with monetary non-neutrality.

Continuum of Sentiment Equilibria. Now suppose that the aggregate price is affected not only by the money supply, but also by consumer sentiment $\mathbf{z}_t$. That is, the relationship between aggregate price and the aggregate state takes the form:

$$\hat{\mathbf{p}}_t = \phi \hat{\mathbf{m}}_t + \mathbf{z}_t. \tag{34}$$

Given equation (34), firm's price setting equation 29 becomes

$$\hat{p}_{jt} = \Omega_3 \times [\mu\phi \hat{\mathbf{m}}_t + \mu \mathbf{z}_t + (1-\mu)\varepsilon_{jt}], \tag{35}$$

where $\Omega_3 \equiv \frac{(\beta\phi+1-\beta)\mu\phi\sigma_m^2 + \beta\mu\sigma_z^2 + \sigma_\varepsilon^2(1-\mu)}{\mu^2\phi^2\sigma_m^2 + \mu^2\sigma_z^2 + (1-\mu)^2\sigma_\varepsilon^2}$. Aggregation on both sides of the equation yields

$$\hat{\mathbf{p}}_t = \Omega_3 \mu \left(\phi \hat{\mathbf{m}}_t + \mathbf{z}_t\right). \tag{36}$$

For this equation to be consistent with equation (34), it requires that $\Omega_3 \mu = 1$, or

$$\sigma_z^2 = \phi(1-\phi)\sigma_m^2 + \frac{(1-\mu)(2\mu-1)}{(1-\beta)\mu^2} \sigma_\varepsilon^2. \tag{37}$$

[10] Note that the possibility of $\phi = 0$ does not arise in the Lucas (1973) model, where the corresponding parameter for ϕ is given by $(1-\alpha) = \frac{1}{1+\gamma\omega}$, which cannot be zero by assumption unless the exogenous supply elasticity parameter γ is infinite. However, in our model, the supply elasticity is endogenously determined.

First, when $\sigma_m^2 = 0$, we are back to the benchmark scenario discussed previously, where $\mu \in (1/2, 1)$ is required to have a sentiment-driven equilibrium. Second, note that σ_z^2 and ϕ are both endogenous to the equilibrium. Therefore, the above equilibrium condition (37) defines a continuum of sentiment-driven equilibria in the volatility σ_z^2 indexed by ϕ. That is, for every possible value of ϕ such that the right-hand side of equation (37) is positive, there is a corresponding stochastic equilibrium path of the sentiment z_t with variance σ_z^2. Third, equation (37) shows that aggregate shocks to fundamentals, such as the variance of the money supply σ_m^2, amplify the volatility of the sentiment for any given value of $\phi \in (0, 1)$ and relax the requirement $\mu > 1/2$.

Interestingly, sentiment-driven equilibria can still emerge even when firms have perfect information about the aggregate price $\hat{\mathbf{p}}_t$: Equation (37) indicates that $\sigma_z^2 = \phi(1-\phi)\sigma_m^2$ when $\mu = 1$ such that the signal is $s_{jt} = \hat{\mathbf{p}}_t$. This is because when there is aggregate uncertainty in the money supply that can shift the aggregate price, the socially coordinated belief (sentiment) about the aggregate price can also become self-fulfilling, even if firms observe $\hat{\mathbf{p}}_t = \mathbf{z}_t$ perfectly. Finally, money is not neutral in a sentiment-driven equilibrium when $\hat{\mathbf{m}}_t \neq 0$.

To see this, note that aggregate prices and output are given by $\hat{\mathbf{p}}_t = \phi\hat{\mathbf{m}}_t + \mathbf{z}_t$, $\hat{\mathbf{y}}_t = (1-\phi)\,\hat{\mathbf{m}}_t - \mathbf{z}_t$, respectively. Thus, a monetary shock raises both aggregate output and the aggregate price, creating the traditional trade-off between output and inflation. In contrast, sentiment shocks raise inflation but reduce aggregate output, functioning similarly to a cost-push shock. Therefore, the existence of sentiment shocks flattens the Phillips curve by introducing a source of variation that moves prices and output in opposite directions.[11]

4 Discussions

This section examines the robustness and scope of our main results by exploring several extensions and clarifications. There are three main objectives: first, to clarify the information structure and timing that underpin the sentiment-driven equilibrium (see Subsections 4.1–4.2); second, to demonstrate that the mechanism survives under alternative information structures, including learning from prices and observing nominal wages (see Subsections 4.3 and 4.4); and third, to show how the mechanism selects equilibria and its limitations under imperfect common knowledge (see Subsections 4.5 and 4.6).

A unifying theme emerges throughout: sentiment-driven equilibria require two conditions.

[11] To illustrate how sentiment can flatten the Phillips curve by shifting it around, combine the aggregate price equation $\hat{\mathbf{p}}_t = \phi\hat{\mathbf{m}}_t + \mathbf{z}_t$ and the aggregate output equation $\hat{\mathbf{y}}_t = (1-\phi)\,\hat{\mathbf{m}}_t - \mathbf{z}_t$. This gives us the covariance between $\hat{\mathbf{p}}_t$ and $\hat{\mathbf{y}}_t$: $Cov\left(\hat{\mathbf{p}}_t, \hat{\mathbf{y}}_t\right) = \phi\left(1-\phi\right)\sigma_m^2 - \sigma_z^2 < \phi\left(1-\phi\right)\sigma_m^2$. Clearly, in the absence of sentiment, the output gap $\hat{\mathbf{y}}$ and the change in the aggregate price level $\hat{\mathbf{p}}$ are positively related, but this positive relationship is significantly neutralized by $\mathbf{z}$. By equation (37), we have $Cov\left(\hat{\mathbf{p}}_t, \hat{\mathbf{y}}_t\right) = \phi\left(1-\phi\right)\sigma_m^2 - \sigma_z^2 = -\frac{(1-\mu)(2\mu-1)}{(1-\beta)\mu^2}\sigma_\varepsilon^2$, which is close to zero if μ is close to $\frac{1}{2}$. Also see Beaudry, Hou and Portier (2024) for an analysis of monetary policy in the context of a flat Phillips curve.

Firstly, agents must face a genuine signal extraction problem—they cannot distinguish between aggregate sentiment shocks and idiosyncratic shocks (or aggregate fundamental shocks) perfectly. Secondly, the structural parameters must ensure positive net feedback from perceived aggregate sentiment to individual decisions. Our price-setting model satisfies both conditions, whereas the Lucas (1973) island model only satisfies the first condition. Subsection 4.2 formalizes this comparison using an abstract signal-extraction framework and derives the precise condition under which positive net feedback is obtained.

4.1 Timeline of events

For simplicity (without loss of generality), in our benchmark model we assumed that households have perfect information about consumer sentiment. The following timeline of events may therefore help to illustrate the sentiment-driven equilibrium in our benchmark model under incomplete information:

1. At the beginning of each period t, all household members observe ε_{jt} and form the common belief that the aggregate price P_t will be determined by consumer sentiment $\mathbf{z}_t$. In other words, consumer sentiment is the belief that $\hat{\mathbf{p}}_t = \mathbf{z}_t$.

2. Firms also believe that consumer sentiment regarding aggregate prices is true, i.e. $\hat{\mathbf{p}}_t = \mathbf{z}_t$. However, firms cannot directly observe $\mathbf{z}_t$ or ε_{jt}. Instead, they receive an imperfect signal, $s_{jt} = \mu \mathbf{z}_t + (1-\mu)\varepsilon_{jt}$.

3. Based on the signal s_{jt}, firms set prices $\hat{p}_{jt}$ without directly observing the nominal wages in island j to maximize expected profits and commit to producing y_{jt} according to expected goods demand. In other words, firms commit to labor demand, with real wages on each island adjusting to clear the labor market.

4. The realization that $\hat{\mathbf{p}}_t = \int \hat{p}_{jt} dj$ is consistent with firms' rational expectations in general equilibrium. This is a self-fulfilling process because all firms' best responses to consumer sentiment $\mathbf{z}_t$ are positive and mutually reinforcing (see equations 25 or 35).

Based on the timeline, Section 3 has already demonstrated the existence of multiple fundamental and sentiment-driven equilibria under incomplete information. A natural question concerns the economic origin of the signal s_{jt} and why it takes the particular form $s_{jt} = \mu \mathbf{z}_t + (1-\mu)\varepsilon_{jt}$. We now show that this signal structure is mathematically equivalent to households' expectations of firms' prices and hence can be obtained naturally from firms' market survey. For simplicity, assume $\hat{\mathbf{m}}_t = 0$. Consider that firm j conducts a market survey of households about its goods price, yielding

$$\tilde{s}_{jt} = \mathbb{E}^h(\hat{p}_{jt}) + \nu_{jt},$$

where v_{jt} is white noise with variance σ_v^2 capturing idiosyncratic sampling error in the firm's market survey, and $\mathbb{E}^h$ denotes households' expectations. Since households' information set contains both ε_{jt} and aggregate sentiments $\mathbf{z}_t$, we conjecture that

$$\mathbb{E}^h(\hat{p}_{jt}) = \phi_\varepsilon \varepsilon_{jt} + \phi_z \mathbf{z}_t.$$

Thus $\tilde{s}_{jt} = \phi_\varepsilon \varepsilon_{jt} + \phi_z \mathbf{z}_t + \nu_{jt}$. Given this signal, the firm's optimal price is

$$\begin{aligned} \hat{p}_{jt} &= \mathbb{E}\left[(\beta \hat{\mathbf{p}}_t + \varepsilon_{jt}) \,|\, (\phi_\varepsilon \varepsilon_{jt} + \phi_z \mathbf{z}_t + \nu_{jt})\right], \\ &= \frac{\phi_\varepsilon \sigma_\varepsilon^2 + \beta \phi_z \sigma_z^2}{\phi_\varepsilon^2 \sigma_\varepsilon^2 + \phi_z^2 \sigma_z^2 + \sigma_v^2} \left(\phi_\varepsilon \varepsilon_{jt} + \phi_z z_t + v_{jt}\right) \end{aligned}$$

By definition, $\tilde{s}_{jt} = \mathbb{E}^h(\hat{p}_{jt}) + \nu_{jt} = \mathbb{E}^h(\hat{p}_{jt} | \varepsilon_{jt}, \mathbf{z}_t) + \nu_{jt}$, which requires

$$\frac{\phi_\varepsilon \sigma_\varepsilon^2 + \beta \phi_z \sigma_z^2}{\phi_\varepsilon^2 \sigma_\varepsilon^2 + \phi_z^2 \sigma_z^2 + \sigma_v^2} \left(\phi_\varepsilon \varepsilon_{jt} + \phi_z z_t\right) + v_{jt} = \phi_\varepsilon \varepsilon_{jt} + \phi_z z_t + v_{jt}$$

or equivalently,

$$\frac{\phi_\varepsilon \sigma_\varepsilon^2 + \beta \phi_z \sigma_z^2}{\phi_\varepsilon^2 \sigma_\varepsilon^2 + \phi_z^2 \sigma_z^2 + \sigma_v^2} = 1. \tag{38}$$

Finally, the equilibrium condition $\hat{\mathbf{p}}_t = \mathbf{z}_t = \int \hat{p}_{jt} dj$ requires

$$\frac{\phi_\varepsilon \sigma_\varepsilon^2 + \beta \phi_z \sigma_z^2}{\phi_\varepsilon^2 \sigma_\varepsilon^2 + \phi_z^2 \sigma_z^2 + \sigma_v^2} \phi_z = 1, \tag{39}$$

which, combined with equation (38), implies $\phi_z = 1$. Substituting $\phi_z = 1$ into equation (38) gives

$$\sigma_z^2 = \frac{\phi_\varepsilon (1 - \phi_\varepsilon) \sigma_\varepsilon^2 - \sigma_v^2}{1 - \beta}. \tag{40}$$

For a sentiment-driven equilibrium to exist, we require $0 < \phi_\varepsilon < 1$ and $\sigma_v^2 < \phi_\varepsilon (1 - \phi_\varepsilon) \sigma_\varepsilon^2$. To see the connection between the above expression and equation (27) in section 3, set $\phi_\varepsilon = \frac{1}{\mu} - 1$ for $\mu \in (1/2, 1)$. Then

$$\sigma_z^2 = \frac{\frac{(1-\mu)(2-\mu)}{\mu^2} \sigma_\varepsilon^2 - \sigma_v^2}{1 - \beta}, \tag{41}$$

which reduces to equation (27) when $\sigma_v^2 = 0$. When there is no sampling error, for any $0 < \phi_\varepsilon < 1$, μ can be defined as $\frac{1}{1+\phi_\varepsilon}$. The information contained in $\tilde{s}_{jt} = \phi_\varepsilon \varepsilon_{jt} + \mathbf{z}_t$ is equivalent to $s_{jt} = \mu \mathbf{z}_t + (1 - \mu) \varepsilon_{jt}$ up to scaling by $\frac{1}{1+\phi_\varepsilon}$.

This example provides a microfoundation for the signal structure in our benchmark model, while also demonstrating that the variable z_t captures households' beliefs about the aggregate price level.

4.2 Why no sunspot in the Lucas model

An interesting question is why the Lucas model does not have a sentiment-driven equilibrium. Ultimately, the validation of sentiment rests with decision-makers who possess incomplete information. In the Lucas model, output producers confuse idiosyncratic shocks with aggregate movements when making decisions. However, their best response coefficient to aggregate output (through the quantity equation) is negative, creating a negative feedback loop that precludes self-fulfilling expectations (see equation 4).[12]

In our model, decisions are made by price setters, and their best response coefficient to the aggregate price (through the quantity equation) is positive (see equations 25 or 35). This positive feedback creates synchronized movements across all agents, making consumer sentiment self-fulfilling in general equilibrium.[13]

To provide a more general discussion of the Lucas model, we can offer a simple, unified framework that explains why there are no sunspots in the Lucas model but there are in ours. Consider the following abstract signal-extraction problem: the optimal decision rule of agent j is given by $x_{jt} = \mathbb{E}\left(\beta \mathbf{x}_t + \theta \varepsilon_{jt} | s_{jt}\right)$, where $\beta < 1$ is the agent's best-response coefficient, ε_{jt} is independent and identically distributed (iid) noise with variance σ_ε^2, and the signal s_{jt} is given by $s_{jt} = a\mathbf{x}_t + b\varepsilon_{jt}$. Aggregation implies $\mathbf{x}_t = \int x_{jt} dj$. Assuming a sentiment-driven equilibrium where the aggregate variable $\mathbf{x}_t$ is driven solely by a sunspot (sentiment) variable $\mathbf{z}_t$ with variance σ_z^2, i.e., $\mathbf{x}_t = \mathbf{z}_t$; and substituting this into the individual decision rule yields

$$x_{jt} = \mathbb{E}\left(\beta \mathbf{z}_t + \theta \varepsilon_{jt} | a\mathbf{z}_t + b\varepsilon_{jt}\right) = \frac{\beta a \sigma_z^2 + \theta b \sigma_\varepsilon^2}{a^2 \sigma_z^2 + b^2 \sigma_\varepsilon^2} \left[a\mathbf{z}_t + b\varepsilon_{jt}\right]. \tag{42}$$

Aggregation implies $\mathbf{x}_t = \frac{\beta a \sigma_z^2 + \theta b \sigma_\varepsilon^2}{a^2 \sigma_z^2 + b^2 \sigma_\varepsilon^2} a\mathbf{z}_t$. For a sentiment-driven equilibrium to exist, we require $1 = \frac{\beta a^2 \sigma_z^2 + \theta b a \sigma_\varepsilon^2}{a^2 \sigma_z^2 + b^2 \sigma_\varepsilon^2}$, or equivalently, $\sigma_z^2 = \frac{b(\theta a - b)\sigma_\varepsilon^2}{(1-\beta)a^2}$. Since $\beta < 1$ and $b > 0$, a sentiment-driven equilibrium is possible only if

$$\theta a - b > 0. \tag{43}$$

In our baseline model, the decision rule 23 corresponds to the parameterization $\theta = 1$, $a = \mu$, and $b = (1 - \mu)$. Condition (43) reduces to $\mu > \frac{1}{2}$. Thus, sentiment-driven equilibria exist when the signal s_{jt} places sufficient weight on the aggregate component $\mathbf{z}$ ($\mu > \frac{1}{2}$). In the Lucas model, the

[12] This holds true regardless of the signal structure. For example, there is no sentiment-driven equilibrium even if the original signal structure is changed to $s_{jt} = \mu \mathbf{p}_t + (1 - \mu)\varepsilon_{jt}$.

[13] Note that the model can be refined by permitting additional aggregate variables to impact firms' wage rates, including labor market frictions, capital accumulation, consumption habits and money velocity. We can demonstrate that, provided the firms' signal s_{jt} does not encompass the entire set of aggregate variables that determine the nominal wage W_{jt}, the sentiment equilibrium remains robust. In other words, our results require incomplete information—not only in the sense of noise in the signal, but also in the sense of an incomplete set of macro variables in the information set.

decision rule (4) corresponds to the parameterization $\theta = \gamma > 0, a = -1$ and $b = 1$. Substituting these values into condition (43) gives $-\gamma - 1 > 0$, which is impossible.

4.3 Learning from prices

This subsection examines the robustness of our sentiment mechanism in the context of an alternative information structure. We consider the extreme case in which firms receive information about the aggregate price itself rather than the components of their marginal cost.

Formally, we set $\sigma_\varepsilon^2 = 0$ and $\mu = 1$, meaning that the firms' signal is

$$s_{jt} = \hat{\mathbf{p}}_t + \nu_{jt},$$

where ν_{jt} is idiosyncratic noise. This is not just a cosmetic difference; it is the most informative signal that firms could reasonably receive about the state of the economy, aside from observing nominal wages directly. By demonstrating that sentiment-driven equilibria survive in this extreme case, we show that our mechanism is not dependent on a specific, fragile information structure.

This robustness exercise also connects our work to Gaballo (2018), who studied a related environment with learning from prices and found multiple fundamental equilibria. We demonstrate that this is also true of our model. Additionally, we demonstrate the existence of a continuum of non-fundamental, sentiment-driven equilibria. We also explore how the economy responds to fundamental and sentiment shocks.[14]

The interpretation of the signal is also connected to Subsection 4.1. Since households have perfect information about both sentiment and monetary shocks, the actual aggregate price $\widehat{\mathbf{p}}_t$ coincides with their expectation of the aggregate price. Consequently, the noisy term v_{jt} in the signal can be interpreted as firm j's sampling error in its market survey of households' aggregate price expectations. Therefore, the signal received by firm j is simply a noisy measure of the survey of households' beliefs about the aggregate price. This connection strengthens the idea that $\mathbf{z}_t$ captures the shared belief among households regarding the aggregate price level and that firms discover this belief through imperfect surveys.

Given the fundamental monetary shocks with variance σ_m^2, each firm's optimal price become $\hat{p}_{jt} = \mathbb{E}[w_t|s_{jt}] = \mathbb{E}[(\beta\hat{\mathbf{p}}_t + (1-\beta)\,\hat{\mathbf{m}}_t)\,|\,(\hat{\mathbf{p}}_t + \nu_{jt})]$. Conjecturing that $\hat{\mathbf{p}}_t = \phi\hat{\mathbf{m}}_t + \mathbf{z}_t$, it follows

[14] King (1983) also provides a detailed discussion of the operating characteristics of a theoretical macroeconomic model in which monetary disturbances are non-neutral and there may be multiple equilibria as a result of imperfect information. Lubik, Matthes and Mertens (2023) also present a microfounded dynamic model with imperfect information. They demonstrate that indeterminacy can arise around a unique the steady state under incomplete information in dynamic NK type models of monetary policy with Taylor rules. Their incomplete information about fundamentals arises from the exogenous specification of two sets of agents, an informed set and an uninformed set, generating signal extraction problems for the agents. Their paper is related to Benhabib, Wang and Wen (2015), Acharya, Benhabib and Huo (2021), as well as this paper. In this paper, multiple equilibria arise under incomplete information generated by sentiment distributions across agents. This creates signal extraction problems resulting in multiple equilibria on a period-by-period basis in the static setting of Lucas-type models. In this context, firms facing idiosyncratic and aggregate shocks to their demand from households must use optimal signal extraction to set their outputs or prices individually. Unlike Lucas's findings, this can result in multiple self-fulfilling equilibria.

that

$$\begin{aligned}\hat{p}_{jt} &= \mathbb{E}\left[\left(\beta(\boldsymbol{\phi}\hat{\mathbf{m}}_{\mathbf{t}}+\mathbf{z}_{\mathbf{t}})+(1-\beta)\,\hat{\mathbf{m}}_t\right)\middle|\left(\phi\hat{\mathbf{m}}_t+\mathbf{z}_t+\nu_{jt}\right)\right] \\ &= \Omega_5\times\left(\phi\hat{\mathbf{m}}_t+\mathbf{z}_t+\nu_{jt}\right), \end{aligned} \tag{44}$$

where $\Omega_5 \equiv \frac{(1-\beta+\beta\phi)\phi\sigma_m^2+\beta\sigma_z^2}{\phi^2\sigma_m^2+\sigma_z^2+\sigma_v^2}$. Integration of $\hat{p}_{jt}$ yields $\hat{\mathbf{p}}_t = \Omega_5\left(\phi\hat{\mathbf{m}}_t+\mathbf{z}_t\right)$. Consistency requires $\Omega_5 = 1$, or $\phi^2\sigma_m^2+\sigma_z^2+\sigma_v^2 = (1-\beta+\beta\phi)\,\phi\sigma_m^2+\beta\sigma_z^2$, which implies

$$\sigma_z^2 = \phi(1-\phi)\sigma_m^2 - \frac{\sigma_v^2}{1-\beta}. \tag{45}$$

If $\sigma_v^2 < \frac{(1-\beta)}{4}\sigma_m^2$, then there is a continuum of sunspot equilibrium indexed by the coefficient $\phi\in(0,1)$. In fact, any $\phi\in\left(\frac{1}{2}\left(1-\sqrt{1-\frac{4}{1-\beta}\frac{\sigma_v^2}{\sigma_m^2}}\right),\frac{1}{2}\left(1+\sqrt{1-\frac{4}{1-\beta}\frac{\sigma_v^2}{\sigma_m^2}}\right)\right)$ constitutes a sentiment-driven equilibrium.

On the other hand, in the absence of sentiment ($z_t = 0$ and $\sigma_z^2 = 0$), we have

$$\hat{p}_{jt} = \mathbb{E}\left[\left(\beta\boldsymbol{\phi}\hat{\mathbf{m}}_{\mathbf{t}}+(1-\beta)\,\hat{\mathbf{m}}_t\right)\middle|\left(\phi\hat{\mathbf{m}}_t+\nu_{jt}\right)\right] = \Omega_6\times\left(\phi\hat{\mathbf{m}}_t+\nu_{jt}\right), \tag{46}$$

where $\Omega_6 \equiv \frac{(1-\beta+\beta\phi)\phi\sigma_m^2}{\phi^2\sigma_m^2+\sigma_v^2}$. Integration of $\hat{p}_{jt}$ yields $\hat{\mathbf{p}}_t(=\phi\hat{\mathbf{m}}_t) = \Omega_6\phi\hat{\mathbf{m}}_t$. Consistency requires $\Omega_6 = 1$, or $\phi\,(1-\phi) = \frac{1}{1-\beta}\frac{\sigma_v^2}{\sigma_m^2}$. Therefore, as in Gaballo (2018), there are three fundamental equilibria with

$$\phi = \left\{0, \frac{1}{2}\left(1\pm\sqrt{1-\frac{4}{1-\beta}\frac{\sigma_v^2}{\sigma_m^2}}\right)\right\}. \tag{47}$$

Unlike in Gaballo's (2018) model, in our model with sentiment, even if $\sigma_v^2 = \sigma_\varepsilon^2 = 0$ and $\mu = 1$ such that $s_{jt} = \hat{\mathbf{p}}_t$, there is still a continuum of equilibria indexed by $\phi\in[0,1]$ (see equation (45)). In this case, firms have perfect knowledge of aggregate $\mathbf{p}_t$, but since fundamental monetary shocks can be mistaken for sentiment, they cannot yet infer nominal wage costs with certainty. The signal extraction problem persists, as does the possibility of self-fulfilling beliefs.[15]

The key insight is that, even when firms directly observe the aggregate price, they cannot perfectly infer the nominal marginal cost, since the aggregate price itself is a combination of monetary fundamentals and sentiment. This creates a signal extraction problem that sustains sentiment-driven equilibria, even when there is no idiosyncratic noise. Learning from prices does not eliminate self-fulfilling beliefs; it merely alters the way in which those beliefs operate. This highlights the robustness of our mechanism: even when using the most natural signal—the aggregate price itself—sentiment-driven fluctuations remain possible.

[15] There is a sharp contrast with Gaballo's model: it does not allow for sentiment as an independent source of uncertainty, whereas ours does. Consequently, our model admits a richer set of equilibria.

4.4 Learning from nominal wage

To generate a signal extraction problem in our Dixit-Stiglitz benchmark model, we assumed, for simplicity, a Calvo-type sticky price setup in which firms must set the price before observe their nominal wages, in order to yield sentiment-driven equilibria.[16] How robust is this result? In this subsection, we relax the sticky price assumption and replace it with another type of signal extraction problem. Specifically, we assume that firms can observe their nominal wages, but cannot distinguish island-specific shocks from aggregate shocks. In particular, we introduce an additional idiosyncratic labor productivity shock that is correlated with the labor supply shock ϵ_{jt}; namely, islands with higher disutility of labor may also exhibit lower labor productivity. We can show that sentiment-driven equilibria also emerge under imperfect information even if the nominal wage can be perfectly observed.[17]

In order to capture the correlation between the two types of idiosyncratic shocks in a simple way, without introducing too many new notations, we modify the production function (14) to

$$Y_{jt} = \epsilon_{jt}^{-\kappa} N_{jt}, \tag{48}$$

where ϵ_{jt} is the same as the household disutility shock and $\kappa > 0$ measures the degree of correlation between the firm-specific labor productivity shock and the household labor supply shock. The firm's optimal price then becomes

$$P_{jt} = \frac{\sigma}{\sigma - 1} \frac{\mathbb{E}\left[\epsilon_{jt}^{\kappa} W_{jt} \left(\frac{1}{P_t}\right)^{1-\sigma} Y_t \middle| W_{jt}\right]}{\mathbb{E}\left[\left(\frac{1}{P_t}\right)^{1-\sigma} Y_t \middle| W_{jt}\right]}. \tag{49}$$

In log-deviation terms, this simplifies to

$$\hat{p}_{jt} = \hat{w}_{jt} + \kappa \mathbb{E}_t \left[\varepsilon_{jt} | \hat{w}_{jt}\right]. \tag{50}$$

As before, in the absence of monetary shocks and through the quantity equation, we have $\hat{w}_{jt} = \beta \hat{\mathbf{p}}_t + \varepsilon_{jt}$ and $\hat{\mathbf{p}}_t = \int_0^1 \hat{p}_{jt} dj$, where $\beta = 1 - \gamma$. In other words, the firm's optimal price depends on the observed nominal wage and the idiosyncratic shock, which is not directly observable, but must be estimated based on the wage signal. This gives rise to a signal extraction problem. If $\kappa = 0$, then $\hat{p}_{jt} = \hat{w}_{jt}$ and the model reverts to that with perfect information.

[16] The Lucas (1973) model involves firms setting output, rather than price, so our microfounded model differs from Lucas'. Instead, we introduce labour types subject to stochastic disutility shocks into household preferences, rather than a single labor type in a single labor market. Therefore, our model also differs from BWW (2015).

[17] In our benchmark model, in which firms set prices, if firms were to observe wages or marginal costs, they would set prices as a constant markup over these costs or wages. This would cause the disappearance of sentiment equilibria, unless productivity shocks were to correlate with idiosyncratic wage shocks.

Now suppose $\kappa > 0$ and that the equilibrium aggregate price follows $\hat{\mathbf{p}}_t = \mathbf{z}_t$, where $\mathbf{z}_t \sim \mathcal{N}(0, \sigma_z^2)$ is a sunspot. This yields

$$\mathbf{z}_t = \beta \hat{\mathbf{p}}_t + \kappa \frac{\sigma_\varepsilon^2}{\beta^2 \sigma_z^2 + \sigma_\varepsilon^2} [\beta \hat{\mathbf{p}}_t] = \Omega_7 \beta \hat{\mathbf{p}}_t, \tag{51}$$

where $\Omega_7 \equiv 1 + \kappa \frac{\sigma_\varepsilon^2}{\beta^2 \sigma_z^2 + \sigma_\varepsilon^2}$. Consistency requires that $\Omega_7 \beta = 1$, or

$$\sigma_z^2 = \frac{\beta\kappa - (1-\beta)}{(1-\beta)\beta^2} \sigma_\varepsilon^2. \tag{52}$$

If $\beta \in (0,1)$ and $\kappa > \frac{1-\beta}{\beta}$, then a stochastic self-fulfilling sentiment equilibrium exists even in the absence of extrinsic uncertainty.

The intuition is similar to that of the baseline model. When firms observe a higher nominal wage, they cannot tell whether this is due to a higher idiosyncratic shock (ϵ_{jt}) or a general increase in the aggregate price level. Due to the signal extraction problem, they partly attribute the rise in the nominal wage to an increase in the aggregate price level. This leads to synchronized price adjustments across all islands. If intrinsic uncertainty (the variance of sunspots) and idiosyncratic shocks satisfy a specific relationship, as described in equation (52), then an anticipated one-percent increase in the aggregate price level generates an actual one-percent increase. This implies that expectations are self-fulfilling.[18]

4.5 Learning in the Lucas Model

This subsection addresses a different question: equilibrium selection. Unlike the previous two subsections, which examined alternative signal extraction problems, here we consider whether agents can learn which equilibrium to coordinate on through adaptive learning. Specifically, we ask whether the multiple equilibria identified in Section 2 survive under constant-gain learning, and which equilibria are stable.

We now demonstrate that, depending on the initial guess for the parameter α, agents' beliefs will converge to one of the equilibria depicted in Figure 1 in the Lucas model.

Suppose that firm j has an initial common belief or guess, α_0, in period 0. Then, in period t, the firm's belief becomes $\mathbf{y}_t = \alpha_t \mathbf{m}_t$ and $\mathbf{p}_t = (1 - \alpha_t)\, \mathbf{m}_t$. This belief leads to a signal extraction problem regarding the firm's output:

$$y_{jt} = \gamma \mathbb{E}[\varepsilon_{jt} | s_{jt}] = \gamma \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + (1-\alpha_t)^2 \sigma_m^2} [(1-\alpha_t)\mathbf{m}_t + \varepsilon_{jt})]. \tag{53}$$

[18] Nevertheless, equilibria driven by sentiment may not necessarily exist in the event of monetary shocks, although multiple fundamental equilibria can emerge.

Integrating over j yields the actual equilibrium output:

$$\mathbf{y}_t = \frac{\sigma_\varepsilon^2 \gamma}{\sigma_\varepsilon^2 + (1-\alpha_t)^2 \sigma_m^2}(1-\alpha_t)\mathbf{m}_t. \tag{54}$$

The firm then updates its coefficient for the next period according to:

$$\alpha_{t+1} = (1-\theta)\alpha_t + \theta f(a_t) \equiv h(\alpha_t), \tag{55}$$

where $0 < \theta < 1$ is a constant gain learning parameter, and $f(\alpha_t) \equiv \left[\frac{\sigma_\varepsilon^2 \gamma}{\sigma_\varepsilon^2 + (1-\alpha_t)^2 \sigma_m^2}(1-\alpha_t)\right]$ is the output coefficient. This represents a weighted average of the prior belief α_t and the coefficient $f(\alpha_t)$ implied by the realized output. A fixed point of this system satisfying $h(\alpha) = \alpha$ requires

$$\alpha = \frac{\sigma_\varepsilon^2 \gamma}{\sigma_\varepsilon^2 + (1-\alpha)^2 \sigma_m^2}(1-\alpha) = f(\alpha). \tag{56}$$

As before, this can yield three solutions (see equations (7) and (8) and Figure 1).

To analyze stability, we compute the derivative using the fixed-point condition (56 for simplification:

$$f'(\alpha) = -\frac{\alpha}{1-\alpha} + 2\alpha^2 \frac{\tau}{\gamma}, \tag{57}$$

where $\tau \equiv \sigma_m^2/\sigma_\varepsilon^2$. We maintain the parameter restriction $\tau > 3(1+\gamma)$, so there are three equilibria as before. The derivative of the updating function is then

$$h'(\alpha) = 1 - \theta + \theta f'(\alpha). \tag{58}$$

A steady state α is locally learnable if $-1 < h'(\alpha) < 1$.

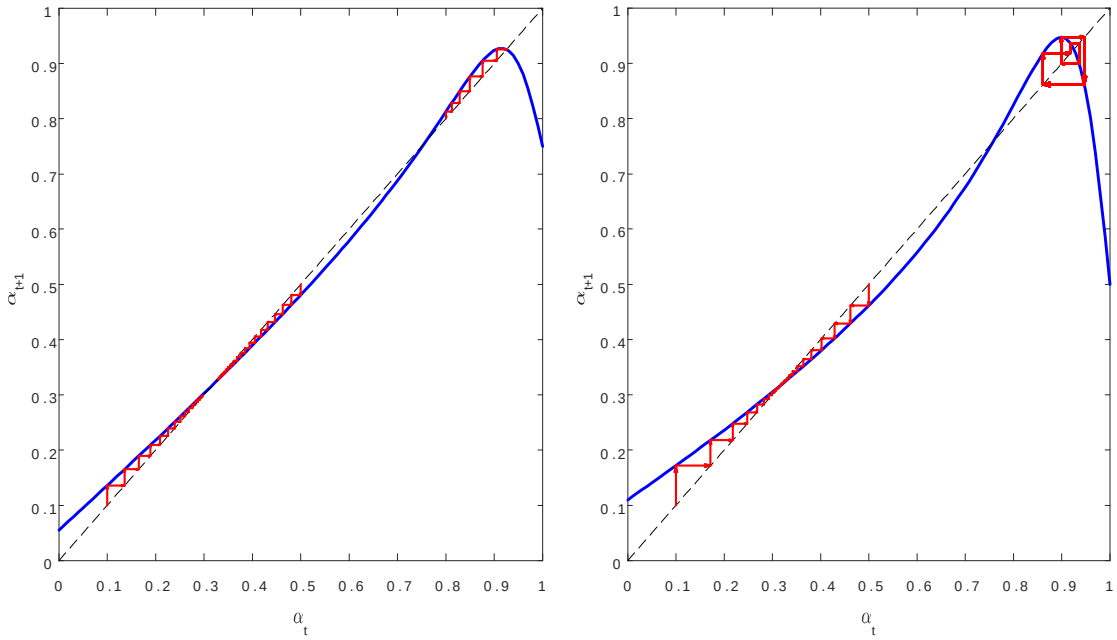


Figure 2. Phase Diagram $\alpha_{t+1} = (1-\theta)\alpha_t + \theta f(a_t)$; $\theta = 0.25$ (Left), $\theta = 0.5$ (right).

Consider the numerical example with three equilibria shown in Figure 1: $\alpha_1 = 0.3170$, $\alpha_2 = 0.7576$ $\alpha_3 = 0.9254$. The derivative at each equilibrium point is given by $f'(\alpha) = [0.4401, 2.0401, -4.7024]$. At α_1, $0 < h'(\alpha_1) < 1$ for all $\theta \in (0,1)$, so α_1 is always learnable. At α_2, $h'(\alpha_2) = 1 + 1.0401\theta > 1$ for all $\theta \in (0,1)$, implying that it is not learnable. At α_3, $h'(\alpha_3) = 1 - 5.7024\theta \in (-1,1)$; clearly, for small values of θ, $1 > h'(\alpha_3) > -1$, so α_3 is learnable. However, as θ increases beyond a critical threshold, $h'(\alpha_3) \leq -1$ and it becomes unlearnable.

For example, the phase diagrams in Figure 2 demonstrate the stability property of the three equilibria in the Lucas model for $\theta = 0.25$ and $\theta = 0.5$, respectively. Specifically, the left window shows that, for $\theta = 0.25$, both α_1 and α_3 are learnable, while the right window shows that, for $\theta = 0.5$, only α_1 is learnable.

In summary, adaptive learning does not eliminate the multiple equilibria of the Lucas model, but rather selects among them. Depending on the initial beliefs and learning gain parameter, the economy may converge to different equilibria. The middle equilibrium is never stable, whereas the outer equilibria are stable for sufficiently small learning gains.

The preceding analysis raises a natural question: does adaptive learning eliminate the sentiment-driven equilibria identified in our baseline model? To address this issue, the appendix examines the stability of the sentiment equilibrium in the context of constant-gain learning. As in the Lucas model exercise, we find that adaptive learning does not eliminate the sentiment equilibrium, but rather selects the positive sentiment equilibrium as the unique stable outcome. The zero-sentiment

(certainty) equilibrium is unstable for $\mu > 1/2$—the same condition required for the existence of a sentiment-driven equilibrium—while the positive sentiment equilibrium is stable for sufficiently small learning gains. This result further supports the robustness of our mechanism: not only are sentiment-driven fluctuations a theoretical possibility, they are also a dynamically stable attractor under plausible learning dynamics.

4.6 Importance of common knowledge

To further demonstrate the importance of common knowledge in establishing sentiment-driven equilibria, we can relax the assumption that agents possess shared knowledge of the structural parameters, such as σ_z^2 and σ_ε^2. For example, consider a scenario in which agents are uncertain about the true variance of the idiosyncratic shock ε_{jt}. Suppose the actual variance is σ_ε^2, but each agent j believes it to be $\chi_j \sigma_\varepsilon^2$, where χ_j is a subjective scaling factor. However, agents continue to share a common belief about the sentiment variance σ_z^2. This exercise is motivated by a large literature following the early work of Mankiw et al (2003) that documents a wide dispersion in inflation beliefs. This heterogeneity in expectations—analogous to heterogeneous beliefs about the variance of shocks in our framework, makes coordination more difficult. Hence it important to see whether it eliminate the self-fulfilling beliefs in driving aggregate prices.

Under this heterogeneous belief structure, solving the signal-extraction problem for agent j as in our abstract model (42) yields

$$x_{jt} = \frac{\beta a \sigma_z^2 + \theta b \chi_j \sigma_\varepsilon^2}{a^2 \sigma_z^2 + b^2 \chi_j \sigma_\varepsilon^2} \left(a \mathbf{z}_t + b \varepsilon_{jt} \right). \tag{59}$$

After aggregation, the equilibrium condition requires

$$\mathbf{1} = \left[\int \frac{\beta a^2 \sigma_z^2 + \theta b a \chi_j \sigma_\varepsilon^2}{a^2 \sigma_z^2 + b^2 \chi_j \sigma_\varepsilon^2} dj \right]. \tag{60}$$

This expression closely resembles that of the baseline model. The critical difference is that the equilibrium condition for σ_z^2 now depends on the distribution of χ_j. For example, suppose χ_j is distributed binomially over $\{0, 2\}$ with equal probability. In this case, the aggregation condition simplifies to $1 = \frac{\rho}{2} + \frac{1}{2} \frac{\beta a^2 \sigma_z^2 + 2\theta b a \sigma_\varepsilon^2}{a^2 \sigma_z^2 + 2 b^2 \sigma_\varepsilon^2}$, which implies

$$\sigma_z^2 = \frac{\left[\theta b a - (2 - \rho) b^2 \right]}{(1 - \beta) a^2} \sigma_\varepsilon^2. \tag{61}$$

For a sentiment-driven equilibrium to exist, we need

$$\theta a > (2 - \beta) b. \tag{62}$$

Since $2 - \beta > 1$, comparing this inequality with the common-knowledge condition ($\theta a > b$) reveals that a lack of common knowledge tightens the parameter restrictions under which sentiment equilibria can arise. In other words, heterogeneity in beliefs about the variance of idiosyncratic shocks makes it slightly more difficult for self-fulfilling sentiment fluctuations to be sustained in equilibrium. Nevertheless, the possibility of sentiment-driven equilibria always remains: sentiment-driven equilibria require positive net feedback and a signal extraction problem, but they do not require perfect common knowledge about all structural parameters.

5 Conclusion

This paper makes several contributions. First, we demonstrate that the Lucas (1973) model exhibits multiple fundamental equilibria driven by extrinsic uncertainty in the money supply due to imperfect information with which to distinguish idiosyncratic and aggregate demand shocks. This finding has not been previously documented in the literature. Second, we provide a microfoundation for this result by embedding the Lucas (1973) model within a standard macroeconomic model. We show that there can be not only multiple fundamental equilibria, but also a continuum of sentiment-driven equilibria based on intrinsic uncertainty, which are not based on randomization across these fundamental equilibria. Third, In a sentiment-driven equilibrium, monetary non-neutrality occurs, and the aggregate price or inflation rate can appear to have a life of its own even with a constant money supply. This seemingly causes the quantity theory of money to fail. This mechanism relies on firms setting prices based on imperfect information about aggregate nominal conditions. Firms' best responses exhibit positive feedback with regard to aggregate price beliefs, enabling equilibrium selection through a common sentiment variable. This contrasts with the Lucas model, in which the relevant best-response slope is negative and sunspot equilibria are precluded. Therefore, our paper offers a clear bridge between Lucas-style signal extraction and Keynesian coordination via sentiments, articulating how imperfect information about aggregate conditions can generate self-fulfilling inflation dynamics without multiple fundamental steady states.

As might be expected, Lucas's results on non-neutrality in his simple model of imperfect information carry over to our comprehensive model featuring sentiment equilibria. The results speak directly to current concerns about fragile inflation expectations and the interpretation of the Phillips curve when expectations/sentiment shift independently of money growth. As the specific effects of more volatile sentiments or monetary policy may differ depending on which equilibrium is realized, the existence of sentiment equilibria may also raise new doubts about the impact and effectiveness of monetary stabilization policies.[19]

[19] Beaudry, Hou and Portier (2024) demonstrate in their work how incorporating a cost channel into monetary policy can shed new light on the relationship between monetary policy and inflation when the Phillips curve is relatively flat. They identify a critical condition under which lax monetary policy can lead to a low inflation trap and discuss how standard inflation-targeting policies may require modification in such circumstances. They also empirically examine

References

[1] Acharya, S., Benhabib, J., and Z. Huo, 2021. The Anatomy of Sentiment-Driven Fluctuations, Journal of Economic Theory, 195, July 2021.

[2] Angeletos, G.M. and Huo, Z., 2021. Myopia and anchoring. American Economic Review, 111(4), pp.1166-1200.

[3] Ascari, Guido, et al. "Endogenous uncertainty and the macroeconomic impact of shocks to inflation expectations." Journal of Monetary Economics 140 (2023): S48-S63.

[4] Beaudry, P., Hou, C. and Portier, F., 2024, "Monetary Policy When the Phillips Curve Is Quite Flat", American Economic Journal (16), January 2024, pages 1-28.

[5] Benhabib, J. and R. Farmer, 1994. Indeterminacy and increasing returns, Journal of Economic Theory 63(1), 19-41.

[6] Benhabib, J. and Mark M. Spiegel. "Sentiments and economic activity: Evidence from US states." The Economic Journal 129.618 (2019): 715-733.

[7] Benhabib, J., Wang, P.F., and Wen, Y., 2015. Sentiments and aggregate demand fluctuations. Econometrica, 83(2), pp.549-585.

[8] Benhabib, J., Wang, P.F., and Wen, Y., 2017. Uncertainty and Sentiment-Driven Equilibria. In Sunspot and Non-linear Dynamics: Essays in Honor of Jean-Michel Grandmont, K. Nishimura, A. Venditti and N. Yannelis, eds., "Studies in Economic Theory" series, Springer Verlag, 2017, NBER Working Paper 18878 (2013) http://www.nber.org/papers/w18878.

[9] Cass, D. and K. Shell, 1983. Do sunspots matter? Journal of Political Economy 91(2), 193-227.

[10] D'Acunto, Francesco, Ulrike Malmendier, and Michael Weber. "What do the data tell us about inflation expectations?." Handbook of economic expectations. Academic Press, 2023. 133-161.

[11] Gaballo, G., 2018. Price Dispersion, Private Uncertainty, and Endogenous Nominal Rigidities. The Review of Economic Studies 85(2) (April) pp. 1070-1110.

[12] Gottardi, Piero and Atsushi Kajii, 1999. "The Structure of Sunspot Equilibria: The Role of Multiplicity," The Review of Economic Studies 66 (3), 713–732.

[13] King, Robert G., 1983. Interest Rate, Aggregate Information, and Monetary Policy. Journal of Monetary Economics 12 (1983) 199-234.

the relevance of the conditions that give rise to these observations.

[14] Lagerborg, Andresa, Evi Pappa, and Morten O. Ravn. "Sentimental business cycles." The Review of Economic Studies 90.3 (2023): 1358-1393.

[15] Leduc, Sylvain, Keith Sill, and Tom Stark. "Self-fulfilling expectations and the inflation of the 1970s: Evidence from the Livingston Survey." Journal of Monetary economics 54.2 (2007): 433-459.

[16] Lucas, R.E., Jr., 1972. Expectations and the Neutrality of Money. Journal of economic theory, 4(2), pp.103-124.

[17] Lucas, R.E., Jr., 1973. Some international evidence on output-inflation tradeoffs. The American economic review, pp.326-334.

[18] Lubik, T., Matthes C. and Mertens E., 2023, Indeterminacy and imperfect information, Review of Economic Dynamics, 2023, vol. 49, 37-57.

[19] Mankiw, N. G., Reis, R., & Wolfers, J. (2003). Disagreement about inflation expectations. NBER macroeconomics annual, 18, 209-248.

[20] Obsfelt, M. and Rogoff K., 1996, Foundations of International Macroeconomics, MIT Press.

[21] Townsend, R., 1983. Forecasting the forecast of others, Journal of Political Economy 91, 546-588.

Appendix

Adaptive Learning and sentiment-driven equilibria

In this appendix, we investigate whether the equilibria driven by sentiment identified in our baseline model persist under adaptive learning. This exercise is similar to the analysis in Section 4.4 for the Lucas model, and it allows us to check that our mechanism is robust. For simplicity, we focus on the case without fundamental shocks, where the signal is $s_{jt} = \mu\hat{\mathbf{p}}_t + (1-\mu)\varepsilon_{jt}$. The rational expectations equilibrium (REE) is defined by $\widehat{\mathbf{p}}_t = \mathbf{z}_t$, with two equilibrium solutions: the certainty equilibrium ($\sigma_z^2 = 0$) and the sentiment-driven equilibrium ($\sigma_z^2 = \frac{(1-\mu)(2\mu-1)}{(1-\beta)\mu^2}\sigma_\varepsilon^2$). We normalize the model such that the sentiment shock $\mathbf{z}_t$ has a unit variance by rewriting $\hat{\mathbf{p}}_t = \sigma_z \mathbf{z}_t$. Then, from the aggregation of firms' optimal prices (25), we have

$$\sigma_z \mathbf{z}_t = \frac{(1-\mu)\sigma_\varepsilon^2 + \beta\mu\sigma_z^2}{(1-\mu)^2\sigma_\varepsilon^2 + \mu^2\sigma_z^2}\mu\sigma_z\mathbf{z}_t.$$

This yields the two REE solutions mentioned above.

We consider the case of adaptive learning. Suppose that, at the beginning of period t, agents conjecture that the aggregate price is $\hat{\mathbf{p}}_t = \sigma_{zt}\mathbf{z}_t$. Aggregating the individual optimal prices yields the realized aggregate price

$$\hat{p}_t = \frac{(1-\mu)\sigma_\varepsilon^2 + \beta\mu\sigma_{zt}^2}{(1-\mu)^2\sigma_\varepsilon^2 + \mu^2\sigma_{zt}^2}\mu\sigma_{zt}.$$

Under adaptive learning with constant gains $g = \theta \in (0,1)$, agents update their perceived aggregate price volatility σ_{zt} according to

$$\sigma_{zt+1} = (1-\theta)\sigma_{zt} + \theta\frac{(1-\mu)\sigma_\varepsilon^2 + \beta\mu\sigma_{zt}^2}{(1-\mu)^2\sigma_\varepsilon^2 + \mu^2\sigma_{zt}^2}\mu\sigma_{zt} \equiv \omega(\sigma_{zt}).$$

The fixed points of the above equation are $\sigma_z = 0$ and a positive solution $\sigma_z = \sqrt{\frac{(1-\mu)(2\mu-1)}{(1-\beta)\mu^2}}\sigma_\varepsilon \equiv \bar{\sigma}_z$. To analyze stability, compute

$$\omega'(0) = 1 - \theta + \theta\frac{\mu}{1-\mu} > 1 - \theta + \theta = 1$$

for $\mu > \frac{1}{2}$, so the zero-sentiment equilibrium is unstable and any small perturbation in sentiment variance will grow, pushing the economy towards the positive-sentiment equilibrium.

We can examine the stability of positive sentiment, $\bar{\sigma}_z = \sqrt{\frac{(1-\mu)(2\mu-1)}{(1-\beta)\mu^2}}\sigma_\varepsilon$, by considering the following: After simplification, we can obtain $\omega'(\bar{\sigma}_z)$:

$$\omega'(\bar{\sigma}_z) = 1 - \frac{2\theta(1-\beta)(2\mu-1)}{\mu(1+\beta)-\beta}.$$

For the positive sentiment equilibrium to be stable, we require $|\omega'(\bar{\sigma}_z)| < 1$, since the second term is positive for $\mu > \frac{1}{2}$ and $\beta < 1$. The condition $\omega'(\bar{\sigma}_z) > -1$ then requires

$$\frac{2\theta(1-\beta)(2\mu-1)}{\mu(1+\beta)-\beta} < 2$$

or we need

$$\theta < \frac{\mu(1+\beta)-\beta}{(1-\beta)(2\mu-1)}.$$

Thus for sufficiently small learning gains θ, the positive sentiment equilibrium is stable.

This appendix has therefore demonstrated that, rather than eliminating sentiment-driven equilibria, adaptive learning selects the positive sentiment equilibrium as the stable outcome. The zero-sentiment equilibrium is unstable for $\mu > 1/2$ — the same condition required for a sentiment-driven equilibrium to exist. For sufficiently small learning gains, the positive sentiment equilibrium is locally stable and serves as the attractor of the learning dynamics. This result is similar to the analysis in Section 4.4 for the Lucas model: in both cases, adaptive learning provides a selection mechanism without eliminating the multiplicity of equilibria. In our model, however, learning selects the sentiment-driven equilibrium as the only stable outcome, which reinforces the robustness of our mechanism.