The Natural Rate of Interest Through a Hall of Mirrors*

Phurichai Rungcharoenkitkul[†] Fabian Winkler[‡]

December 21, 2023

Abstract

We propose a novel explanation for persistent movements in the natural rate of interest (r-star) based on two-sided learning between the central bank and the private sector. We analyze a New-Keynesian model where both learn about r-star from each other. When both sides fail to recognise that their actions influence the other's beliefs, a "hall-of-mirrors" effect arises that causes persistent shifts in r-star in response to cyclical shocks. The effect can explain the post-2008 decline in r-star even if long-run fundamentals had not changed. Conversely, a surge in inflation accompanied by monetary policy tightening can induce a persistent r-star increase.

JEL Classification: E52, E58, D83

Keywords: Natural rate of interest, learning, overreaction, dispersed information, long-term rates, demand shocks, monetary policy shocks.

*We thank Leonardo Melosi for helpful comments on an earlier draft; our discussants Stefania d'Amico, Afrasiab Mirza, Alp Simsek and David Vestin; Claudio Borio, Jenny Chan, Giovanni Favara, Etienne Gagnon, Ben Johannsen, Esa Jokivuolle, Nigel McClung, Hyun Song Shin, Min Wei; and seminar/conference participants at the Federal Reserve Bank of Chicago, the 2022 AEA annual meeting, Bank for International Settlements, Bank of Finland, Bank of Thailand, the Expectations in Dynamic Macroeconomic Models conference at the Czech National Bank, CEF 2021 Conference, EEA Annual Congress 2021, the European Winter Meeting of the Econometric Society 2021, the 52nd MMF Annual Conference and Swiss National Bank. All remaining errors are ours. The views herein are solely those of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, the Bank for International Settlements, or of any other person associated with the Federal Reserve System or the Bank for International Settlements.

[†]Bank for International Settlements (Basel, Switzerland): phurichai.rungcharoenkitkul@bis.org

[‡]Board of Governors of the Federal Reserve System (Washington, DC, United States.): fabian.winkler@frb.gov

1 Introduction

Few concepts have had greater influence on recent monetary policy debate than the natural rate of interest, or r-star—the real interest rate consistent with output equaling potential and stable inflation. In the decades since the 1980s, real interest rates in advanced economies have fallen by more than 5 percentage points. Interest rates fell sharply further in the wake of the Great Financial Crisis (GFC) in 2008, as major central banks cut their policy interest rates to record lows in a bid to support economic recovery. Nominal and real interest rates then stayed persistently low in the subsequent decade as global inflation remained subdued even after demand recovered. Standard macroeconomic theory rationalises the co-existence of limited price pressure and very low interest rates by a decline in r-star. This diagnosis raises a policy problem because, given an effective lower bound (ELB) on the nominal policy rate, it implies that central banks are less able to provide policy accommodation when the need arises. Such concerns have led some central banks to introduce unconventional policy measures, and more recently to review their monetary policy frameworks with a view to regaining policy space.

Several hypotheses have been proposed to explain a persistent fall in r-star, including a fall in trend productivity growth, increased longevity and higher demand for safe assets, among others. These explanations invoke different changes in economic *fundamentals* that raise real desired savings or lower desired investment, putting downward pressure on the equilibrium real interest rate. Empirically, there is little consensus, however, on the relative importance of these factors. The literature that evaluates these competing explanations without imposing a priori theoretical restrictions is relatively scant, and tends to find only limited explanatory power of various saving-investment factors consistently over long samples (see Borio et al. (2022) and Lunsford and West (2019)). The lack of conclusive empirical evidence is perhaps not surprising given the inherent identification challenge. Not only is r-star unobservable, but it is also a theoretical construct—it can only be estimated by taking a stand on the correct model of the economy. This leaves open the possibility that other factors may well be relevant secular drivers of real interest rates.

This paper proposes an alternative explanation of persistent movements in r-star that is based on endogenous *beliefs* and informational feedback. The central idea is that r-star depends on beliefs which can evolve in a persistent way when the central bank and the private sector learn from each other. When the central bank adjusts the policy rate, it sends a signal about r-star that the private sector incorporates into consumption-saving decisions. This in turn affects macroeconomic outcomes that feed into the central bank's inference about r-star. When both the central bank and the private sector underestimate the importance of this informational feedback, r-star can become endogenous to cyclical perturbations, including those of monetary policy.

We formalise this idea by adding imperfectly observed stochastic trends to the canonical New Keynesian model. These trends are the exogenous fundamentals that determine the real interest rate in the long run. Both the private sector and the central bank learn about these fundamentals from their own private information, as well as from public observations of output, inflation and interest rates. Agents draw on public observations because they provide useful signals of the other side's private information. To our knowledge, this setup with non-nested information sets of the central bank and the private sector is new to the literature.

Our simple extension has non-trivial implications. We show that r-star becomes endogenous to cyclical fluctuations and monetary policy because agents partially attribute their unexpected variations to useful information about long-run r-star fundamentals. Moreover, we highlight a particularly plausible yet dangerous situation in which neither the central bank nor the private sector is fully aware of the informational feedback caused by two-sided learning. Specifically, we examine the case where each side does not take into account the fact that the other side is also learning from them.[1] Under a reasonable calibration, r-star beliefs overreact to cyclical macroeconomic shocks in a persistent and sizeable manner. We call this overreaction a *hall-of-mirrors* effect in reference to Bernanke (2004). Because the central bank and the private sector rely on one another when forming beliefs, both can end up confusing the effects of their own actions with useful information.

To illustrate the mechanism, suppose that the central bank cuts interest rates sharply in response to a recession. Facing uncertainty, private agents attribute a part of this policy adjustment to the central bank having received information about a lower natural interest rate, even if no such information has arrived. As the private sector updates their beliefs and lowers their r-star estimate, output and inflation fall. Not internalizing the two-sided learning dynamic, the central bank now interprets this additional demand shortfall as an indication of a lower r-star, hence adjusts its own r-star estimate downward and cuts interest rate further. The private sector again interprets the additional policy easing as information of a further r-star reduction, and so on. Both sides misinterpret the macroeconomic effects of their own actions as genuine information. They are staring into a hall of mirrors.

[1] An alternative modelling approach is carried out in the working paper version of our paper (Rungcharoenkitkul and Winkler, 2021). There, we assume that the central bank and the private sector are aware of two-sided learning, but overestimate the precision of the other side's private signal. As a result, they are also not fully aware of the degree to which the other side is learning from them. The assumption that people believe others to be better informed than is actually the case relates to the psychology concepts of *informational influence* or *social proof* (Cialdini et al., 1999), a cognitive bias thought to arise particularly in situations of high uncertainty.

Despite its simplicity, our model can explain a range of salient empirical facts in the post-GFC period. The model can quantitatively explain much of the decline of US long-term real interest rates between 2008 and 2019 without appealing to shifts in fundamentals. Moreover, it rationalizes the empirical excess sensitivity of long-term forward real rates to monetary policy surprises, a seeming violation of the long-run monetary neutrality. This sensitivity arises naturally in our model, as the private sector's r-star beliefs (hence the de facto r-star) are endogenous to monetary policy actions. Finally, the model also explains the puzzling persistence of forecast errors in measured interest rate expectations, which in our model arises from agents' incomplete understanding of the two-sided learning process between the central bank and the private sector.

The hall-of-mirrors effect has important implications for current monetary policy debates. The extraordinary monetary policy measures in the decade after the GFC were guided in no small part by policymakers' beliefs that r-star had substantially fallen, for reasons outside their control. But with the hall-of-mirrors effect, an aggressively accommodative policy strategy exacerbates the very problem policymakers are trying to solve as it causes r-star beliefs to fall endogenously and persistently. Conversely, rapid policy tightening in response to high inflation, as during the recovery from the COVID-19 pandemic, can persistently raise r-star beliefs and long-term interest rates even after inflationary pressures dissipate. Our model thus calls for greater recognition of the unintended consequences of policy strategy and communications.

The macroeconomic literature has extensively studied cases in which only the central bank learns about economic fundamentals from the private sector. Prominent contributions are Orphanides (2003), Cukierman and Lippi (2005) and Primiceri (2006) and Nimark (2008). Orphanides and Williams (2007, 2008) allow for imperfect information on behalf of the private sector, though only about the short-run dynamics of the economy. On the empirical side, the well-known r-star estimation procedures of Laubach and Williams (2003) and others (e.g. Holston et al., 2017; Johannsen and Mertens, 2021) also belong in this category, since they estimate r-star from macroeconomic and financial variables that reflect private sector information and expectations. However, these empirical studies implicitly assume that r-star is exogenous to monetary policy.

On the flip side, a more recent strand of the literature has examined the case in which only the private sector learns about economic fundamentals from the central bank. This gives rise to what is called the signalling channel of monetary policy, which has been prominently documented empirically by Nakamura and Steinsson (2018). Tang (2015), Melosi (2016), Angeletos et al. (2020) and Hillenbrand (2022) provide the theoretical basis for this information channel.

Our paper forms a bridge between the two strands of the literature above. To our knowledge, it is the first in which both the central bank and the private sector are learning about uncertain fundamentals underpinning r-star from each other.

The idea that the central bank and private sector expectations can reinforce each other has been explored in the monetary economics literature, but not in the context of r-star as we do here. Bernanke and Woodford (1997) argue that if the central bank targets private sector inflation forecasts to steer actual inflation, indeterminacy can obtain from a positive feedback loop between expectations. In our model, the equilibrium is always determinate as fundamentals anchor long-run r-star expectations, but amplification of noise can still create a persistent wedge between the de facto r-star and the level consistent with fundamentals. Morris and Shin (2002) argue that the information provided by monetary policy communications can crowd out dispersed information in the private sector, preventing an efficient aggregation of information. In our model, the main source of inefficient information aggregation comes from the fact that the central bank and the private sector are relying on each other without internalising the informational feedback, thus generating more powerful consequences than in Morris and Shin (2002).

Our model also relates to an emerging literature on the possibility that r-star could be endogenous to monetary policy. In Rungcharoenkitkul et al. (2019), a monetary policy regime that focuses unduly on short-term output can exacerbate the financial boom-bust cycle, resulting in lower equilibrium output and interest rates in the long run. In Mian et al. (2020), the natural rate of interest is lower when demand is constrained by over-indebtedness, which can be result from monetary policy accommodation. Similarly in Beaudry and Meh (2021), low interest rates can push the economy into an ELB trap in which r-star is endogenously low. In our model, r-star is endogenous not because of fundamental economic mechanisms, but because of mutual learning and endogenous information acquisition. The notorious practical difficulties in assessing r-star speak to the importance of having a model where learning is a central feature.

The remainder of this paper is organised as follows. The next section discusses empirical evidence that motivates our analysis. Section 3 sets up the basic macroeconomic framework modified to accommodate incomplete information, and establishes the modified r-star concept. Section 4 builds intuition by analysing a tractable static version of the model and deriving key qualitative results. Section 5 lays out the full dynamic version of our model, detailing the analogous equilibrium concept and solution methods. Section 6 discusses our quantitative simulation results and assesses the macroeconomic relevance of the hall-of-mirrors effect. Section 7 examines potential policy implications. Section 8 concludes.

2 Motivating evidence

Is our proposed hall-of-mirrors effect simply a theoretical curiosity? The answer would be yes if the natural rate hypothesis held true, i.e. if the short-term real interest rate expected to prevail in the long run is independent of monetary policy. This hypothesis, essentially a form of long-run monetary neutrality, implies that private agents form r-star beliefs independently of monetary policy actions and communications. The prediction, if validated empirically, would indeed rule out the hall-of-mirrors effect.

There is, however, strong evidence that monetary policy affects market expectations of interest rates over very long horizons. Hanson and Stein (2015) and Hanson et al. (2018) document that monetary policy news have a surprisingly strong effects on forward real interest rates in the distant future, interpreted by the authors as a correlation between risk premia and policy shocks.[2] However, accounting for risk premia does not entirely eliminate this puzzling sensitivity of expected short-term rate in the long future to monetary policy. The left panel in Figure 1 plots the one-year nominal bond yield (black line), a proxy for the policy interest rate and its near-term outlook, against the forward real rate of the same maturity nine years ahead (blue line). The latter is constructed from a risk-neutral yield curve, where risk premia have been removed as in Adrian et al. (2013), thus providing a better estimate of the expected short-term interest rate component. This risk-neutral long forward rate thus serves as a reasonable proxy for r-star. There is a high correlation between the two series, which is evidently driven by variations of interest rates over the monetary policy cycles. In each episode of persistent tightening or loosening of the short-term interest rate, the real long rate follows suit. The highly cyclical pattern of this r-star proxy is seemingly at odds with the concept of an exogenous long-run real interest rate.

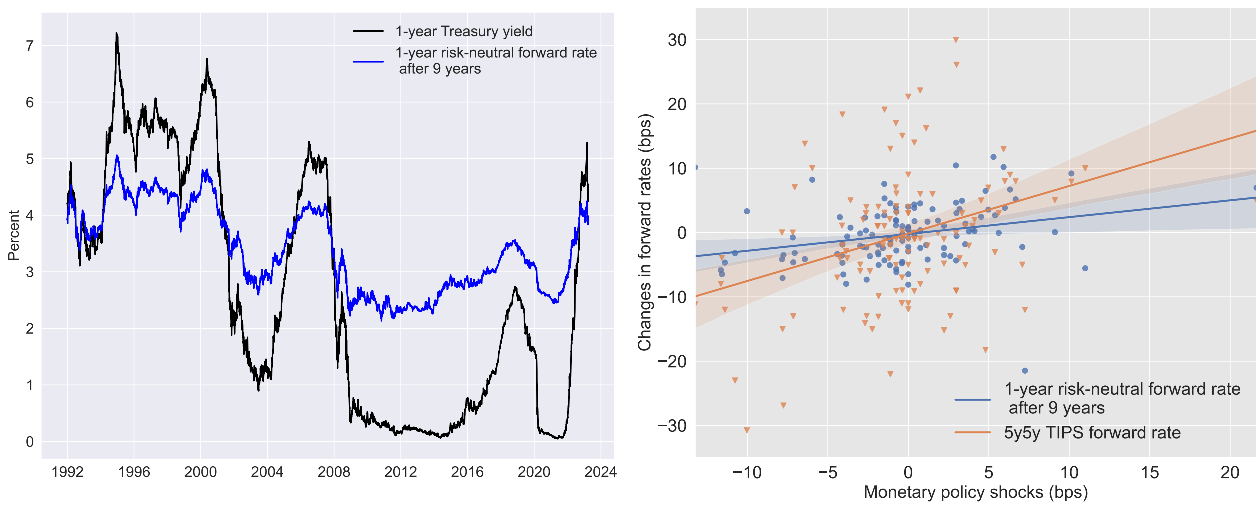
As a more stringent test, we gauge the causal impact of monetary policy by using high frequency-identified monetary policy surprises to examine how the long forward rate responds to monetary policy surprises immediately after FOMC meetings. The right panel in Figure 1 shows significant positive responses of risk-neutral long forward rates to monetary policy surprises.[3] Relatedly, Hillenbrand (2022) documents that the change in 10-year nominal yields around FOMC meetings explains the entire decline in 10-year yields over the last thirty years. Monetary policy thus seems to impart a significant effect on the market expectations

[2] Hanson and Stein (2015) estimate a regression of changes in forward interest rates on changes in 2-year nominal yields on FOMC announcement dates, and find that a 100 bps change in 2-year nominal yields translates to a 40 bps change in real forward rate at the 10-year horizon.

[3] The positive response is stronger if one uses the 5-year 5-year real forward rates from the TIPS market, though part of this responsiveness may owe to the risk premium component as noted in Hanson and Stein (2015). At the same time, the result rules out excess sensitivity of long-run inflation expectations to monetary policy as an explanation.

Figure 1: Relation of long-term rates and monetary policy.

(a) Correlation between long and short rates. (b) Sensitivity of long rates to policy surprises.



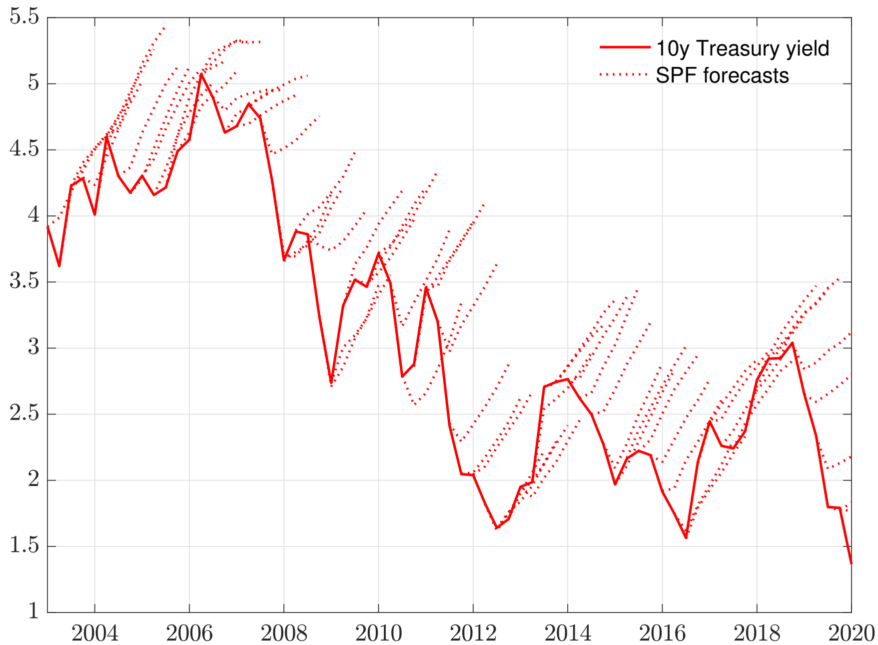
Note: Left panel: The black line shows the nominal one-year US Treasury bond yield. The blue line shows the 9y1y forward interest rate. Right panel: Blue dots and fitted line trace the relationship between changes in the 9y1y-neutral forward interest rate around policy surprise events (2-day window), and the size of monetary policy shocks during those events. Orange dots and line use the 5y5y real forward interest rate from the TIPS market as the forward rate. Shaded areas are 90 percent confidence bands. The period covered in both cases is December 2003 to June 2019. Policy shock series are from Kearns et al. (2018). The forward rate in both panels is from the risk-neutral yield curve as in Adrian et al. (2013).

of steady-state interest rate.

There is also evidence that expectations about long-term rates do not conform to the rational expectations hypothesis, as has been documented previously in the literature (e.g. Coibion and Gorodnichenko, 2015). Figure 2 plots the time-series of the 10-year US Treasury yield alongside forecasts from the Survey of Professional Forecasters. The long-term interest rate declined continuously throughout the sample, by over 3 percentage points from its peak. Yet, forecasters consistently expected the decline in long-term yields to reverse each time they were surveyed, leading to systematic forecast errors. These persistent errors are particularly puzzling considering that many proposed fundamental drivers of r-star, such as life expectancy or dependency ratios, follow slow-moving and predictable trends which should, in turn, make the r-star trends forecastable.

As we will show, these seemingly puzzling empirical patterns arise naturally in the model of two-sided learning about r-star. Long-end forward rate is sensitive to the policy rate because the private sector learns from the central bank's actions, effectively making r-star endogenous to monetary policy. And because both are learning from each other but may not realise that the other is doing the same, r-star can shift in a persistent yet unpredictable

Figure 2: Trend decline in long-term yield was largely unforeseen.



Note: The solid line is the 10-year Treasury yield, while dotted lines represent the projected paths of 10-year Treasury yield according to the Survey of Professional Forecasters. The start of each line marks the current yield as of the survey date, and is hence on the solid line by definition.

way. We later show via simulation that the model can moreover account for these empirical patterns in quantitative terms.

3 Macroeconomic environment

We now introduce the model and show how the de facto natural real interest rate is fundamentally an expectation. By influencing how agents make consumption and saving decisions, this expectation also dictates how inflation and output respond to monetary policy.

3.1 The New Keynesian model with unobserved trends

Our model is the standard New Keynesian model, but with incomplete information about stochastic trends. A representative household solves the utility maximisation problem

$$\max_{\{C_t,N_t,B_t\}_{t=0}^{\infty}} E_0^h \sum_{t=0}^{\infty} \beta^t \Xi_t \left(\frac{C_t^{1-\sigma}}{1-\sigma} - A_t^{1-\sigma} \frac{N_t^{1+\varphi}}{1+\varphi} \right)$$
$$\text{s.t.} \quad P_t C_t + B_t = (1+i_{t-1}) B_{t-1} + P_t (W_t N_t + T_t + \Pi_t)$$

by choosing consumption C_t, labour supply N_t, and nominal bond holdings B_t which are in zero net supply and yield a nominal return of i_t. Consumption C_t is an aggregate of differentiated goods:

$$C_t \equiv \left(\int_0^1 C_{it}^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}}$$

which gives rise to a standard CES demand function. The household takes as given the price level P_t, the real wage W_t, dividends from firms Π_t and any lump-sum transfers from the government T_t. The utility function is affected by shocks to the rate of time preference Ξ_t. To ensure a balanced growth path with trend productivity shocks, the disutility of labour also depends on $A_t^{1-\sigma}$, where A_t is the aggregate productivity level.

Importantly, the expectation operator E^h is conditional on the information set of private agents, namely the household and firms. This information set is potentially incomplete and different from that of the central bank. The expectations also need not coincide with rational expectations once we introduce misperception later on.

Differentiated goods are produced by a continuum of firms $i \in [0, 1]$ with the technology

$$Y_{it} = A_t N_{it}^{1-\alpha}$$

and sold at the price P_{it} (whose CES sum over i equals P_t). Firms are subject to Calvo pricing frictions and can only re-optimise their prices with probability $1 - \theta$. Firms' revenues are subsidised at a rate τ_t. In steady state, this subsidy is set to the value $\bar{\tau}$ that ensures efficiency, while random fluctuations around this value act as cost-push shocks. Firms distribute their profits to the household. The government consists of a central bank that sets the nominal interest rate i_t, and a fiscal authority that collects taxes on firms and distributes the proceeds lump-sum to the household.

Productivity growth consists of a persistent and a temporary component. Defining $a_t \equiv \log(A_t)$, we posit:

$$\begin{aligned} a_{t+1} &= a_t + g_t + \epsilon_{at+1}, & \epsilon_{at+1} &\sim \mathcal{N}\left(0, \sigma_a^2\right) \\ g_{t+1} &= \rho_g g_t + \epsilon_{gt+1}, & \epsilon_{gt+1} &\sim \mathcal{N}\left(0, \sigma_g^2\right). \end{aligned}$$

The shock to the discount factor $\xi_t \equiv \log(\Xi_t)$ similarly consists of a persistent and a temporary component, but also a moderately persistent part:

$$\xi_{t+1} = \xi_t - z_t - u_{ht} - \epsilon_{\xi,t+1}, \qquad \epsilon_{\xi,t+1} \sim \mathcal{N}\left(0, \sigma_\xi^2\right)$$
$$z_{t+1} = \rho_z z_t + \epsilon_{zt+1}, \qquad \epsilon_{z,t+1} \sim \mathcal{N}\left(0, \sigma_z^2\right)$$
$$u_{ht+1} = \rho_h u_{ht} + \epsilon_{ht+1}, \qquad \epsilon_{h,t+1} \sim \mathcal{N}\left(0, \sigma_h^2\right).$$

Throughout the paper, we will set $\rho_g = \rho_z = 1$, so that g_t and z_t are both random walks. To be clear, we do not intend to claim that productivity growth and the natural real rate are literally random walks. Rather, we choose this value in accordance with the empirical literature on estimating r-star, which has documented that the random walk assumption fits the data well in small samples.[4]

A key departure from the standard setup arises from agents' incomplete information about the processes a_t and ξ_t. In particular, everyone can observe the current productivity level a_t and the current preference ξ_t as well as u_{ht}. But they cannot separately observe the subcomponents $g_t, \epsilon_{at}, z_t, \epsilon_{\xi t}$. As a result, agents cannot disentangle movements in a_t and ξ_t that are attributable to the permanent components g_t and z_t from those that are due to the temporary shocks ϵ_{at} and $\epsilon_{\xi t}$.

3.2 The belief-driven natural interest rate

We now derive the natural interest rate with incomplete information. Log-linearising the first-order conditions and solving the model leads to the familiar Euler equation:

$$E_t^h\left[\Delta y_{t+1}\right] = \frac{1}{\sigma}\left(i_t - E_t^h\left[\pi_{t+1}\right] + E_t^h\left[\Delta \xi_{t+1}\right]\right) \tag{3.1}$$

where $i_t - E_t^h\left[\pi_{t+1}\right]$ is the ex-ante real interest rate from the perspective of private agents. Evaluating this equation under flexible prices, where output is at its natural level $y_t^* = a_t$ and $E_t^h\left[\Delta y_{t+1}^*\right] = E_t^h\left[g_t\right]$, one can back out out the corresponding level of the real interest rate under flexible prices as

$$E_t^h\left[\sigma g_t + z_t\right] + u_{ht}.$$

[4]Technically, ξ_t and g_t have to be bounded in order to guarantee that expected discounted utility remains finite. Imposing such bounds would introduce a non-linearity that would render the filtering problems in the model computationally prohibitive. Instead, we derive the linearised model equations for $\rho_g, \rho_z < 1$ and then set these parameters to one in the linearised model.

We define r-star as the private sector's expectation of this real interest rate in the long run. If trend growth g_t and the trend discount rate z_t were fully observed, then r-star would be:

$$r_t^{**} = \sigma g_t + z_t. \tag{3.2}$$

This notation follows Laubach and Williams (2003), but we denote this object "r-double-star": These are the *fundamentals* driving r-star. They are exogenous but unobservable. The *de facto* nautral rate in the economy under incomplete information is:

$$r_t^* = E_t^h \left[\sigma g_t + z_t\right] = E_t^h \left[r_t^{**}\right]. \tag{3.3}$$

The *de facto* r-star is an expectation, and as such it is endogenous to changes in the private sector's information. It is this endogenous expectation that actually determines aggregate demand. Denoting the output gap by $\tilde{y}_t \equiv y_t - y_t^*$, one obtains the familiar IS curve:

$$E_t^h \left[\Delta \tilde{y}_{t+1}\right] = \frac{1}{\sigma} \left(i_t - E_t^h \left[\pi_{t+1}\right] - r_t^* - u_{ht}\right). \tag{3.4}$$

The second equation of the linearised model is the Phillips curve:

$$\pi_t = \beta E_t^h \left[\pi_{t+1}\right] + \kappa \tilde{y}_t + u_{pt} \tag{3.5}$$

where $\kappa > 0$ is a function of other primitive parameters (see Appendix A and Galí (2015) for detailed derivation). The cost-push shock $u_{pt} \sim \log \tau_t$, which is observed by private sector agents, is assumed to follow a normal AR(1) process with autocorrelation ρ_p and innovation variance σ_p^2.

We close the model by assuming that the central bank sets the nominal interest rate according to a standard Taylor-type rule with inertia:

$$i_t = \rho_i i_{t-1} + (1 - \rho_i) \left(\hat{r}_t^* + \phi_\pi \pi_t + \phi_y \tilde{y}_t + u_{ct}\right). \tag{3.6}$$

The monetary policy shock u_{ct} is observed by the central bank but not by the private sector. It is assumed to follow a normal AR(1) process with autocorrelation ρ_c and innovation variance σ_c^2.

Like private agents, the central bank cannot directly observe r_t^{**} and must form an estimate to set policy:

$$\hat{r}_t^* \equiv E_t^c \left[r_t^{**}\right] \tag{3.7}$$

where E_t^c denotes the expectation with respect to the central bank's information set and inference. In our model, the information sets of the private sector and the central bank will not be nested, so that r_t^* and $\hat{r}_t^*$ are not identical.[5]

This incomplete-information recast of the New Keynesian model yields two key insights:

1. The *de facto* natural rate of interest relevant to the economy, r_t^*, is belief-dependent. It is whatever private agents expect the long-run level of interest rates to be. It is only in the special case where private agents perfectly observe the fundamentals r_t^{**} (and understand the model correctly) that r_t^* is exogenous.

2. The *de facto* natural interest rate r^* is not necessarily the same as the estimate $\hat{r}^*$ used by the central bank to guide monetary policy. The two coincide only when the central bank and the private sector share the same beliefs, but in general the private sector and the central bank will disagree about r-star.

In the next section, we let both the central bank and the private sector learn from each other through observing macroeconomic outcomes.

4 The hall-of-mirror effect: Building intuition

In this section, we illustrate the hall-of-mirror effect in a static version of the New Keynesian model discussed above. This allows us to focus on the mutual learning problem and develop intuition for how the mechanism operates. The central insights carry over to the dynamic setting in the next section.

4.1 A static model

Assume that the economy returns to full employment from period 1 onwards, so that $E_0^h[\tilde{y}_t] = E_0^c[\tilde{y}_t] = 0$ for all $t \geq 1$, and that the Taylor rule has no inertia, $\rho_i = 0$. By the Phillips curve equation 5.8, both the central bank and households expect inflation to return to zero in period 1. The model then becomes effectively static, and can be summarised in terms of period-0 variables, omitting time subscripts:

$$\tilde{y} = -\frac{1}{\sigma}(i - r^* - u_h) \tag{4.1}$$

$$\pi = \kappa\tilde{y} + u_p \tag{4.2}$$

$$i = \hat{r}^* + \phi_\pi \pi + \phi_y \tilde{y} + u_c \tag{4.3}$$

[5] An alternative assumption is that the central bank uses the second-order belief $E_t^c E_t^h[r_t^{**}]$ as the intercept of the Taylor rule. Even then, the two-sided learning dynamics persist as the second-order belief still contains useful information about the central bank's private information. Our simulations of this variant of the model, discussed briefly in Section 7, yield similar results.

$$0 = \hat{r}^* + u_c - r^* - u_h + \phi_\pi u_p + (\sigma + \kappa\phi_\pi + \phi_y)\,\tilde{y} \tag{4.4}$$

$$i = \hat{r}^* + \phi_\pi u_p + (\kappa\phi_\pi + \phi_y)\,\tilde{y} + u_c \tag{4.5}$$

where $r^* \equiv E^h[r^{**}]$ and $\hat{r}^* \equiv E^c[r^{**}]$. We assume that the fundamentals r^{**} and macroeconomic shocks are normally distributed:

$$r^{**} \sim \mathcal{N}(0,1) \tag{4.6}$$

$$u_i \sim \mathcal{N}\left(0, \sigma_{ui}^2\right), \quad i = c, h. \tag{4.7}$$

The stochastic terms (r^{**}, u_c, u_h) are mutually independent. The prior on r^{**} has zero mean and unit variance without loss of generality.

Solving (4.1)-(4.3) gives

$$\tilde{y} = \frac{1}{\lambda}\left(r^* - \hat{r}^* + u_h - u_c - \phi_\pi u_p\right) \tag{4.8}$$

$$\pi = \frac{\kappa}{\lambda}\left(r^* - \hat{r}^* + u_h - u_c + \frac{\sigma + \phi_y}{\kappa} u_p\right) \tag{4.9}$$

$$i = \frac{\sigma}{\lambda}\left(\hat{r}^* + u_c\right) + \left(1 - \frac{\sigma}{\lambda}\right)\left(r^* + u_h\right) + \phi_\pi\left(\frac{\sigma + 2\phi_y}{\lambda}\right) u_p \tag{4.10}$$

where $\lambda = \sigma + \phi_\pi\kappa + \phi_y$. The output gap and inflation increase with the difference $r^* - \hat{r}^*$, because the stance of monetary policy matters only relative to the de-facto neutral real rate. As a result, disagreement about r-star can cause the output gap to deviate from zero, even in the absence of demand and policy shocks. Furthermore, the interest rate that prevails in equilibrium becomes a weighted average between the beliefs of the private sector r^* and those of the central bank $\hat{r}^*$.

To form beliefs about the natural interest rate, the private sector and the central bank rely on two sources of information. First, the private sector "h", and the central bank "c" each receives a signal about r^{**}:

$$s_i = r^{**} + \epsilon_i, \quad \epsilon_i \sim \mathcal{N}\left(0, \sigma_{\epsilon i}^2\right), \; i = c, h \tag{4.11}$$

where ϵ_c and ϵ_h are mutually independent.[6] The variance $\sigma_{\epsilon i}^2$ can be zero, which corresponds to i having full information about r^{**}; it can also be infinity, which corresponds to i having no private information about r^{**}. Each side can only observe their own signal.

The second source of information is macroeconomic outcomes $\tilde{y}$, π and i. Observing these outcomes allows each side to extract information about the private signal of the other.

[6] In the fully dynamic model, we allow for correlated private signals as well as public signals about r-star.

The information content can be summarized easily by rearranging the equilibrium conditions (4.8) and (4.10) in terms of two sufficient statistics:

$$a_h \equiv E^h\left[r^{**}\right] + u_h = i + \sigma\tilde{y}. \tag{4.12}$$

$$a_c \equiv E^c\left[r^{**}\right] + u_c = i - \phi_\pi \pi - \phi_y \tilde{y}. \tag{4.13}$$

These are endogenous signals of expectations and of the private information contained in them. Notice that observational noise arises from demand shocks u_h and policy shocks u_c. It is worth pointing out that, by contrast, the cost-push shock u_p has no bearing on the inference problem for r-star. A cost-push shock moves the output gap and inflation in opposite directions, and agents can thus easily distinguish it from r-star movements.

4.2 Equilibrium with common knowledge learning

We can cast the mutual learning problem in terms of the two "agents" in our model—the private sector and the central bank—forming expectations about the random variable r^{**} conditional on (s_i, u_i, a_j) where $j \neq i$. The inference problem of agent i is non-trivial because a_j is endogenous to i's expectations.

Due to the Gaussian structure of the fundamentals and signals, beliefs in equilibrium will depend linearly on the signals and noises. We conjecture, and subsequently verify, that agent i's belief of agent j's expectation takes the linear form:

$$E^j\left[r^{**}\right] = \alpha_j s_j + \beta_j s_i + \gamma_j u_j + \delta_j u_i. \tag{4.14}$$

We solve agent i's signal extraction problem given this belief. Agent i's expectation is

$$E^i\left[r^{**}\right] = E\left[r^{**} \mid s_i, u_i, a_j\right]$$

where $a_j = E^j\left[r^{**}\right] + u_j$ with $E^j\left[r^{**}\right]$ given by (4.14). Under agent i's beliefs, the vector (r^{**}, s_i, u_i, a_j) is normally distributed. The optimal filtering solution is:

$$E_i\left[r^{**}\right] = g_{si} s_i + g_{ai}\left(a_j - \beta_j s_i - \delta_j u_i\right). \tag{4.15}$$

The gain parameters are given by the following expression:

$$\begin{pmatrix} g_{si} \\ g_{ai} \end{pmatrix} = \frac{1}{\alpha_j^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \left(1+\gamma_j\right)^2 \sigma_{uj}^2\left(\sigma_{\epsilon i}^2 + 1\right)} \begin{pmatrix} \alpha_j^2 \sigma_{\epsilon j}^2 + \left(1+\gamma_j\right)^2 \sigma_{uj}^2 \\ \alpha_j \sigma_{\epsilon i}^2 \end{pmatrix}. \tag{4.16}$$

In an equilibrium with rational expectations, agent i's conjecture (4.14) coincides with the actual expectation formation of agent j. Substituting (4.14) into (4.15) yields:

$$E^i\left[r^{**}\right] = g_{si}s_i + g_{ai}\left(\alpha_j s_j + \left(1+\gamma_j\right)u_j\right). \tag{4.17}$$

Indeed, this has the functional form of the conjecture in (4.14). Comparing coefficients for $i = c, h$ yields the following equilibrium conditions:

$$\begin{pmatrix} \alpha_i \\ \beta_i \\ \gamma_i \\ \delta_i \end{pmatrix} = \begin{pmatrix} g_{si} \\ g_{ai}\alpha_j \\ 0 \\ g_{ai}\left(1+\gamma_j\right) \end{pmatrix}. \tag{4.18}$$

The equilibrium expectation under common knowledge learning is thus given by:

$$E^i\left[r^{**}\right] = g_{si}s_i + g_{ai}\left(g_{sj}s_j + u_j\right). \tag{4.19}$$

The equilibrium conditions (4.18) are a non-linear system of equations because g_{si} and g_{ai} depend on α_j through (4.16). The following proposition shows that an equilibrium always exists[7] and that the parameters of the reaction functions are bounded.

Proposition 1 (Common knowledge learning)**.** *The equilibrium defined by* (4.16) *and* (4.18) *exists and satisfies* $0 \leq g_{si} \leq 1$ *and* $0 \leq g_{ai} < 1$*. Furthermore,* $g_{si} = 1$ *if and only if* $\sigma_{\epsilon i} = 0$.

Proof. See Appendix B.1. □

Consider two special cases. The first arises when the private sector has perfect information, while the central bank has no direct source of information and must only rely on macroeconomic outcomes to infer r-star. This case underlies the empirical approach of filtering r-star with a macroeconomic model to gauge the neutral stance of monetary policy (e.g. Laubach and Williams (2003), Holston et al. (2017) and others). In our setting, this situation would be captured by $\sigma^2_{\epsilon h} = 0$ and $\sigma^2_{\epsilon c} = \infty$, and would lead to one-sided learning by the central bank only. The second special case is the reverse: The central bank has perfect information, while the private sector has no direct information and has to rely on the central bank's policy actions to infer r-star. This situation gives rise to the signalling channel of monetary policy Nakamura and Steinsson (2018).[8] This would be captured by $\sigma^2_{\epsilon h} = \infty$ and

[7] We can also rule out "nonfundamental" equilibria in which expectations would not conform to the form conjectured in (4.14) and instead coordinate on a sunspot variable (Benhabib et al., 2015; Chan, 2020).

[8] This special case also applies if the central bank has imperfect information, but can perfectly observe the private sector's expectation. This situation arises in Hillenbrand (2022).

$\sigma^2_{\epsilon c} = 0$, so that only the private sector would be learning.

The general case, where both agents learn about the determinants of r-star from each other, is unexplored in the literature to our knowledge. Under common knowledge learning, we can rearrange the equilibrium expression (4.17) to write agent i's expectation as a function of her observables:

$$E^i\left[r^{**}\right] = \left(1 - g_{ai}g_{aj}\right) g_{si}s_i - g_{ai}g_{aj}u_i + g_{ai}\left(E^j\left[r^{**}\right] + u_j\right). \tag{4.20}$$

This general case is depicted in the left panel of Figure 3. Both sides have useful private information about r-star and try to learn from each other. The red line traces the central bank's estimate $\hat{r}^*$ as a function of r^*. It has a positive slope g_{ac}. Intuitively, when r^* increases, the central bank observes higher output and inflation for a given level of the nominal interest rate and revises its own estimate $\hat{r}^*$ upwards. The blue line traces the private sector belief r^* as a function of $\hat{r}^*$. It has a positive slope $1/g_{ah}$. Intuitively, when $\hat{r}^*$ increases, the private sector observes higher interest rates for given levels of output and inflation and revises up its own beliefs of long-term real rates. This *de facto* r-star is endogenous to monetary policy.

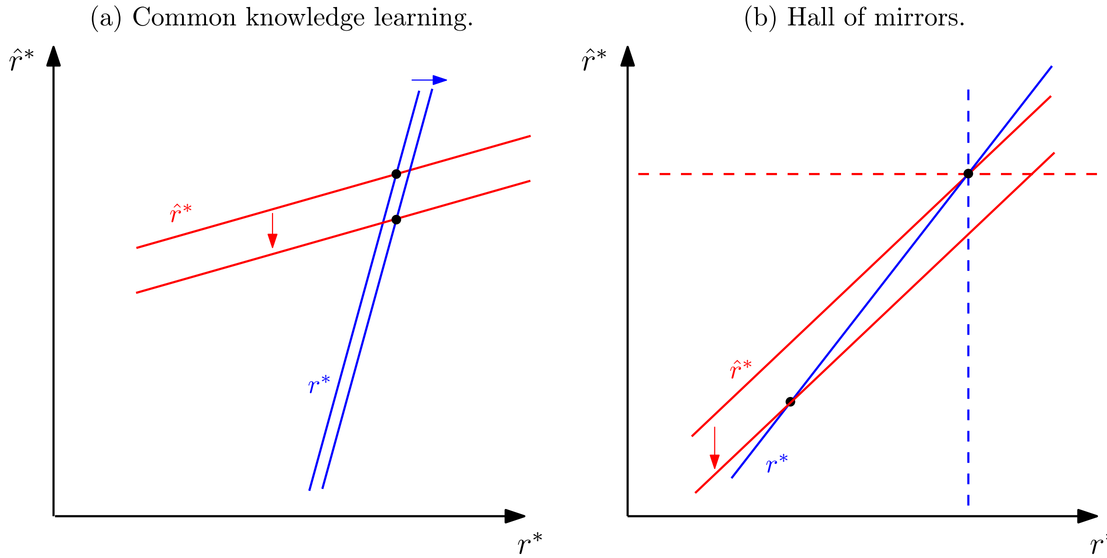
The r-star beliefs of both sides now depend on cyclical shocks. In the figure, a negative demand shock u_h shifts the red line down, resulting in lower $\hat{r}^*$. The central bank cannot readily distinguish between cyclical and permanent economic forces, and as a result assigns some weight to the possibility of a reduction in the natural interest rate. At the same time, the private sector observes the demand shock and knows that it has no bearing on r-star. It rationally corrects its reaction function in anticipation of the decline in $\hat{r}^*$: The blue line shifts slightly to the right. In equilibrium, r^* is unchanged, as shown in equation (4.17). Thus, only the central bank's expectations are affected by cyclical demand shocks. Likewise, only the private sector expectations are affected by cyclical monetary policy shocks under common knowledge learning, as the central bank correctly filters out the effects of its own shock on private sector expectations.

When both sides learn, agents can sometimes mistake a cyclical shock for an r-star shock, as they cannot differentiate the two from the others' perspectives. But the magnitude of this mistake is limited in equilibrium because each side is fully aware that the other side is also learning and correctly understands the other's reaction function.

4.3 Hall-of-mirrors equilibrium

We now consider the case where each side is unaware of the fact that the other is also learning from itself. Specifically, each side believes that the other side relies exclusively on

Figure 3: Determination of expectations with two-sided learning.



Note: Each panel shows the central bank's estimate of r-star $\hat{r}^* \equiv E^c[r^{**}]$ as a function of the noisy observation of the private sector expectation $a_h \equiv r^* + u_h$ (red lines), and the private sector's *de facto* r-star $r^* \equiv E^c[r^{**}]$ as a function of the noisy observation of the central bank expectation $a_c \equiv \hat{r}^* + u_c$ (blue lines). Arrows depict the change in beliefs after a negative shock to demand u_h.

its private signal when forming expectations about r-star. In reality, both sides are learning from each other. Ignoring the two-sided learning problem amounts to a deviation from fully rational expectations. The central bank tries to learn r-star from private sector actions, but does not take into account the fact that the underlying private sector's expectations are endogenous to monetary policy actions. Likewise, the private sector tries to learn r-star from observed monetary policy, but does not take into account the endogeneity of the central bank's r-star beliefs to demand fluctuations. Each side thus mistakes the other's reaction to its own actions for genuine information, generating a feedback loop that distorts information aggregation and amplifies noise. We call such mechanism the *hall-of-mirrors effect*.

Each agent $i = \{c, h\}$ now mistakenly believes that the expectation of agent $j \neq i$ only depends on agent j's private signal:

$$E_j^{|i}[r^{**}] = g_{sj}^{|i} s_j \tag{4.21}$$

where the perceived optimal gain is $g_{sj}^{|i} = 1/\left(1 + \sigma_{sj}^2\right)$. As a result, agent i perceives the endogenous signal in her own signal extraction problem to be $a_j = g_{sj}^{|i} s_j + u_j$. The solution of agent i's filtering problem is given by the following modification of (4.20):

$$\hat{E}_i[r^{**}] = g_{si} s_i + g_{ai}\left(\hat{E}_j[r^{**}] + u_j\right). \tag{4.22}$$

Here, $\hat{E}^i[r^{**}]$ denotes the subjective expectation, which differs from the rational expectation. The gain coefficients are still determined by Equation (4.16), but where the true equilibrium parameter α_j is replaced by the perceived parameter $g_{sj}^{|i}$. This reaction function differs from that under common knowledge learning in two important ways. First, the gain parameters are different from those in (4.20) because agents misjudge the informativeness of the endogenous signals. Second, the correction for feedback is missing here because agents are not aware of the fact that both sides are learning.

When the central bank and the private sector ignore the two-sided learning problem, neither has correct beliefs about how expectations are determined in equilibrium. To find this equilibrium, we can repeatedly substitute (4.22) into itself, switching the roles of i and j, and obtain:

$$\begin{aligned}\hat{E}_i[r^{**}] &= g_{si}s_i + g_{ai}\left(g_{sj}s_j + g_{aj}\left(\hat{E}_i[r^{**}] + u_i\right) + u_j\right) \\ &= g_{si}s_i + g_{ai}(g_{sj}s_j + u_j) + \underbrace{\sum_{k=1}^{\infty} g_{ai}g_{aj}\left(g_{si}s_i + u_i + g_{ai}(g_{sj}s_j + u_j)\right)}_{\text{hall-of-mirrors effect}}. \end{aligned} \tag{4.23}$$

Compared to the corresponding expression (4.19) under rational expectations, there now appears an extra term that captures an informational feedback loop and represents the hall-of-mirrors effect. Consider a temporary demand shock $u_h = 1$. The private sector does not react to this shock, knowing that it is unrelated to r^{**}. The expectation of the central bank, however, rises by g_{ac} and the policy rate rises by the same amount. Because the private sector does not take into account that the central bank is learning, it is surprised by the tightening in policy and partly attributes this observation to the central bank having received information that r-star is higher. The private sector thus adjusts its expectation by $g_{ah}g_{ac}$. A a result, aggregate demand strengthens, which leads the central bank to further adjusts its expectations by $g_{ac}g_{ah}^2$, resulting in a further adjustment $g_{ac}^2 g_{ah}^2$ of the private sector, and so on. In equilibrium, these effects give rise to the sum in (4.23). Because the gain parameters g_{ai} and g_{aj} are always strictly below one, the equilibrium is still unique.

We can characterize the hall-of-mirrors equilibrium as follows.

Proposition 2 (Hall-of-mirrors effect). *If σ_{uc} and σ_{uh} are sufficiently large, then the hall-of-mirrors equilibrium has the following properties relative to the common knowledge learning equilibrium:*

1. *The weight on private information g_{si} is lower and the weight on the other side's expectations g_{ai} is higher, $i = \{c, h\}$.*

2. *Equilibrium expectations overreact to cyclical demand and monetary policy shocks u_h and u_c.*

Proof. See Appendix B.2. □

The hall-of-mirrors effect is graphically illustrated in the right panel of Figure 3. The slope of the central bank's reaction function (red solid line) is g_{ac} and the slope of the private sector's reaction function (blue solid line) is $1/g_{ah}$. These slopes are closer to one than in the left panel as each side overweighs the acquisition of information through the other side's expectations. More importantly, each side also thinks that the reaction function of the other side is completely insensitive to its own expectations (dashed lines). The panel shows how the equilibrium adjusts to a negative demand shock u_h. The central bank's reaction function shifts down, as the central bank revises its r-star estimate downward and cuts interest rates. But unlike with common knowledge learning, the private sector's reaction function does not shift to account for the informational feedback. As a result, the private sector lowers its own estimate of r-star, prompting a fall in output and inflation. The central bank interprets this demand shortfall as a further indication that r-star has fallen and lowers its own estimate, and so on. This process continues until the new equilibrium is reached with a much lower r-star. The larger fall in r-star obtains because both sides misread reactions to their own actions as useful information, and are reacting to a reflection of themselves. They are staring into a hall of mirrors.

5 Dynamic model

To set the stage for a full-fledged quantitative exercise, we now turn to the dynamic version of our model. For generality, we add a public source of information observable by all agents, and also allow for a general autocorrelation structure of economic fundamentals and signal noise. The information structure and the associated qualitative equilibrium results all carry over from the static model presented above. We will show that the hall-of-mirrors effect not only amplifies noise to r-star beliefs, but also generates a very persistent deviation of the de facto r-star from its underlying fundamentals.

5.1 Fundamentals and exogenous signals

In the dynamic model, the central bank and the private sector need to form expectations of the fundamental determinants of real interest rates, $r_t^{**} = \sigma g_t + z_t$, which form a random walk process

$$r_t^{**} = r_{t-1}^{**} + v_t,\ v_t \sim \mathcal{N}\left(0, \sigma_r^2\right) \tag{5.1}$$

with $\sigma_r^2 = \sigma^2\sigma_g^2 + \sigma_z^2$.

At the start of each period t, the central bank and the private sector $(i = c, h)$ receive privately observed signals s_{it} about the fundamentals:

$$s_{it} = r_t^{**} + e_{it} \tag{5.2}$$

$$e_{it} = \rho_{ei} e_{it-1} + \epsilon_{it},\ \epsilon_{it} \sim \mathcal{N}\left(0, \sigma_{\epsilon i}^2\right). \tag{5.3}$$

In addition, we include the possibility of both agents observing a public signal x_t of the same form:

$$x_t = r_t^{**} + f_t \tag{5.4}$$

$$f_t = \rho_f f_{t-1} + \eta_t,\ \eta_t \sim \mathcal{N}\left(0, \sigma_\eta^2\right). \tag{5.5}$$

Apart from these signals, there are three transient macroeconomic shocks in the model: The demand shock u_{ht}, the cost-push shock u_{pt} and the monetary policy shock u_{ct}. Each follows an AR(1) process:

$$u_{kt} = \rho_k u_{kt-1} + \nu_{kt},\ \nu_{kt} \sim \mathcal{N}\left(0, \sigma_{uk}^2\right).$$

The private sector is assumed to observe the demand and cost-push shocks u_{ht} and u_{pt} but not the policy shock u_{ct}. Meanwhile, the central bank observes the policy shock u_{ct}, but not the demand and cost-push shocks u_{ht} and u_{pt}.

We collect the vector of exogenous states in $Z_t = (r_t^{**}, e_{ht}, e_{ct}, f_t, u_{ht}, u_{pt}, u_{ct})'$ and the vector of exogenous shocks in $q_t = (v_t, \epsilon_{sht}, \epsilon_{ct}, \eta_t, \nu_{ht}, \nu_{pt}, \nu_{ct})'$. Then we can write

$$Z_t = RZ_{t-1} + q_t,\ q_t \sim \mathcal{N}\left(0, \Sigma_q\right) \tag{5.6}$$

with $R = \operatorname{diag}\left(1, \rho_{e1}, \rho_{e2}, \rho_f, \rho_{uh}, \rho_{up}, \rho_{uc}\right)$ and $\Sigma_q = \operatorname{diag}\left(\sigma_r^2, \sigma_{\epsilon 1}^2, \sigma_{\epsilon 2}^2, \sigma_\eta^2, \sigma_{uh}^2, \sigma_{up}^2, \sigma_{uc}^2\right)$.

5.2 Macroeconomic outcomes and endogenous signals

The private sector determines inflation and output according to the system of equations consisting of (3.4)–(3.6). We write this system of equations entirely from the perspective of the private sector information set:

$$\tilde{y}_t = E_t^h [\tilde{y}_{t+1}] + \frac{1}{\sigma} E_t^h [\pi_{t+1} + r_t^{**}] + \frac{1}{\sigma} (u_{ht} - i_t) \tag{5.7}$$

$$\pi_t = \beta E_t^h [\pi_{t+1}] + \kappa \tilde{y}_t + u_{pt} \tag{5.8}$$

$$i_t = \rho_i i_{t-1} + (1 - \rho_i) \left(E_t^h E_t^c r_t^{**} + \phi_\pi \pi_t + \phi_y \tilde{y}_t + E_t^h [u_{ct}] \right) \tag{5.9}$$

The private sector observes the nominal interest rate i_t in addition to its private signal s_{ht} and the public signal x_t, and it also observes the demand and cost-push shocks u_{ht} and u_{pt}. But it does not separately observe $E_t^c r_t^{**}$ and u_{ct} in the policy rule (5.9) and has to estimate these objects.

The central bank observes current inflation π_t and the output gap $\tilde{y}_t$ in addition to its private signal s_{ct} and the public signal x_t. It determines the nominal interest rate i_t according the Taylor rule as a function of inflation, the output gap, its current-period estimate of r-star, as well as the monetary policy shock u_{ct}. We assume that the coefficients ϕ_π and ϕ_y in the monetary policy rule are such that they yield a unique solution to (5.7)–(5.9) under full information.

5.3 Equilibrium with common knowledge learning

Solving the dynamic model requires keeping track of higher-order beliefs explicitly. We define the belief of order zero as the true fundamentals: $E_{it}^{(0)} [Z_t] = Z_t$. The first-order beliefs of agent $i = c, h$ are her expectations about the fundamentals: $E_{it}^{(1)} [Z_t] = E_{it} [Z_t]$. For $n \geq 1$, her $n + 1$-th order belief is defined as the belief about the n-th order belief of agent j: $E_{it}^{(n+1)} [Z_t] = E_{it} \left[E_{jt}^{(n)} [Z_t] \right]$, $j \neq i$.

For each agent $i = c, h$, we denote with X_{it} the relevant states for agent i, which are the beliefs of all orders $n = 0, 1, 2, \ldots$ of Z_t of agent j:

$$X_{it} = \left(E_{jt}^{(n)} Z_t \right)_{n=0}^{\infty}. \tag{5.10}$$

Agent i has to form beliefs about X_{it}. She enters period t with a prior belief about X_{it-1} which is distributed as $\mathcal{N}(m_{it-1}, P_i)$.[9] She then observes her own private signal s_{it}, the public signal x_t, as well as the macroeconomic outcomes $(\tilde{y}_t, \pi_t, i_t)$. Additionally, the central bank observes u_{ct}, while the private sector observes u_{ht} and u_{pt}.

[9] We assume that enough time has passed for the prior variance to reach its time-invariant level, in keeping with much of the literature.

Agent i's posterior belief takes the form

$$X_{it} \mid i, t-1 \sim \mathcal{N}\left(m_{it}, P_i\right). \tag{5.11}$$

To characterize this belief, we solve the signal extraction problem using a conjecture on the other agent's belief m_{jt}. We guess, and later verify, that agent j's belief evolves according to:

$$m_{jt} = \Phi_j m_{jt-1} + \Psi_j m_{it-1} + \Omega_j Z_t. \tag{5.12}$$

Because $X_{it} = \left(Z_t', m_{jt}'\right)'$, this guess directly gives a state space equation for agent i's filtering problem. The exact expressions are provided in Appendix C.

The observation equation of agent i is given by

$$Y_{it} = H_{zi} Z_t + H_{mi} m_{jt} = H_i X_{it}. \tag{5.13}$$

Computing the mapping H_i is more involved than in the static model, because observations of macroeconomic outcomes now depend on the fundamentals as well as first- and higher-order beliefs. For example, the output gap, which the central bank uses as an endogenous signal about r-star, depends on the private sector's expectation of the central bank's expectation of r-star through the expected path of interest rates. Appendix C solves for the macroeconomic outcomes of the model as a function of private sector beliefs and solves for the mappings H_i.

Equations (5.12) and (5.13) form a linear filtering problem, the solution of which is given by a Kalman filter. The solution to this filtering problem has the form:

$$m_{it} = A_i m_{it-1} + G_i Y_{it}. \tag{5.14}$$

After substituting (5.13), the relationship $Y_{it} = (Z_{it}', m_{it}')'$, and the guess (5.12) for m_{jtt}, the above filtering equation satisfies the guess (5.12) for agent i. We have thus found an equilibrium.

In practice, we compute the equilibrium numerically following Nimark (2008). Our algorithm starts from an initial guess for (5.12), and then repeatedly computes the corresponding filtering problem and replaces the solution (5.14) as the guess, until the solution converges to a fixed point. For the algorithm to be computationally feasible, we have to truncate the infinite sequence of higher-order beliefs contained in m_{it} to some finite level N that represents the highest order of beliefs stored in memory and assume $E_{it}^{(N+1)}\left[Z_t\right] \approx E_{it}^{N}\left[Z_t\right]$. We have found this approximation to be valid when $N = 12$ in our simulations; increasing N further leaves the equilibrium dynamics virtually unchanged.

5.4 Hall-of-mirrors equilibrium

As in the static version of our model, we can compute the corresponding equilibrium when each agent is ignorant that the other is also learning. Specifically, each agent i believes that the other agent j does not make use of the endogenous macroeconomic signals when forming expectations.

Agent i's perceived law of motion of beliefs, we first run the same algorithm used to compute the common knowledge learning equilibrium, but where we remove the endogenous signals $(\tilde{y}_t, \pi_t, i_t)$ from agent j's observation equation (5.13). The results of this computation is the evolution of beliefs *as perceived* by agent i. The associated optimal filtering problem for agent i given this perceived evolution of beliefs takes the form in (5.14) but with different coefficient matrices that we denote with $A_i^{|i}$ and $G_i^{|i}$. From the solution, we can also recover signal matrices $H_{mj}^{|i}$ and $H_{zj}^{|i}$, which represent the signals that are available to agent j, even though agent i thinks that some of them are ignored by agent j. Appendix C describes this solution in more detail.

To proceed from the perceived law of motion to the equilibrium, we then write the filtering equation (5.14) as:

$$\begin{aligned} m_{it} &= A_i^{|i} m_{it-1} + G_i^{|i}\left(H_{Zi}^{|j} Z_t + H_{mi}^{|j} m_{jt}\right) \\ &= A_i^{|i} m_{it-1} + G_i^{|i} H_{Zi}^{|j} Z_t + G_i^{|i} H_{mi}^{|j}\left(A_j^{|j} m_{jt-1} + G_j^{|j}\left(H_{Zj}^{|i} Z_t + H_{mj}^{|i} m_{it}\right)\right) \\ &= \left(I - G_i^{|i} H_{mi}^{|j} G_j^{|j} H_{mj}^{|i}\right)^{-1}\left(A_i^{|i} m_{it-1} + G_i^{|i} H_{mi}^{|j} A_j^{|j} m_{jt-1} + G_i^{|i}\left(H_{Zi}^{|j} + H_{mi}^{|j} G_j^{|j} H_{Zj}^{|i}\right) Z_t\right). \end{aligned} \tag{5.15}$$

To obtain the first line, we substitute out the signals Y_{it} that agent i receives, which include the endogenous macroeconomic signals. To obtain the second line, we re-apply the first equality with the roles of i and j reversed. The third line collects terms. The resulting expression for m_{it} describes the dynamic hall-of-mirrors equilibrium.

5.5 Calibration

We calibrate the dynamic model to explore the quantitative implications of the hall-of-mirrors effect. Table 1 shows the calibrated model parameters. Macroeconomic coefficients pertaining to the IS curve, the Phillips curve and the policy rule are standard in the literature (e.g. Billi, 2011). In particular, the standard deviations of the cyclical shocks imply unconditional volatility of the output gap and inflation that matches the data over the last several decades. The process for r-star fundamentals r^{**} is chosen to be consistent with prior studies. The standard deviation of annual changes to r_t^{**} is set to 0.2 percent, in

Table 1: Calibrated parameters.

Parameter	Symbol	Value	Parameter	Symbol	Value
Inverse EIS	σ	6	Initial value of r_t^{**}	r_0^{**}	2.35 %
Phillips curve slope	κ	0.015	S.d. of r_t^{**} shock	σ_r	0.05
Discount factor	β	0.9941	Steady-state inflation	π^*	2 %
Rule coefficient on inflation	ϕ_π	1.5	Autocorr. of cost-push shock	ρ_p	0.8
Rule coefficient on output gap	ϕ_y	0.125	S.d. of cost-push shock	σ_p	0.05
Rule coefficient on lagged rate	ρ_i	0.5	Autocorr. of public signal noise	ρ_f	0.8
Autocorr. of policy shock	ρ_c	0.5	Autocorr. of private signal noise	ρ_{ei}	0.8
S.d. of policy shock	σ_c	0.1	S.d. of public signal noise	σ_η	0.9
Autocorr. of demand shock	ρ_h	0.8	S.d. of private signal noise	$\sigma_{\epsilon i}$	0.9
S.d. of demand shock	σ_h	0.17			

line with estimates by Holston et al. (2017).

Parameters governing the information flow about r-star fundamentals lack clear empirical analogues, necessitating some subjectivity in calibration. To gauge the potential strength of the hall-of-mirrors mechanism, we deliberately set these parameters to obtain significant learning gains from observing the other's actions. We set the autocorrelation parameters and the standard deviations of the signal noises to $\rho = 0.8$ and $\sigma = 0.9$—identically for the public signal as well as the private signals of central bank and the private sector. The amount of autocorrelation implies that telling noise and signal apart takes considerable time. The standard deviation is chosen such that, in our main simulation exercise aimed at capturing the period between 2008 and 2020, r-star estimates in the hall-of-mirrors equilibrium decline by about 1.5 percentage points while the true fundamentals r^{**} stay unchanged. Despite this rather aggressive calibration that attributes the entire fall in r-star estimates during this period to the hall-of-mirrors effects, our values imply reasonable amounts of subjective uncertainty about r-star. The subjective standard deviation of r^{**} estimates is 1.4 percent at an annual rate, which is very close to the standard error filtered r-star estimates in Holston et al. (2017) of 1.5 percent.

In Appendix E, we will examine the sensitivity of our results to alternative parameter choices. Broadly speaking, the strength of the hall-of-mirrors effects depends on several factors. First, it is increasing in the volatility of r-star fundamentals, because greater uncertainty promotes learning from each other's actions. Second, it is decreasing in the volatility of macroeconomic shocks, because those shocks act as noise that makes it harder to learn about r-star from each other. Third, it is a non-monotonic function of the private signal quality. If the private signals are very uninformative (so that the other side knows nothing useful) or extremely informative (so that one's own information suffices to know the true state), then there are no incentives to learn from each other. These incentives, and therefore the strength of the hall-of-mirrors effect, are maximal with moderately informative information that is dispersed across agents.

6 Simulation results

6.1 GFC counterfactuals

Our main simulation exercise focuses on the decade following the GFC, a period commonly associated with a notable decline in the natural interest rate, large adverse demand shocks, and extraordinary monetary policy accommodation. To illustrate the potential strength of the hall-of-mirror effect, we adopt the stark assumption that the underlying r-star fundamentals (hence the true r-star, r^{**}) stay constant throughout the post-GFC period. Shifts in the perceived r-star are then entirely due to learning.

In this simulation, all realized shocks are set to zero with the exception of the demand and cost-push shocks u_{ht} and u_{pt}. For these two shocks, we choose a sequence of realizations so that, in the hall-of-mirrors equilibrium, the paths of the output gap and inflation exactly replicate the time series of the CBO output gap and core PCE inflation in the data over the period 2007:Q4–2019:Q1. We then simulate the model under three informational assumptions. Under *full information*, the true r-star r^{**} is fully observable by everyone. Under *common knowledge learning*, both the central bank and the private sector have incomplete information about r^{**}, learn from each other and are aware that the other side is also learning. Finally, under the *hall-of-mirrors* case, both agents learn about r^{**} from each other, but neither internalises the fact that the other is also learning.

Figure 4 compares the simulation results across the three cases. Under full information (labeled FI, dotted lines), the central bank and the private sector share the same correct belief that r^{**} remains unchanged. Contractionary cyclical shocks thus have no bearing on r-star expectations, and have standard effects on the output gap and inflation. In response, the central bank lowers the policy rate using the model's Taylor-type policy rule. In this case, the rule has a constant r-star term.

In the common knowledge learning case (labeled CK, dashed lines), the central bank does not know the sources of shocks behind the observed declines in output and inflation. It attributes part of these declines to a negative r-star shock, thus lowering its r-star estimate (dashed red line, upper left panel). Meanwhile, the private sector understands that the central bank is learning and that changes in monetary policy are induced entirely by cyclical shocks, hence keeps its r-star estimate unchanged (dashed blue line). With the r-star beliefs of the private sector fixed, the central bank's low r-star beliefs act as an additional source of policy accommodation, lowering real interest rate relative to under full information. The resulting boost to output lifts inflation (lower panels), in turn pushing the nominal interest rate above the previous case (upper right panel).

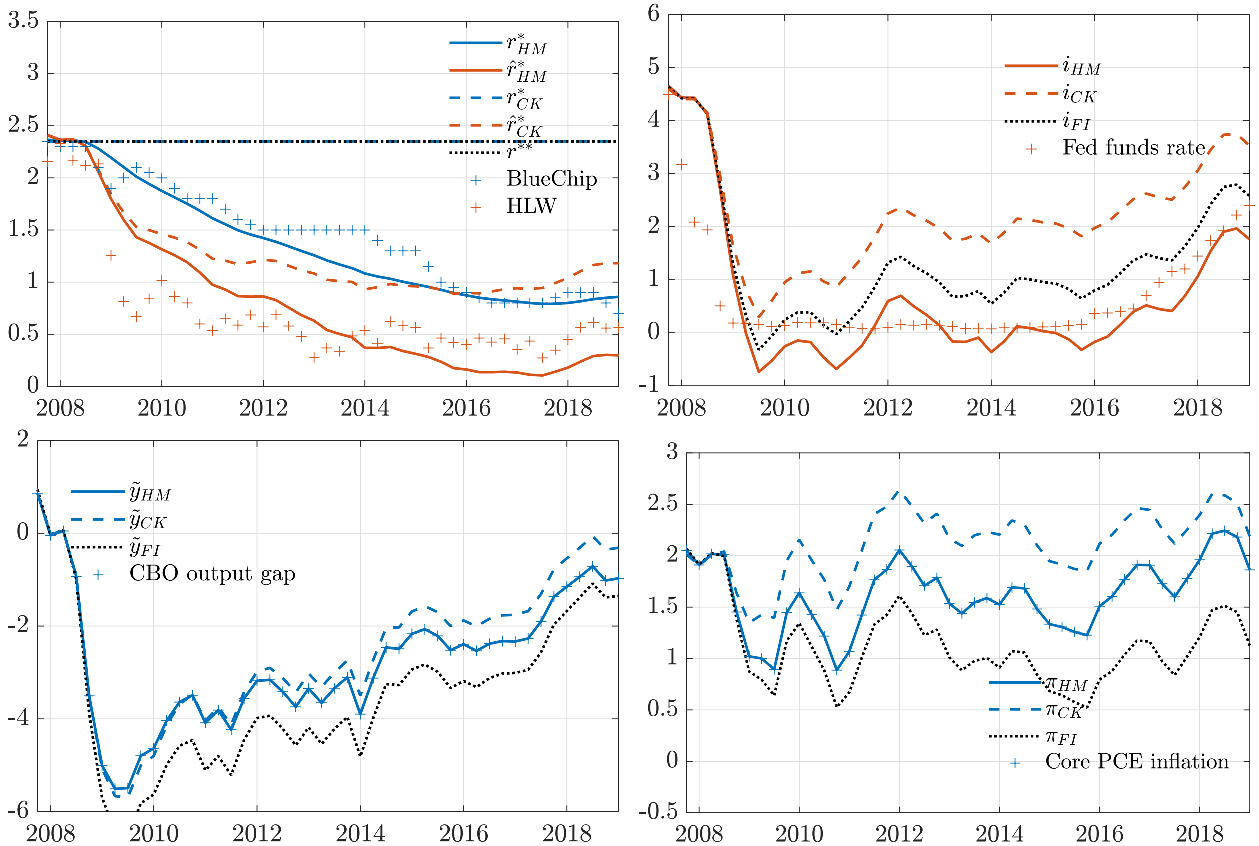
In the hall-of-mirrors case (labeled HM, solid lines), the policy rate stays lower for longer relative to the previous two cases. With neither agent aware that the other is learning from itself, the adverse cyclical shocks set in motion highly persistent self-reinforcing changes in both agents' beliefs about r-star trends. As before, the central bank rationalises the observed output fall partly by a lower r-star and cuts the policy rate more aggressively than the Taylor rule's response to inflation and the output gap implies. What is new is that the private sector, now unaware of the central bank's learning, interprets this policy action as partly reflecting the central bank having received information about r-star falling. It thus revises down its own r-star estimate, resulting in lower output and inflation for any given level of interest rates. The central bank observes these weaker economic outcomes and further revises down its r-star estimate, leading to still lower interest rates. The private sector then revises its r-star estimate further down and so on. The positive learning feedback continues and keeps r-star from both agents' perspectives low throughout the following decade. Despite the central bank cutting interest rates more aggressively than under common knowledge, output and inflation end up being lower. The reason is precisely the lower de facto r-star, which moves the goalpost and makes it harder for monetary policy to induce an expansionary effect.[10]

The hall-of-mirrors simulation captures key features of macroeconomic data in the post-GFC period (Figure 4, cross marks). By construction, simulated output and inflation are exactly the same as in the data. More importantly, the simulated r-star series line up remarkably well with their empirical analogues. The model-generated central bank's r-star assessment closely tracks the benchmark estimate from Holston et al. (2017) (top left panel), declining by almost 2 percentage points over the decade. Similarly, the simulated private sector's r-star also matches up well with its empirical counterpart, measured by the Blue Chip consensus estimate of the real interest rate at longer horizons.[11] The simulations predict a lower r-star from the central bank's perspective than that from the private sector's viewpoint, in line with the empirical pattern. The reason is that the central bank simultaneously

[10]Note how under the hall-of-mirrors and common knowledge cases, the central bank's r-star beliefs lie below those of the private sector, implying more policy accommodation than under full information where the r-star beliefs coincide. This explains why output and inflation are lowest under full information. As we show later, this result is fragile once we take into account the zero lower bound on the nominal interest rate.

[11]The estimate is the Blue Chip Economic Indicators consensus forecast of the average value over the forecast period of the federal funds rate minus the corresponding forecast for the average annual change in the GDP price index. The forecast period covers the five-year period that begins with the first quarter of the seventh year after the survey year.

Figure 4: Simulation of r-star beliefs and economic outcomes during the post-GFC period.



Note: HM denotes the hall-of-mirrors learning, CK denotes common knowledge learning, and FI denotes full information. The simulations are based on a sequence of demand and cost-push shock that are constructed so that the paths for the output gap and inflation in the hall-of-mirrors equilibrium replicate the data series shown. Other shocks are set to zero. Parameters used for the calibration are shown in Table 1.

misreads cyclical shocks as an r-star shock and misinterprets subsequently lower output and inflation as indicating further falls in r-star. Meanwhile, the private sector is aware of the true cyclical shocks, so commits an inference error only because it misinterprets an easier monetary policy as conveying news about a lower r-star. Finally, the policy rate path matches the data remarkably well. Despite the fact that we are using a Taylor-type rule, the simulation captures the "lower for longer" policy of the time, rationalized in the model by a sustained decline in the central bank's estimate of r-star.

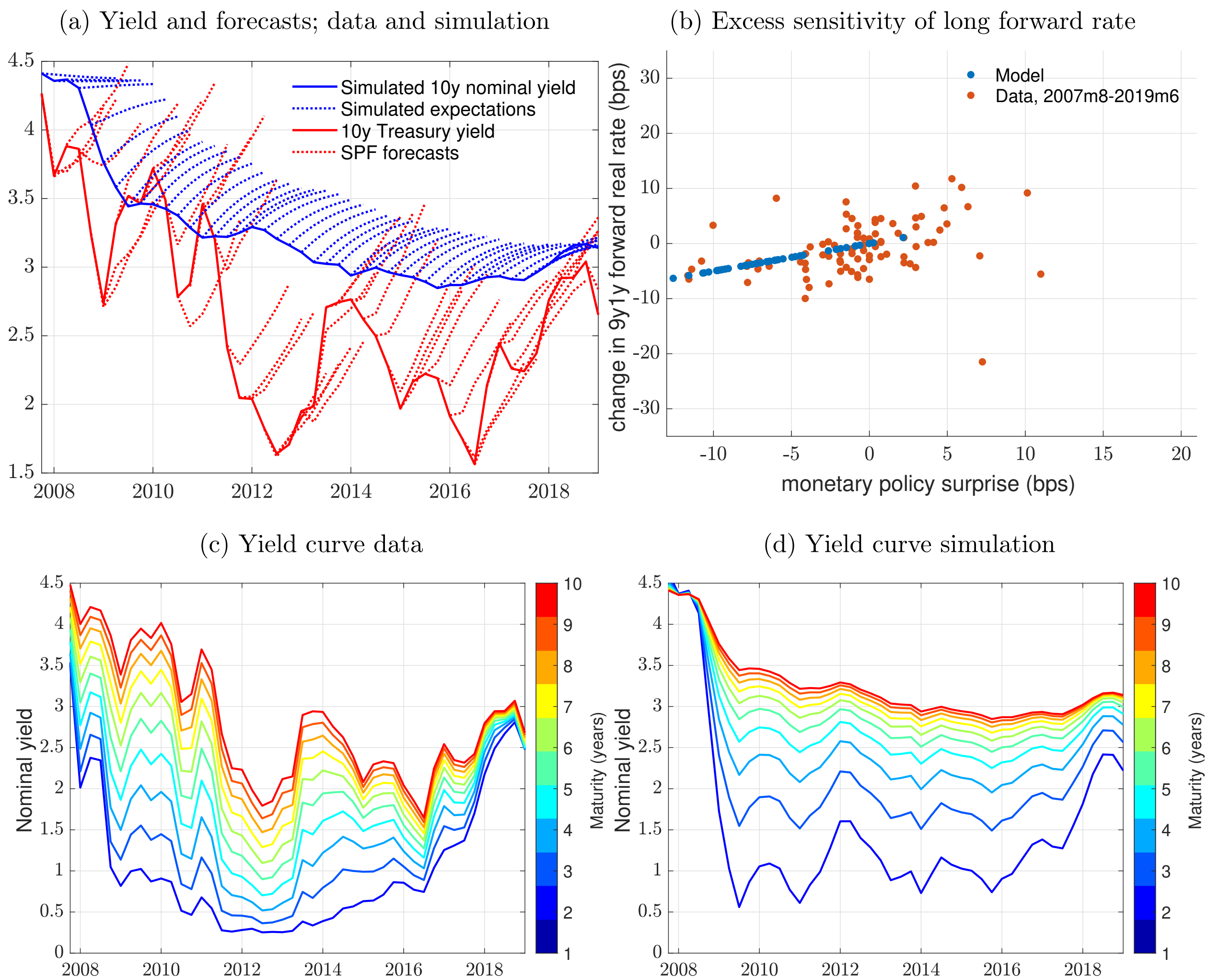
The model is also able to reproduce salient financial market patterns consistent with the stylised facts discussed in Section 2. We start by simulating zero-coupon bond yields at different maturities assuming zero term premia, as well as their projections under the hall-of-mirrors equilibrium. The top-left panel of Figure 5 shows the simulated path of ten-year nominal yields together with the private sector's projection at each point in time. We compare these with actual ten-year Treasury yields and SPF forecasts thereof during the post-GFC period. As in the data, the 10-year yield forecasts systematically fail to predict the persistent decline in actual yields. Note that expectations of interest rates *underreact*, even though expectations of r-star *overreact*. The reason is that private sector agents are unaware of the predictable overreaction of their own r-star expectations, and therefore continue to expect interest rates to recover while they keep falling. It is only when r-star expectations stabilize that the bias in interest rate forecasts all but disappears in our model.

With r-star being endogenous to cyclical developments, the model naturally rationalises the apparent violation of money neutrality seen in the data. The top right panel of Figure 5 shows that the simulated long-term (9y1y) forward real rates react to monetary policy surprises, with a slope that is remarkably similar to the empirical counterpart. Note that monetary policy surprises arise in the model even though all policy shocks u_{ct} are zero, because the private sector does not anticipate the full extent of the central bank's r-star revision following cyclical shocks u_{ht}. Cyclical shocks therefore correlate with policy surprises, consistent with the empirical evidence in Bauer and Swanson (2022). In turn, policy surprises influence the private sector's r-star beliefs, generating the correlation between policy surprises and the long forward rate shown in the top right panel.[12]

Finally, the two bottom panels of Figure 5 show that the broad movements of the yield curve match up with the data. The short-term interest rate fell sharply in the wake of the GFC, prompting a decline in longer-term interest rates. But as short-term interest rates began to rise, long-term interest rates remained relatively sticky, picking up only slowly.

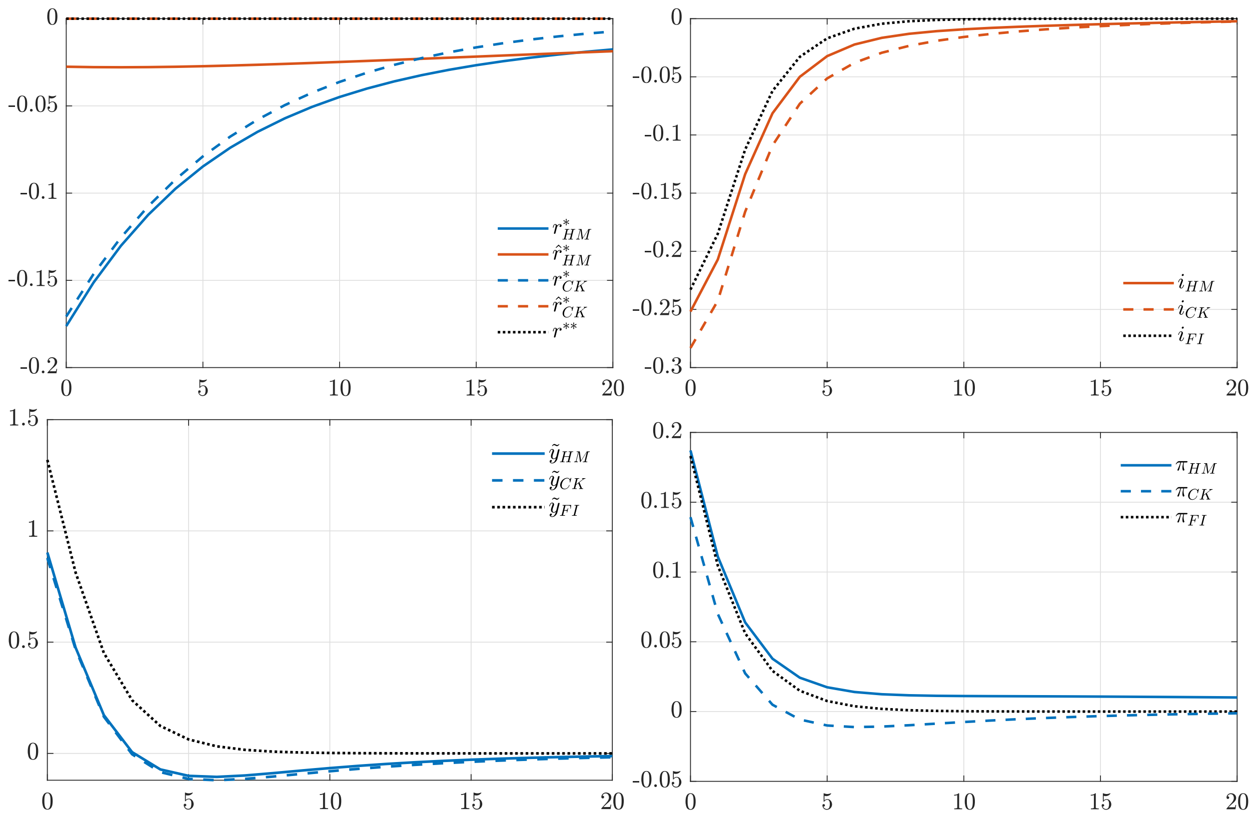
[12]In the model, macroeconomic shocks and monetary policy surprises are realized simultaneously. We use a decomposition of the overall surprise in each period, and of the corresponding movements in yields, that best isolates the policy surprise component. The construction of this decomposition is documented in Appendix D.

Figure 5: Simulation of yields and expectations during the post-GFC period.



Note: Simulation of yield curves and forward interest rates in the hall-of-mirrors equilibrium, constructed using expected paths of interest rate and assuming that the expectation hypothesis holds. For the construction of monetary policy surprises and associated yield movements see Appendix D. The simulations are based on the same sequence of shocks as those underlying Figure 4.

Figure 6: Responses to an expansionary monetary policy shock.



Note: Responses to a one-time accommodative monetary policy shock. The size of the shock is normalized to achieve a 25 basis point reduction in the policy rate i_t on impact in the hall-of-mirrors equilibrium. HM denotes the hall-of-mirrors learning, CK denotes common knowledge learning, and FI denotes full information.

These stylised facts are natural predictions of the hall-of-mirror hypothesis following a large demand shock, as the endogenous fall in r-star persists long after the shock dissipates (bottom right panel).

6.2 Monetary policy shocks

Monetary policy shocks can set off a similar chain reaction that prompts a persistent shift in r-star. Figure 6 shows the effects of an expansionary shock to the Taylor rule u_{ct}. Unexpected policy easing causes the private sector to revise down its r-star estimate, pushing the natural rate relevant for the economy lower. This revision occurs under both common knowledge learning and the hall-of-mirrors, and, as a result, output in these cases rises by less than under full information. The unexpected easing of monetary policy has the unintended effect of moving the goalpost and weakening the stimulative effect away from the full-information setup. Note that this differs from the preceding case of purely cyclical shocks where the de facto r-star does not change under common knowledge learning.

Monetary policy shocks nonetheless induce a longer-lasting r-star impact under the hall-of-mirrors case relative to the common knowledge learning. With the hall-of-mirrors, the central bank mistakenly attributes weaker-than-expected demand partly to the private sector having received information about lower r-star, not realising that it reflects the endogenous impact of monetary policy action on private sector expectations. As a result, the central bank revises down its own r-star estimate, setting in motion the information feedback that leads to persistently lower r-star assessments by both parties. In our calibration, a policy shock that lowers the policy rate by 25 basis points implies an immediate reduction in the de-facto r-star of more than 15 basis points. The impact dissipates slowly over time and more so under the hall-of-mirrors case. After 10 quarters, the private sector still believes that r-star is 5 basis points lower than actually the case.

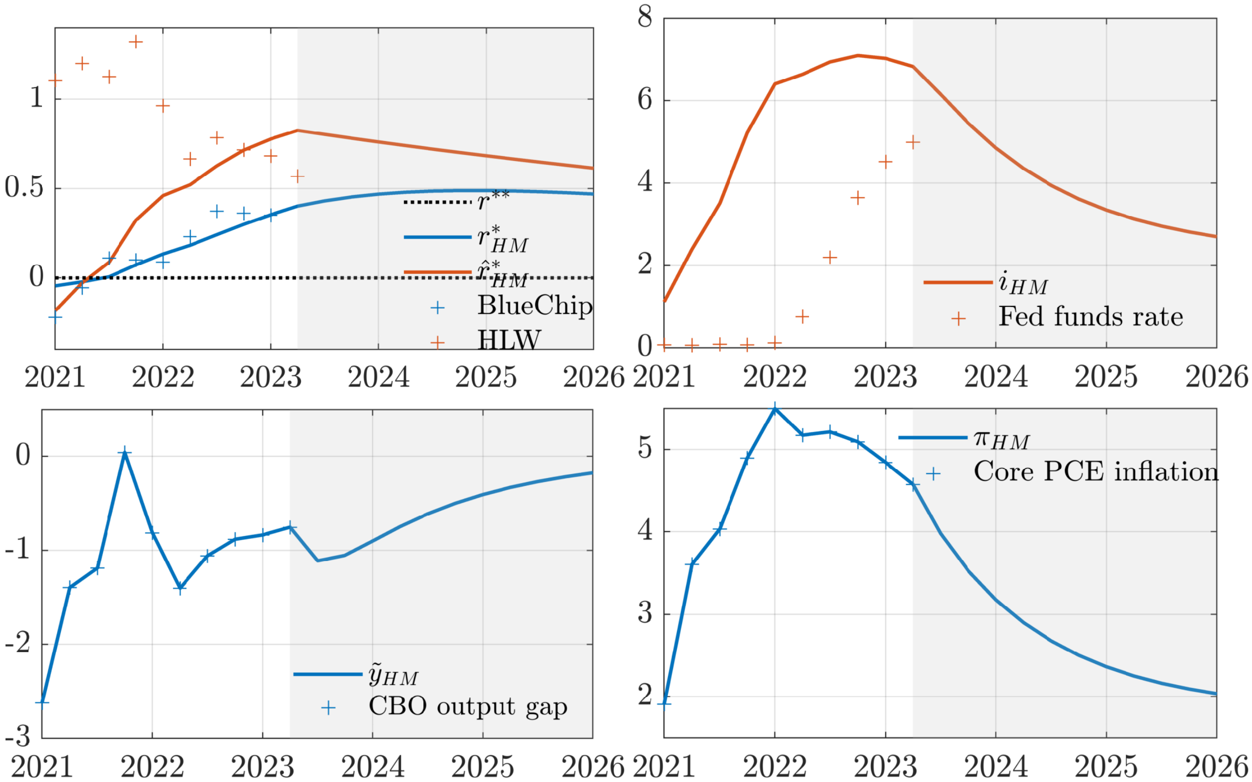
6.3 Inflationary shocks after the COVID pandemic

We now turn to the more recent period starting in 2021 when global inflation surged during the post-pandemic recovery. In the United States, core PCE inflation topped 5.5% year-over-year in February 2022, prompting a sharp monetary policy tightening by the Federal Reserve. Central banks elsewhere took similar actions in a bid to contain inflation, which globally reached its highest levels since the 1970s. The large increase in global policy rates triggered a debate about whether the trend decline in interest rates may be reversing. How does the hall-of-mirrors effect contribute to this debate?

We reprise the post-GFC exercise in the post-Covid context. Specifically, we set demand and cost-push shocks such that the hall-of-mirrors equilibrium produces the inflation and output gap paths as observed in 2021Q1–2023Q2 (holding all other shocks at zero). We assume initial beliefs align with underlying fundamentals, and keep the full-information r-star constant at zero throughout the simulation, i.e. $r_0^{**} = 0$. The choice of this value serves to isolate the effects of the cyclical shocks on r-star beliefs. We also run the simulation for a few more years without shocks to illustrate the model-implied persistence of changes in r-star beliefs.

As Figure 7 shows, the simulation predicts an uptick in r-star expectations during the post-pandemic recovery. In response to the demand and cost-push shocks, private sector r-star beliefs rise by about half a percentage point between 2021 and 2023 (upper left panel), while those of the central bank rise somewhat more. The hall-of-mirrors mechanism operates in the same way as during the GFC period, but in reverse. The central bank and the private sector attribute the causes of inflation surge partly to an increase in r-star, a set of beliefs

Figure 7: The hall-of-mirrors effect in an inflationary scenario.



Note: HM denotes the hall-of-mirrors learning. Demand and cost-push shocks from 2021Q1 through 2023Q2 are constructed so that the paths for the output gap and inflation in the hall-of-mirrors equilibrium replicate the data series shown. Other shocks are set to zero. Parameters used for the calibration are as in Table 1, except for r_t^{**} which is constant at zero to isolate the effects of cyclical shocks. Starring in 2023Q3 (shaded areas), the model is simulated without shocks.

that get entrenched by rising inflation and higher interest rates. Even as shocks dissipate going forward, the model predicts r-star beliefs to stay persistently above their initial levels by about half a percentage point after a few years.[13]

Comparing with the data, the simulated private sector r-star beliefs align well with the Blue Chip survey as shown in the upper left panel, both rising by about half a percentage point during 2021-2022 (market-implied proxies, such as the risk-neutral forward rate, rose by more during this period). The simulated central bank's r-star is met with a mixed success, however, perhaps reflecting estimation challenges during the pandemic. The estimate of Holston et al. (2017) declined slightly over this period (shown in the graph), while other estimates have risen notably after the pandemic, including Lubik and Matthes (2015) and the FOMC medium-term dot plot. The path of the model-implied policy rate was also different from that of the federal funds rate especially in 2021 (top right panel), though this owes to the Federal Reserve's commitment to keep the policy rate low after the pandemic shock which is a departure from the usual policy rule.

The simulation shows that the apparent increase in long-term yields and r-star expectations could partly be a mirage caused by learning feedback unrelated to structural features of the economy. Similar to how adverse demand shocks may have kept r-star persistently low in the post-GFC period, the recent inflationary shocks may well prompt a reversal of this trend.

7 Monetary policy implications

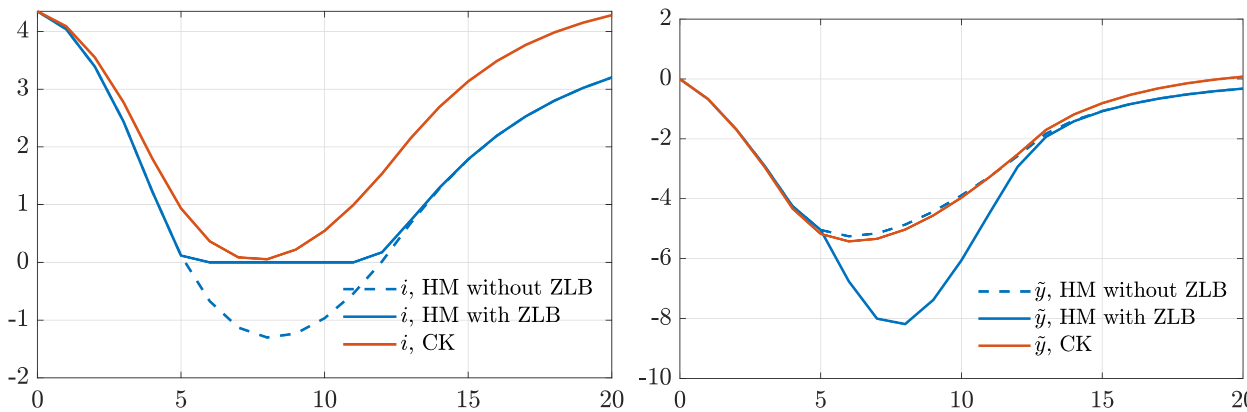
The endogenous interactions between r-star and monetary policy have important policy implications. In this section, we consider the costs of an endogenous decline in r-star in the presence of a lower bound on nominal interest rates. We discuss how the conduct of monetary policy can amplify or mitigate the hall-of-mirrors effect.

7.1 Endogenous ELB episodes

A decline in r-star raises deflationary concerns because, when the nominal interest rate is constrained by effective lower bound (ELB), the ability of central banks to stimulate demand is diminished. In response to these concerns, policymakers have attempted to intensify the degree of accommodation by other means, e.g. unconventional monetary policy measures or

[13] If the true r-star fundamentals were higher than the initial beliefs, subsequent updated beliefs would drift upwards over time to narrow the initial gap. In Appendix F, we simulate the model over the entire period 2007:Q4–2023:Q2 on the assumption that true r-star fundamentals remain constant at their pre-GFC level of 2.4 percent. In that exercise, r-star beliefs after the pandemic rise to about 1.2 percent.

Figure 8: The hall-of-mirrors effect in the presence of an ELB constraint..



Note: Simulation based on a sequence of 14 quarters of negative demand shocks. "HM" denotes the hall-of-mirrors equilibrium. "CK" denotes the common knowledge learning equilibrium.

communications to keep rates low even after adverse shocks abate. In our setup, however, policy accommodation can induce an endogenous fall in the de facto r-star and aggravate the very problem that it is trying to solve.

To illustrate this problem, we introduce an ELB constraint to our model. For tractability, we make the simplifying assumption that both the private sector and the central bank continue to receive the same information about each other's expectations whether the ELB is binding or not.[14]

In Figure 8, we simulate the response of the model to a sequence of negative cyclical demand shocks u_{ht} that are designed to push the nominal interest rate against the ELB for a prolonged period of time. The figure reveals that the macroeconomic costs of the adverse shocks are distinctly higher in the hall-of-mirrors equilibrium compared to the case of common knowledge learning. In the latter case (red lines), r^* remains unchanged and the interest rate stays just above the ELB: The central bank is left with enough policy space to counter the recession according to its interest rate rule. With the hall-of-mirrors effect (solid blue lines), however, r-star endogenously falls, reducing policy space and resulting in a much more severe recession as the ELB binds for about eight quarters. The reduction in r-star occurs *because* the central bank responds to the adverse shocks by lowering the policy rate, and the private sector misinterprets this accommodation as a signal that r-star has fallen.

The large macroeconomic costs result from the interaction of the hall-of-mirrors effect

[14]For the private sector, the assumption amounts to the shadow rate of interest—the nominal interest rate the central bank would like to set absent the ELB—being observable, e.g. through policy communications. Bodenstein et al. (2022) show that learning can be much slower when the shadow rate is unobserved. For the central bank, the assumption amounts to output and inflation carrying the same information about r-star regardless of whether the ELB is binding or not. Many empirical models estimating r-star, including Laubach and Williams (2003), embed this assumption.

Table 2: Effect of central bank behavior on r-star misperception.

Standard deviation	HH expectational errors $r_t^* - r_t^{**}$	CB expectational errors $\hat{r}_t^* - r_t^{**}$	Disagreement $r_t^* - \hat{r}_t^*$
Baseline	1.57	1.51	0.50
Direct communication of $\hat{r}^*$	1.59	1.54	0.49
First-difference rule	1.54	1.43	0.81
$E_{ct}[r^*]$ instead of $\hat{r}^*$	1.70	1.51	0.82
No CB misperception	1.46	1.44	0.59
Common knowledge	1.41	1.45	0.49
Full information	0.00	0.00	0.00

Note: Percent annual rates. Standard deviations are computed by simulating the model for 100,000 periods with random shocks. "Direct communication of $\hat{r}^*$" refers to a variant of the model where the central bank reveals its r-star estimate $\hat{r}_t^*$ without noise to the private sector. "$E_{ct}[r^*]$ instead of $\hat{r}^*$" refers to a variant where the Taylor rule intercept is $E_{ct}E_{ht}r^{**}$ instead of $E_{ct}r^{**}$. "First-difference rule" refers to a variant where the central bank follows a first-difference rule which does not rely on $\hat{r}_t^*$. "No CB misperception" refers to a variant where the central bank has rational expectations while the private sector suffers from misperception as in the baseline model.

and the ELB. Without the ELB (dashed blue lines), the reduction in output is very similar to that under common knowledge learning as the policy rate falls in accordance with the lower r-star path.

7.2 Monetary policy rules and communication

How does the conduct of monetary policy influence the hall-of-mirrors effect? The effect is caused by both the private sector and the central bank not internalizing that they are learning from each other. An obvious remedy is for central banks to fully internalize the extent to which the private sector's r-star expectations are shaped by policy actions. In our model, this leads the central bank to rely more on their own information about economic fundamentals to judge the appropriate level of interest rates in the longer run, and less on measures of private sector expectations. Another possible mitigating factor is a greater awareness by the central bank that monetary policy and its communication can have an unintended effect on the formation of public expectations. In the r-star context, more communication is not always better.

We arrive at these conclusions by evaluating several variants of our model via Monte Carlo simulation. For each variant, we simulate the model with random shocks and report the simulated unconditional standard deviation of three objects: The expectational error of the private sector $r^* - r^{**}$, the expectational error of the central bank $\hat{r}^* - r^{**}$, and the disagreement $\hat{r}^* - \hat{r}^*$ between the two. Table 2 shows the results.

The first row of the table shows our baseline model. The second row shows a variant in which the central bank communicates its r-star estimate publicly and without noise in addition to its use in the Taylor rule. This more transparent communication *worsens* the hall-of-mirrors problem relative to the baseline, as the expectational errors increase in magnitude. Communicating more is detrimental in this two-sided learning environment because it leads the private sector to react even more to the central bank's expectations, thereby strengthening the informational feedback loop.

The third row shows a variant where the central bank follows the first-difference rule of Orphanides and Williams (2007). This rule does not have an intercept and thus reveals no information at all about the central bank's r-star estimate.[15] Therefore, this case is the opposite of the previous one: the central bank withholds all information about r-star from the public. The table shows that expectational errors become smaller as a result. Because less information gets shared, there is also more disagreement between the central bank and the private sector. But more disagreement is not a sign of informational inefficiency in our model, because the hall-of-mirrors effect can lead the central bank and the private sector to agree on an incorrect value of r-star.

The fourth row shows a variant in which the central bank does not use its own r-star estimate $\hat{r}_t^* = E_{ct} r_t^{**}$ as the intercept of the Taylor rule (3.6), but rather its best estimate of the private sector's r-star expectation $E_{ct} r_t^* = E_{ct} E_{ht} r_t^{**}$. This is a natural alternative choice, because r^* is the de-facto r-star that determines aggregate demand. Choosing its estimate as the rule intercept could in principle improve macroeconomic stabilization. However, this choice does not alleviate the hall-of-mirrors problem. The table shows that household expectational errors are larger than in the baseline model while those of the central bank are unchanged. The volatility of inflation and the output gap (not reported) are not materially reduced. Even though the central bank's private information enters the interest rate rule less directly, the private sector still tries to learn this private information. As a result, the perils of ignoring the two-sided learning problem remain.

The fifth row shows a variant where the central bank does not suffer from misperception. Here, the central bank not only knows the true precision of all the signals in the economy, but is also aware of the private sector's misperception and corrects for it in its inference problem. As a result, the central bank effectively breaks the informational feedback loop, resulting in expectational errors that are much smaller than in the baseline simulation. Nevertheless, the private sector continues to ignore the fact that the central bank is learning, pays too much attention to the central bank's actions, and too little attention to its own private

[15] Specifically, the inertial Taylor rule (3.6) is replaced by $i_t = i_{t-1} + \phi_{pi} \pi_t + \phi_y \hat{y}_t + u_{ct}$.

information. As a consequence, the expectational errors are still somewhat larger than in the sixth row of the table, which reports common knowledge learning. The final row shows that expectational errors and disagreement are zero when r_t^{**} is publicly observed.

8 Conclusion

We extend the canonical New-Keynesian model to an incomplete information setup where both the central bank and the private sector have to learn about the determinants of r-star. In such a world, beliefs and two-way learning are an important driver of persistent changes in real interest rates. Crucially, we have shown that ignoring the double learning problem, or even underestimating its importance, can lead to large and persistent movements of r-star in response to purely cyclical shocks.

In practice, it is difficult to determine if r-star shifts come from slow-moving structural forces or endogenous two-way learning by the private sector and the central bank. But the policy implications cannot be more different. If the-hall-of-mirrors effect is indeed responsible for a trend decline in real interest rates, then extraordinary monetary policy accommodation may have been part of the problem, in moving the goalpost and raising the bar for what it takes to stimulate the economy. At the same time, the recent surge in global inflation may offer an opportunity to escape from such an endogenous liquidity trap. With the hall-of-mirrors effect working in reverse, monetary policy tightening can lead to persistently higher r-star beliefs.

The simplicity of our model comes at the expense of staying silent on important channels through which r-star beliefs could affect the economy. For example, shifts in r-star beliefs could distort investment decisions, not captured in our model without capital. Accounting for such channels in future research could shed further light on the real effects of r-star beliefs and implications for optimal monetary policy.

References

Adrian, T., R. Crump, and E. Moench (2013): "Pricing the Term Structure with Linear Regressions," *Journal of Financial Economics*, 110, 110–38.

Angeletos, G.-M., L. Iovino, and J. La'O (2020): "Learning Over the Business Cycle: Policy Implications," *Journal of Economic Theory*, 190, 105115.

Bauer, M. D. and E. T. Swanson (2022): "A Reassessment of Monetary Policy Surprises and High-Frequency Identification," Working Paper 29939, National Bureau of Economic Research.

Beaudry, P. and C. Meh (2021): "Monetary Policy, Trends in Real Interest Rates and Depressed Demand," Bank of Canada Staff Working Papers 2021-27.

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

Bernanke, B. (2004): "What Policymakers Can Learn from Asset Prices," Speech before the Investment Analysts Society of Chicago, Chicago, Illinois.

Bernanke, B. and M. Woodford (1997): "Inflation Forecasts and Monetary Policy," *Journal of Money, Credit and Banking*, 29, 653–84.

Billi, R. M. (2011): "Optimal Inflation for the US Economy," *American Economic Journal: Macroeconomics*, 3, 29–52.

Bodenstein, M., J. Hebden, and F. Winkler (2022): "Learning and Misperception of Makeup Strategies," *Journal of Economic Dynamics and Control*, 139, 104417.

Borio, C., P. Disyatat, J. Juselius, and P. Rungcharoenkitkul (2022): "Who so Low for so Long? A Long-Term View of Real Interest Rates," *International Journal of Central Banking*, 18, 47–87.

Chan, J. (2020): "Monetary Policy and Sentiment-Driven Fluctuations," working paper.

Cialdini, R. B., W. Wosinska, D. W. Barrett, J. Butner, and M. Gornik-Durose (1999): "Compliance with a Request in Two Cultures: The Differential Influence of Social Proof and Commitment/consistency on Collectivists and Individualists," *Personality and Social Psychology Bulletin*, 25, 1242–1253.

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

Cukierman, A. and F. Lippi (2005): "Endogenous Monetary Policy with Unobserved Potential Output," *Journal of Economic Dynamics and Control*, 29, 1951–1983.

Galí, J. (2015): *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications*, Princeton University Press.

Hanson, S. G., D. O. Lucca, and J. H. Wright (2018): "The Excess Sensitivity of Long-Term Rates: A Tale of Two Frequencies," Staff Report 810, Federal Reserve Bank of New York.

Hanson, S. G. and J. C. Stein (2015): "Monetary Policy and Long-Term Real Rates," *Journal of Financial Economics*, 115, 429–48.

Hillenbrand, S. (2022): "The Fed and the Secular Decline in Interest Rates," Working paper.

Holston, K., T. Laubach, and J. C. Williams (2017): "Measuring the Natural Rate of Interest: International Trends and Determinants," *Journal of International Economics*, 108, S59–S75.

——— (2023): "Measuring the Natural Rate of Interest After Covid-19," Staff Report 1063, Federal Reserve Bank of New York.

Johannsen, B. K. and E. Mertens (2021): "A Time-Series Model of Interest Rates with the Effective Lower Bound," *Journal of Money, Credit and Banking*, 53, 1005–1046.

Kearns, J., A. Schrimpf, and D. Xia (2018): "Explaining Monetary Spillovers: The Matrix Reloaded," BIS Working Papers 757.

Laubach, T. and J. C. Williams (2003): "Measuring the Natural Rate of Interest," *Review of Economics and Statistics*, 85, 1063–1070.

Lubik, T. A. and C. Matthes (2015): "Calculating the Natural Rate of Interest: A Comparison of Two Alternative Approaches," Economic Brief 15-10, Federal Reserve Bank of Richmond.

Lunsford, K. G. and K. D. West (2019): "Some Evidence on Secular Drivers of US Safe Real Rates," *American Economic Journal: Macroeconomics*, 11, 113–39.

Melosi, L. (2016): "Signalling Effects of Monetary Policy," *Review of Economic Studies*, 84, 853–884.

Mian, A., L. Straub, and A. Sufi (2020): "Indebted Demand," NBER Working Papers 26940.

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

Nakamura, E. and J. Steinsson (2018): "High-Frequency Identification of Monetary Non-Neutrality: The Information Effect," *Quarterly Journal of Economics*, 133, 1283–1330.

Nimark, K. (2008): "Monetary Policy with Signal Extraction from the Bond Market," *Journal of Monetary Economics*, 55, 1389–1400.

Orphanides, A. (2003): "Monetary Policy Evaluation with Noisy Information," *Journal of Monetary Economics*, 50, 605–631.

Orphanides, A. and J. C. Williams (2007): "Robust Monetary Policy with Imperfect Knowledge," *Journal of Monetary Economics*, 54, 1406–1435, carnegie-Rochester Conference Series on Public Policy: Issues in Current Monetary Policy Analysis November 10-11, 2006.

——— (2008): "Learning, Expectations Formation, and the Pitfalls of Optimal Control Monetary Policy," *Journal of Monetary Economics*, 55, S80–S96.

Primiceri, G. E. (2006): "Why Inflation Rose and Fell: Policy-Makers' Beliefs and U. S. Postwar Stabilization Policy*," *Quarterly Journal of Economics*, 121, 867–901.

Rungcharoenkitkul, P., C. Borio, and P. Disyatat (2019): "Monetary Policy Hysteresis and the Financial Cycle," BIS Working Papers 817.

Rungcharoenkitkul, P. and F. Winkler (2021): "The Natural Rate of Interest through a Hall of Mirrors," BIS Working Papers 974.

Tang, J. (2015): "Uncertainty and the Signaling Channel of Monetary Policy," Working Paper 15-8, Federal Reserve Bank of Boston.

Online Appendix

Appendix A New Keynesian Model with Incomplete Information

A.1 Model building blocks

Households

The representative household solves

$$\max_{C_t,N_t} E_0^h \sum_{t=0}^{\infty} \beta^t e^{\xi_t} \left(\frac{C_t^{1-\sigma}}{1-\sigma} - A_t^{1-\sigma} \frac{N_t^{1+\varphi}}{1+\varphi} \right) \tag{A.1}$$

$$\text{s.t.} \quad P_t C_t + Q_t B_t \leq B_{t-1} + W_t N_t + D_t \tag{A.2}$$

where C_t is aggregate consumption. Importantly, E^h is the mathematical expectations taken with respect to households' information set, economic model and beliefs about the relative accuracy of their information. We elaborate on this in the next section.

Solving this problem leads to the first-order conditions

$$\frac{W_t}{P_t} = C_t^{\sigma} A_t^{1-\sigma} N_t^{\varphi} \tag{A.3}$$

$$Q_t = \beta E_t^h \left[\left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \frac{P_t}{P_{t+1}} e^{\Delta \xi_{t+1}} \right] \tag{A.4}$$

plus a transversality condition. These conditions can be written in log-linear form as

$$w_t - p_t = \sigma c_t + (1-\sigma) a_t + \varphi n_t \tag{A.5}$$

$$c_t = E_t^h(c_{t+1}) - \frac{1}{\sigma} \left[i_t - E_t^h(\pi_{t+1}) - \rho + \Delta \xi_{t+1}) \right] \tag{A.6}$$

where $i_t \equiv -\log Q_t$, $\rho \equiv -\log(\beta)$, $\pi_{t+1} \equiv p_{t+1} - p_t$, and small letters denote logs of relevant variables. Households also solve a sub-problem, which arises from their preference for variety. The aggregate consumption is posited to be a CES sum of differentiated goods

$$C_t \equiv \left(\int_0^1 C_t(i)^{\frac{\epsilon-1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}} \tag{A.7}$$

Households maximise this for a given level of expenditure

$$\int_0^1 P_t(i) C_t(i) di \tag{A.8}$$

which gives rise to

$$C_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} C_t \tag{A.9}$$

and can be used to verify that equation A.8 is equal to $P_t C_t$, consistent with the budget constraint A.2.

Production

Firms $i \in [0, 1]$ produce differentiated goods, with an identical technology

$$Y_t(i) = A_t N_t(i)^{1-\alpha} \tag{A.10}$$

where we leave the process for $a_t \equiv \log(A_t)$ unspecified for now (this will be one source of uncertainty, along with z_t).

Price setting

Assume Calvo pricing where firms can re-optimise and adjust prices with probability $1 - \theta$. This gives rise to the sticky price formulation of aggregate price in a log-linearised term (around steady-state inflation of zero):

$$\pi_t = (1 - \theta)(p_t^* - p_{t-1}) \tag{A.11}$$

where p_t^* is the log price set by re-optimising firms, which must take into account how long they will remain with the price once it has been reset. This can be shown to be given by

$$p_t^* = \mu + (1 - \beta\theta) \sum_{k=0}^{\infty} (\beta\theta)^k E_t^h(\psi_{t+k|t}) \tag{A.12}$$

where $\psi_{t+k|t}$ is the log marginal cost in period $t + k$ for a firm that last reset in period t, and $\mu \equiv \log(\frac{\epsilon}{\epsilon-1})$ is desired gross markup. Note that again we use the expectations operator E_t^h, with the assumption that firms and households share the same information set, economic model, and perception about the accuracy of their signals relative to the central bank's.

Equilibrium

Consider first the derivation of the Phillips curve. The individual firm's marginal cost $\psi_{t+k|t}$ is wage at $t + k$ minus marginal product of labour in $t + k$ for a firm resetting in t. These need to be solved in the general equilibrium, and will depend on future employment, hence output and price. It can be shown that

$$\psi_{t+k|t} = \psi_{t+k} - \frac{\alpha\epsilon}{1-\alpha}(p_t^* - p_{t+k}) \tag{A.13}$$

where $\psi_{t+k} \equiv \int_0^1 \psi_t(i) di$ is the cross-sectional average marginal cost.

Combining equations A.12 and A.13, one can write p_t^* in a recursive form as

$$p_t^* = \beta\theta E_t^h(p_{t+1}^*) + (1-\beta\theta)\left(p_t - \frac{1-\alpha}{1-\alpha+\alpha\epsilon}\hat{\mu}\right) \tag{A.14}$$

where $\hat{\mu} \equiv \mu_t - \mu$ is the deviation between the average markup $\mu_t \equiv p_t - \psi_t$ and the desired markup. Plugging this into equation A.11, we get

$$\pi_t = \beta E_t^h(\pi_{t+1}) - \left(\frac{(1-\theta)(1-\beta\theta)(1-\alpha)}{\theta(1-\alpha+\alpha\epsilon)}\right)\hat{\mu}_t \tag{A.15}$$

To derive $\hat{\mu}$, note that the average markup μ_t depends on output and productivity:

$$\mu_t = \left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)(a_t - y_t)\, a_t + \log(1-\alpha) \tag{A.16}$$

Under flexible prices, $\mu_t = \mu$ obtains as firms can set markup frictionlessly. Inverting the equation leads to the natural output y_t^n definition:

$$y_t^n = a_t + \psi_y \tag{A.17}$$

where $\varphi_y \equiv -(1-\alpha)(\mu - \log(1-\alpha))/(\sigma(1-\alpha)+\varphi+\alpha)$. Thus:

$$\hat{\mu_t} = -\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)(y_t - y_t^n) \tag{A.18}$$

Substituting this into A.15 results in the New Keynesian Phillips curve

$$\pi_t = \beta E_t^h(\pi_{t+1}) + \kappa\tilde{y}_t \tag{A.19}$$

where $\tilde{y}_t \equiv y_t - y_t^n$ is the output gap and $\kappa \equiv \left(\frac{(1-\theta)(1-\beta\theta)(1-\alpha)}{\theta(1-\alpha+\alpha\epsilon)}\right)\left(\sigma + \frac{\varphi+\alpha}{1-\alpha}\right)$.

To derive the IS curve, note that the goods market equilibrium condition

$$Y_t = C_t \tag{A.20}$$

implies that the Euler equation A.6 is given by

$$y_t = E_t^h(y_{t+1}) - \frac{1}{\sigma}\left[i_t - E_t^h(\pi_{t+1}) - \rho + \Delta\xi_{t+1}\right] \tag{A.21}$$

We use the following specification for productivity process a_t and that of the stochastic discount factor ξ_t:

$$
\begin{aligned}
a_t &= a_{t-1} + g_{t-1} + \epsilon_{y^*t} \\
g_t &= g_{t-1} + \epsilon_{gt} \\
\xi_t &= \xi_{t-1} - z_{t-1} - \sigma u_{ht-1} - \epsilon_{\xi t} \\
u_{ht} &= \rho_h u_{ht-1} + \epsilon_{ht} \\
z_t &= z_{t-1} + \epsilon_{zt}
\end{aligned}
$$

We can now rewrite the Euler equation in terms of output gaps to make r-star appear:

$$
\begin{aligned}
\tilde{y}_t &= E_t^h(\tilde{y}_{t+1}) + E_t^h(y_{t+1}^n) - y_t^n - \frac{1}{\sigma}\left[i_t - E_t^h\pi_{t+1} + E_t^h\Delta\xi_{t+1} - \rho\right] && \text{(A.22)} \\
&= E_t^h(\tilde{y}_{t+1}) + \psi_{ya}E_t^h\Delta a_{t+1} - \frac{1}{\sigma}\left[i_t - E_t^h\pi_{t+1} + E_t^h\Delta\xi_{t+1} - \rho\right] && \text{(A.23)} \\
&= E_t^h(\tilde{y}_{t+1}) - \frac{1}{\sigma}\left[i_t - E_t^h\pi_{t+1} - E_t^h(r_t^*)\right] + \varepsilon_t^y && \text{(A.24)}
\end{aligned}
$$

where the natural interest rate is defined by

$$
r_t^* = \rho + \sigma g_{t+1} + z_{t+1} \tag{A.25}
$$

The model is closed by the Taylor rule

$$
i_t = E_t^c(r_t^*) + \phi_\pi \pi_t + \phi_y \tilde{y}_t + \varepsilon_t^m \tag{A.26}
$$

where E^c is the mathematical expectations taken with respect to the central bank's information set, economic model and beliefs about the accuracy of its information relative to that of households.

Appendix B Proofs

B.1 Proof of Proposition 1

Proof. Equation (4.18) implies $\gamma_j = 0$ and $\alpha_j = g_{sj}$.Therefore:

$$
\begin{pmatrix} g_{si} \\ g_{ai} \end{pmatrix} = \frac{1}{g_{sj}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \sigma_{uj}^2\left(\sigma_{\epsilon i}^2 + 1\right)} \begin{pmatrix} g_{sj}^2\sigma_{\epsilon j}^2 + \sigma_{uj}^2 \\ g_{sj}\sigma_{\epsilon i}^2 \end{pmatrix}.
$$

It is immediate that $g_{si} \geq 0$. By symmetry, $g_{sj} \geq 0$ and this in turn implies $g_{ai} \geq 0$ as well. It is also clear that

$$
g_{sj}^2\sigma_{\epsilon j}^2 + \sigma_{uj}^2 \leq g_{sj}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \sigma_{uj}^2\left(\sigma_{\epsilon i}^2 + 1\right)
$$

and therefore $g_{si} \leq 1$. This inequality is binding if and only if

$$
0 = \sigma_{\epsilon i}^2\left(g_{sj}^2\left(1 + \sigma_{\epsilon j}^2\right) + \sigma_{uj}^2\right).
$$

Because we assume $\sigma_{uj}^2 > 0$, this can be the case if and only if $\sigma_{\epsilon i}^2 = 0$. This establishes that $g_{si} = 1$ if and only if $\sigma_{\epsilon i}^2 = 0$.

We now show that $g_{ai} < 1$. Because $\sigma_{uj}^2 > 0$, we have that

$$
\begin{aligned}
g_{ai} &< \frac{g_{sj}\sigma_{\epsilon i}^2}{g_{sj}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right)} \\
&= \frac{1}{g_{sj}}\frac{\sigma_{\epsilon i}^2}{\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right)} \\
&= \frac{g_{si}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \sigma_{ui}^2\left(\sigma_{\epsilon j}^2 + 1\right)}{g_{si}^2\sigma_{\epsilon i}^2 + \sigma_{ui}^2}\frac{\sigma_{\epsilon i}^2}{\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right)} \\
&\leq \frac{\sigma_{\epsilon i}^2 g_{si}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right)\sigma_{ui}^2}{\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right)\left(g_{si}^2\sigma_{\epsilon i}^2 + \sigma_{ui}^2\right)} \\
&= 1.
\end{aligned}
$$

It is left to show existence of the equilibrium. Start by noting that $g_{si} \in [0,1]$ and that it is a non-increasing function of g_{sj}, in fact:

$$
\frac{\partial g_{si}}{\partial g_{sj}} = -\frac{2g_{sj}\sigma_{uj}^2\sigma_{\epsilon i}^2}{\left(g_{sj}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \sigma_{uj}^2\left(\sigma_{\epsilon i}^2 + 1\right)\right)^2} \leq 0.
$$

Because this holds for $i = c, h$, there exists at least one set of values (g_{sh}, g_{sc}) for which all equilibrium conditions are satisfied. The second derivative of g_{si} with respect to g_{sj} is:

$$
\frac{\partial^2 g_{si}}{\partial g_{sj}^2} = -2\sigma_{uj}^2\sigma_{\epsilon i}^2\frac{\sigma_{uj}^2\left(\sigma_{\epsilon i}^2 + 1\right) - 3g_{sj}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right)}{\left(g_{sj}^2\left(\sigma_{\epsilon i}^2 + \sigma_{\epsilon j}^2 + \sigma_{\epsilon i}^2\sigma_{\epsilon j}^2\right) + \sigma_{uj}^2\left(\sigma_{\epsilon i}^2 + 1\right)\right)^3}
$$

Therefore, g_{si} has at most one inflection point. It follows that there are at most three sets of values (g_{sh}, g_{sc}) satisfying the equilibrium conditions. □

B.2 Proof of Proposition 2

Proof. Under the hall-of-mirrors assumption, agent i thinks that agent j forms expectations as:

$$
E^j\left[r^{**}\right] = g_{sj}^{|i}s_j \tag{B.1}
$$

where the perceived gain is $g_{sj}^{|i} = 1/\left(1 + \sigma_{sj}^2\right)$ and $g_{aj}^{|i} = 0$.

Agent i's optimal filtering problem is then:

$$
E^i\left[r^{**}\right] = g_{si}s_i + g_{ai}\left(E^j\left[r^{**}\right] + u_j\right) \tag{B.2}
$$

where the gain parameters are given by the same expression as in (4.16), but where g_{sj} is replaced with $g_{sj}^{|i}$. In what follows, the superscript CK denotes the corresponding gains under common knowledge learning.

It is clear that $g_{sj}^{|i} > g_{sj}^{CK}$ because agent j is perceived to rely on their private signal alone. Then it also follows from (4.16) that $g_{si} < g_{si}^{CK}$: Each side pays less attention to their own private information.

For the gain g_{ai}, we can establish its derivative with respect to $g_{sj}^{|i}$ as:

$$\begin{aligned}\frac{dg_{ai}}{dg_{sj}^{|i}} &\sim \left(g_{sj}^{2|i}\left(\sigma_{si}^2+\sigma_{sj}^2+\sigma_{si}^2\sigma_{sj}^2\right)+\sigma_{uj}^2\left(1+\sigma_{si}^2\right)\right)\sigma_{si}^2 - 2g_{sj}^{2|i}\sigma_{si}^2\left(\sigma_{si}^2+\sigma_{sj}^2+\sigma_{si}^2\sigma_{sj}^2\right)\\ &\sim \sigma_{uj}^2\left(1+\sigma_{si}^2\right) - g_{sj}^{2|i}\left(\sigma_{si}^2+\sigma_{sj}^2+\sigma_{si}^2\sigma_{sj}^2\right).\end{aligned}$$

Therefore, the gain g_{ai} is increasing in $g_{sj}^{|i}$ if σ_{ui}^2 is large enough relative to $\sigma_{\epsilon i}^2$ and $\sigma_{\epsilon j}^2$. In that case, $g_{ai} > g_{ai}^{CK}$.

Now, recall that under common knowledge learning, equilibrium expectations are given by

$$E^i\left[r^{**}\right] = g_{si}^{CK} s_i + g_{ai}^{CK}\left(g_{sj}^{CK} s_j + u_j\right). \tag{B.3}$$

whereas under the hall-of-mirrors equilibrium, they are given by:

$$E^i\left[r^{**}\right] = \frac{1}{1-g_{ai}g_{aj}}\left(g_{si}s_i + g_{ai}\left(g_{sj}s_j + g_{aj}u_i + u_j\right)\right). \tag{B.4}$$

Because $g_{ai} > g_{ai}^{CK}$ holds if σ_{ui}^2 is large enough, it also holds that

$$\frac{g_{ai}}{1-g_{ai}g_{aj}} > g_{ai} > g_{ai}^{CK}$$

and the equilibrium reaction to u_j is larger under the hall-of-mirrors than under common knowledge learning. Additionally, the reaction to u_i is strictly positive under the hall-of-mirrors, whereas it is zero under common knowledge learning. □

Appendix C Solution of the dynamic model

The solution to the macro side of the model (5.7)–(5.9) can be written in the following form:

$$\begin{pmatrix}\tilde{y}_t\\ \pi_t\\ i_t\end{pmatrix} = \Gamma i_{t-1} + \Theta E_t^h Z_t + \sum_{s=0}^{\infty}\theta_s E_t^h\left[E_{t+s}^c\left[Z_{t+s}\right]\right]. \tag{C.1}$$

The coefficient matrices Γ, Θ and θ_s depend on the parameters as well as on the matrix A_z. Their values can be computed using standard solution methods. We use Chris Sims's Gensys procedure, which has the advantage that it directly yields the coefficients θ_s without the need to specify a process for $E_t^h\left[E_{t+s}^c\left[Z_{t+s}\right]\right]$. Future values $E_t^h\left[E_{t+s}^c\left[Z_{t+s}\right]\right]$ for $s \geq 0$ need to be carried over because the evolution of these expectations will be endogenous to the belief formation process.

We can rewrite the state X_{it} that player i has to learn about in a recursive form:

$$X_{it} = A_{Xi} X_{it-1} + B_{qi} q_t + B_{mi} m_{it-1}. \tag{C.2}$$

where the linear maps A_{Xi}, B_{qi} and B_{mi} depend on the guess in (5.12) as well as R in (5.6). Start by observing that $X_{it} = \left(Z_t', m_{jt}'\right)'$, so that we can define linear maps C_z,C_m, D_z and D_m for which

$$X_{it} = C_z Z_t + C_m m_{jt} \tag{C.3}$$

$$\begin{pmatrix} Z_t \\ m_{jt} \end{pmatrix} = \begin{pmatrix} D_z \\ D_m \end{pmatrix} X_{it}. \tag{C.4}$$

In practice, we cut the infinite hierarchy of higher-order beliefs at some finite level N, so that $Z_t \in \mathbb{R}^7$ and $m_{it}, X_{it} \in \mathbb{R}^{7N}$. This means that (C.4) holds only approximately: In defining the map D_m, we use the approximation $E_{it}^{(N+1)}\left[Z_t\right] \approx E_{it}^{N}\left[Z_t\right]$.

With this notation and the guess (5.12), we can write

$$\begin{aligned} X_{it} &= C_z Z_t + C_m \left(\Phi_j m_{jt-1} + \Psi_j m_{it-1} + \Omega_j Z_t\right) \\ &= \left(\left(C_z + C_m \Omega_j\right) R D_z + C_m \Phi_j D_m\right) X_{it-1} + \left(C_z + C_m \Omega_j\right) q_t + C_m \Psi_j m_{it-1} \\ &= A_{Xi} X_{it-1} + B_{qi} q_t + B_{mi} m_{it-1}. \end{aligned} \tag{C.5}$$

We now describe how to find the observation equations (5.13). The household's observation problem is straightforward. The household observes $s_{ht}, x_t, u_{ht}, u_{pt}$, as well as the nominal interest rate i_t. From the monetary policy rule (5.9), one can see that observing i_t, as well as $\tilde{y}_t$ and π_t (which are the household's own choice variables) is equivalent to observing $E_t^c z_t + u_{ct}$ each period. We can therefore write the household observation as $Y_{ht} = H_h X_{ht}$ with a matrix H_h that is independent of equilibrium beliefs:

$$H_h = \begin{pmatrix} 1 & 1 & 0 & & & & & \\ 1 & 0 & 0 & 1 & 0 & & & \\ & & & 0 & 1 & 0 & & \\ & & & & 0 & 1 & 0 & \\ & & & & & 0 & 1 & \end{pmatrix} D_z + \begin{pmatrix} 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ 0 & 0 & \cdots \\ 1 & 0 & \cdots \end{pmatrix} D_m. \tag{C.6}$$

The central bank's signalling problem is more complicated. Its information about the household's expectation comes from observing $\tilde{y}_t$ and π_t, which are themselves equilibrium outcomes that depend on the household's beliefs in a non-trivial way. We first note that i_{t-1} is also in the central bank's information set. Using (C.1), we can express the macroeconomic outcomes of the model as:

$$
\begin{aligned}
\begin{pmatrix} \tilde{y}_t \\ \pi_t \\ i_t \end{pmatrix} - \Gamma i_{t-1} &= \Theta E_t^h \left[Z_t \right] + \sum_{s=0}^{\infty} \theta_s E_t^h \left[E_{t+s}^c \left[Z_{t+s} \right] \right] \\
&= \Theta D_z m_{ht} + \sum_{s=0}^{\infty} \theta_s D_z E_t^h \left[m_{ct+s} \right] \\
&= \Theta D_z m_{ht} + \sum_{s=0}^{\infty} \theta_s D_z \left(A_{Xh} + B_{mh} \right)^s m_{ht} \\
&= \underbrace{\left(\Theta D_z + \sum_{s=0}^{\infty} \theta_s D_z \left(A_{Xh} + B_{mh} \right)^s \right)}_{=M_h} D_m X_{ct}. \qquad \text{(C.7)}
\end{aligned}
$$

Therefore $Y_{ct} = H_c X_{ct}$, where the matrix H_c is endogenous to the belief formation process:

$$
H_c = \begin{pmatrix} 1 & 0 & 1 & & & & & \\ 1 & 0 & 0 & 1 & 0 & & & \\ & & & & 0 & 0 & 1 \\ & & & & & & 0 \\ & & & & & & 0 \end{pmatrix} D_z + \begin{pmatrix} 0 & & \\ 0 & & \\ 0 & & \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} M_h D_m. \qquad \text{(C.8)}
$$

The household's expectations of future macroeconomic outcomes, which we use to compute nominal and real yields in the simulations, can be found through the recursion:

$$
E_t^h \begin{pmatrix} \tilde{y}_{t+s} \\ \pi_{t+s} \\ i_{t+s} \end{pmatrix} = \Gamma E_t^h i_{t+s-1} + M_h \left(A_{Xh} + B_{mh} \right)^s m_{ht}, \; s \geq 0. \qquad \text{(C.9)}
$$

Equations (C.2) and (5.13) form a standard linear filtering problem, the solution of which is given by the Kalman filter. The optimal filtering equation is:

$$
m_{it} = \left(I - G_i H_i \right) \left(A_{Xi} + B_{mi} \right) m_{it-1} + G_i Y_{it} \qquad \text{(C.10)}
$$

The Kalman gain G_i, as well as the time-invariant posterior covariance matrix P_i, can be computed using the following formula:

$$
\begin{aligned}
P_i^- &= A_{Xi} P_i A_{Xi}' + B_{qi} \Sigma_q B_{qi}' && \text{(C.11)} \\
S_i &= H_i P_i^- H_i' && \text{(C.12)} \\
G_i &= P_i^- H_i' S_i^{-1} && \text{(C.13)} \\
P_i &= P_i^- - G_i S_i G_{it}' && \text{(C.14)}
\end{aligned}
$$

In practice, one iterates on these four equations to find the fixed point of this system of equations.

We can now find the equilibrium with common knowledge learning and verify our conjectures (5.12) and (5.13). In a common knowledge equilibrium, agent i's beliefs of (5.12) and of the signal matrix H_i are correct. We can thus express the vector of signals Y_{it} in terms of past beliefs and the current state:

$$\begin{aligned} Y_{it} &= H_i \left(C_z Z_t + C_m m_{jt} \right) \\ &= H_i \left(\left(C_z + C_m \Omega_j \right) Z_t + C_m \Phi_j m_{jt-1} + C_m \Psi_j m_{it-1} \right). \end{aligned} \tag{C.15}$$

Substituting this expression into (C.10) gives an expression for m_{it} that verifies our guess (5.12). The equilibrium coefficients can be found using the following system of equations:

$$\Phi_i = \left(I - G_i H_i \right) A_{Xi} + B_{mj} \tag{C.16}$$
$$\Psi_i = G_i H_i C_m \Phi_j \tag{C.17}$$
$$\Omega_i = G_i H_i \left(C_z + C_m \Omega_j \right). \tag{C.18}$$

To compute the equilibrium numerically, we use the following iterative algorithm:

1. Start with an initial guess $\left(\Phi_i, \Psi_i, \Omega_i, H_i \right)_{i=c,h}$.

2. For $i = c, h$:

   (a) compute the law of motion for X_{it} from (C.5);

   (b) compute the Kalman matrices P_i^-, S_i, G_i, P_i from (C.11)–(C.14);

   (c) compute Φ_i, Ψ_i, Ω_i from (C.16)–(C.18).

3. Compute the signal matrix H_c according to (C.21). The matrix H_h stays the same across iterations.

4. Iterate on steps 2. and 3. until convergence.

We now turn to the hall-of-mirrors equilibrium. Here, the central bank and the private sector disagree on the flow of information in the economy. To solve agent i's perceived law of motion of beliefs, we use the same procedure as above for the common knowledge learning equilibrium, but with modified signal matrices. Specifically, for $i = h$, the private sector assumes the central bank ignores the information contained in inflation and the output gap. The signals observed by the private sector take the same form $Y_{ht} = H_{ht} X_{ht}$ as in (C.6). but the signals observed by the central bank (as perceived by the private sector) are now $Y_c t = H_c^{|h} X_{ct}$ with

$$H_c^{|h} = \begin{pmatrix} 1 & 0 & 1 & & & & \\ 1 & 0 & 0 & 1 & 0 & & \\ & & & & 0 & 0 & 1 \end{pmatrix} D_z. \tag{C.19}$$

For $i = c$, the central bank assumes that the privte sector ignores the information contained in the interest rate, so that $Y_h t = H_h^{|c} X_{ht}$ with

$$H_h^{|c} = \begin{pmatrix} 1 & 1 & 0 & & & & & \\ 1 & 0 & 0 & 1 & 0 & & & \\ & & & 0 & 1 & 0 & & \\ & & & & 0 & 1 & 0 & \end{pmatrix} D_z. \tag{C.20}$$

The central bank itself has beliefs about how macroeconomic outcomes depend on shocks through private sector expectations. These beliefs are given by:

$$H_c^{|c} = \begin{pmatrix} 1 & 0 & 1 & & & & & \\ 1 & 0 & 0 & 1 & 0 & & & \\ & & & & 0 & 0 & 1 & \\ & & & & & & \Theta_{1\cdot} & \\ & & & & & & \Theta_{2\cdot} & \end{pmatrix} D_z + \begin{pmatrix} 0 & & \\ 0 & & \\ 0 & & \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} M_h^{|c} D_m. \tag{C.21}$$

Here, the appearance of $\Theta_{1\cdot}$ and $\Theta_{2\cdot}$ reflects an assumption that the central bank expects the reaction to a monetary policy shock u_{ct} to be the same as under full information (the belief about u_{ct} embedded in m_{ht} is always zero if the private sector does not use i_t as a signal). The term $M_h^{|c}$ is the counterpart of M_h in (C.7) evaluated under the perceived law of motion of the central bank.

To proceed from the perceived law of motions to the equilibrium, we then write the filtering equation (C.10) as:

$$\begin{aligned} m_{it} &= \left(\Phi_i^{|i} - G_i^{|i} H_i^{|i} C_m \Psi_j^{|i}\right) m_{it-1} + G_i^{|i} Y_{it} \\ &= \left(\Phi_i^{|i} - G_i^{|i} H_i^{|i} C_m \Psi_j^{|i}\right) m_{it-1} + G_i^{|i} H_i^{|j} \left(C_z Z_t + C_m m_{jt}\right). \end{aligned} \tag{C.22}$$

To obtain the second line, we substitute out the actual signals Y_{it} that agent i receives, which depend on the expectations of agent j and are hence given by $Y_{it} = H_i^{|j} X_{it}$. The above expression also holds when the roles of i and j are reversed, and we can use this fact to substitute out m_{jt}. We then obtain the actual law of motion of beliefs describing the hall-of-mirrors equilibrium:

$$m_{it} = \Phi_i m_{it-1} + \Psi_i m_{jt-1} + \Omega_i Z_t \tag{C.23}$$

where the actual transition coefficients are given by:

$$\Phi_i = \left(I - G_i^{|i} H_i^{|j} C_m G_j^{|j} H_j^{|i} C_m\right)^{-1} \left(\Phi_i^{|i} - G_i^{|i} H_i^{|i} C_m \Psi_j^{|i}\right) \tag{C.24}$$

$$\Psi_i = \left(I - G_i^{|i} H_i^{|j} C_m G_j^{|j} H_j^{|i} C_m\right)^{-1} G_i^{|i} H_i^{|j} C_m \left(\Phi_j^{|j} - G_j^{|j} H_j^{|j} C_m \Psi_i^{|j}\right) \tag{C.25}$$

$$\Omega_i = \left(I - G_i^{|i} H_i^{|j} C_m G_j^{|j} H_j^{|i} C_m\right)^{-1} G_i^{|i} H_i^{|j} \left(I + C_m G_j^{|j} H_j^{|i}\right) C_z. \tag{C.26}$$

Macroeconomic outcomes in this equilibrium are again given by (C.9), but where A_h and Ψ_c are replaced by $A_h^{|h}$ and $\Psi_c^{|h}$, respectively, in (C.9) and in the formula for M_h above.

Appendix D Model counterpart to high-frequency policy surprises

In the empirical literature, monetary policy surprises and their effects on asset prices are constructed by measuring changes in short- term interest rates and other asset prices in narrow intra-day windows around announcement dates. Identification obtains from the making sure that during the window, no macroeconomic news other than the policy announcement are realized. In our model, there is no direct counterpart to an announcement window, as all macroeconomic shocks are realized simultaneously, at the same time as expectations are updated.

To construct a counterpart to policy surprises in the model, we proceed as follows. In each period, we construct the household's belief about the state X_{ht} after observing the public signal x_t, the private signal s_{ht}, as well as the demand and cost-push shocks u_{ht}, u_{pt}, but before observing the current interest rate i_t. Denote this belief by $\bar{m}_{ht}$. When the prior is distributed as $\mathcal{N}\left(m_{ht-1}, P_h\right)$, then the posterior $\bar{m}_{ht}$ can be found using the analogous equations to (C.10) of the Kalman filter:

$$\bar{m}_{ht} = \left(I - \bar{G}_h \bar{H}_h\right)\left(A_h + C_m \Psi_c\right) m_{ht-1} + \bar{G}_h \bar{Y}_{ht} \tag{D.1}$$

$$\bar{G}_h = P_h^{-} \bar{H}_h' \left(\bar{H}_h P_h^{-} \bar{H}_h'\right)^{-1}. \tag{D.2}$$

Here, the signals before the observation of the interest rate are:

$$\bar{Y}_{ht} = \begin{pmatrix} s_{ht} \\ x_t \\ u_{ht} \\ u_{pt} \end{pmatrix} = \bar{H}_h X_{ht} = \left(\begin{array}{ccccccc} 1 & 1 & 0 & & & & \\ 1 & 0 & 0 & 1 & 0 & & \\ & & & 0 & 1 & 0 & \\ & & & & 0 & 1 & 0 \\ & & & & & 0 & 1 \end{array}\right) D_z X_{ht}. \tag{D.3}$$

We now take the information that is revealed through observing the current interest rate as the policy surprise, which comprises both information about the policy shock u_{ct} as well as about the central bank's beliefs $\hat{r}_t^*$. The effect of the surprise on private sector expectations is simply $m_{ht} - \bar{m}_{ht}$. The surprise in $E_t^h\left(\tilde{y}_{t+s}, \pi_{t+s}, i_{t+s}\right)$, which can be used to compute the short-term policy surprise in i_t as well as the announcement effects on the nominal and real yield curve, then takes the form:

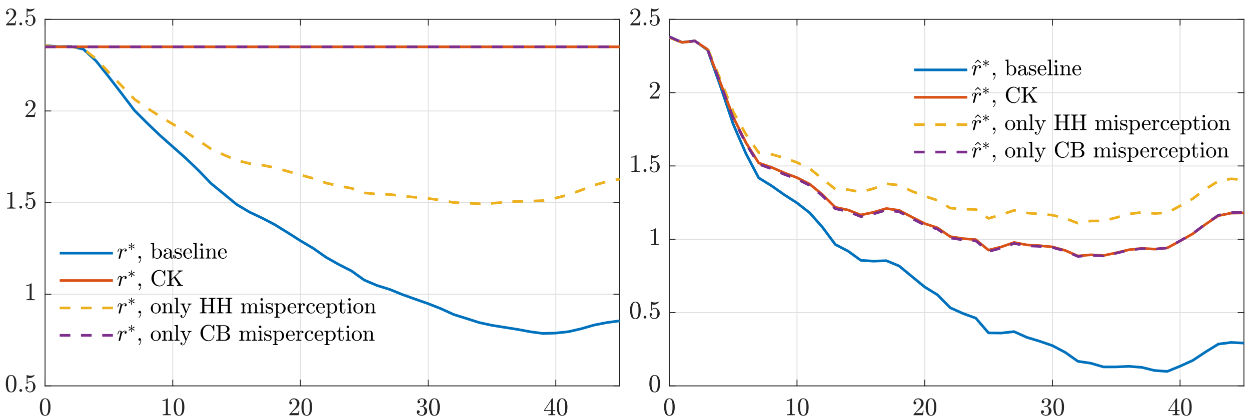
$$M_h \left(A_h + C_m \Psi_c\right)^s \left(m_{ht} - \bar{m}_{ht}\right),\ s \geq 0. \tag{D.4}$$

Appendix E Sensitivity to alternative information structures

In this section, we examine the sensitivity of the strength of the hall-of-mirrors effect to alternative assumptions about the information structure. First, we show that the hall-of-mirrors effect is greatly reduced if only one agent is ignorant of the two-sided learning problem. Second, we show that the hall-of-mirrors effect remains strong across a wide range of reasonable information parameters, pointing to robustness of the mechanism.

We first repeat the simulation of Figure 4 using the same sequence of shocks, but assuming only one or none of the two learning agents suffers from the misperception that the other agent does not learn from macroeconomic outcomes. The corresponding simulated outcomes on the private sector's and the central bank's estimates of r-star (r^* and $\hat{r}^*$ respectively) are shown in Figure E.1.

Figure E.1: Disappearance of the hall-of-mirrors effect with one-sided misperception.



Note: The simulation labeled "baseline" refers to the simulation shown in Figure 4. The other simulations use the same shocks, but assume that only one or neither agent ignores the double learning problem.

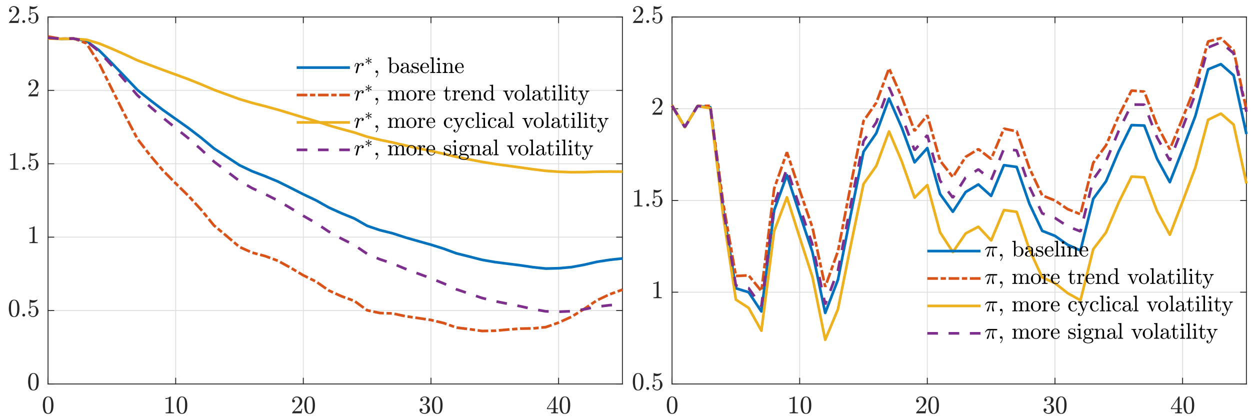
The simulations labeled "baseline" and "CK" replicates the hall-of-mirrors and common knowledge learning simulations shown in Figure 4. The simulation labeled "only HH misperception" assumes that only the private sector (i.e. the household) is unaware that the central bank is learning from macroeconomic outcomes. The central bank is fully aware of the private sector's learning problem, including its misperception, and can correct for belief distortions affecting its observation of the output gap and inflation. Figure E.1 reveals that the hall-of-mirrors effect is greatly reduced in this case. Private sector misperception of r-star is only about half as strong, while the central bank's beliefs actually move a little less than under common knowledge learning. When the central bank is aware of the private sector's misperception, it effectively reduces the strength of the informational feedback loop.

In the simulation labeled "only CB misperception", it is the central bank that is unaware of the fact that the private sector is learning from observing interest rates. The private sector can correct its beliefs for the central bank's misperception, and also perfectly observes and understands the demand shocks hitting the economy. As a result, its r-star beliefs are unchanged. The central bank still misinterprets the adverse demand shocks as a fall in r-star, but by just as much as under common knowledge learning.

Next, we assess the sensitivity of the hall-of-mirrors effect to signal precision. We repeat the simulation of Figure 4 keeping the same shocks, but considering several alternative information parameters. The corresponding simulated outcomes on the private sector's de facto r-star and the central bank's estimate of r-star are shown in Figure E.2.

The first simulation labeled "baseline" is identical to that shown in Figure 4. In the second simulation labeled "more volatile fundamentals", the underlying r-star determinants

Figure E.2: Sensitivity to information parameters.



Note: The simulation labeled "baseline" refers to the simulation shown in Figure 4. In each of the other simulations, one parameter is changed relative to this baseline. For "more trend volatility", the volatility σ_r of r^{**} is set to 0.075 instead of 0.05. For "more cyclical volatility", the standard deviation of the three cyclical shocks u_{ht}, u_{pt}, u_{ct} by a factor of 1.5. For "more signal volatility", the standard deviation of the noise in signals e_{ht}, e_{ct} is increased by a factor of 1.5.

(i.e. the trends of productivity and discount factor changes) become more volatile. Our baseline calibration has $\sigma_r = 0.05$, corresponding to quarterly changes in annualized r-star of 0.2 percent, which is at the lower end of the estimates in Holston et al. (2017). Increasing this value to $\sigma_r = 0.075$, towards the upper bound of their estimates, leads to a further endogenous reduction of half a percentage point in r_t^*. Intuitively, when it is harder to pin down the true drivers of r-star, agents rely more on learning from each other, strengthening the hall-of-mirrors effect.

In the third simulation labeled "more cyclical volatility", the standard deviation of each of the cyclical shocks u_{ct}, u_{pt}, u_{ht} is raised by 50 percent. The result is that the amount of misperception (for the same realizations of the shocks) is cut in half. The reason is that the cyclical shocks do not only cause business cycle fluctuations, but also act as noise in the endogenous signals about the beliefs of the central bank and the private sector. More noise reduces the incentives to learn from each other and therefore weakens the hall-of-mirrors effect.

In the final simulation labeled "more signal volatility", we increase the standard deviations of the noise terms in each of the private signals s_{ht}, s_{ct}, x_t by 50 percent. This change increases the amount of misperception, but also leads both the private sector and the central bank to pay less attention to their own information. The result is a decline in r-star of similar magnitude but somewhat higher persistence than in the baseline.

Appendix F Full-sample simulation

Here, we present a simulation that combines the post-GFC exercise and the post-pandemic exercise in the paper. We initialize r-star fundamentals at $r_0^{**} = 2.35\%$ and choose a sequence of cost-push and demand shocks u_{pt}, u_{ht} to exactly match inflation and the output gap in the calibrated hall-of-mirrors model to core PCE inflation and the CBO output gap in the data 2007Q4-2023Q2, spanning the GFC and the pandemic. All other shocks are set to zero.

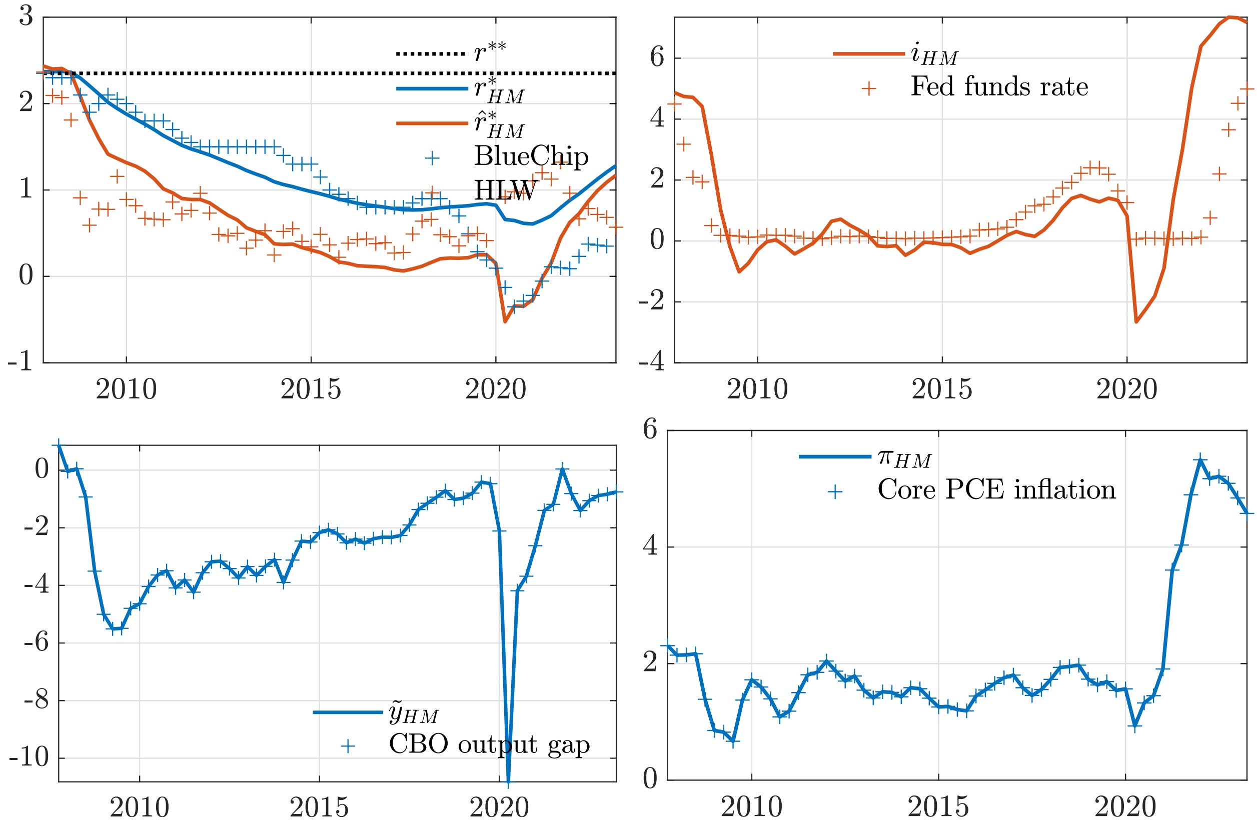
Figure F.3 shows the simulated series of r-star beliefs, interest rates, the output gap and inflation. By construction of the exercise, the simulated model outcomes are identical to those shown in the main text in Section 6.3 through 2019. Inflation falls below 1 percent and then stays persistently below 2 percent through 2019, the output gap plunges then gradually recovers, interest rates remain near zero until 2016, and r-star beliefs gradually decline to about 0.5 percent, similar to the data.

When the pandemic causes an unprecedentedly sudden decline in economic activity, coupled with a modest fall in inflation, the de-facto r-star of the private sector declines only about 0.2 percentage point while the r-star estimate of the central bank falls by about a percentage point. As a result, the simulated policy rate drops below minus 2 percent. But When inflation surges again in 2021, the policy rate quickly quickly rises and reaches a peakf of more than 6 percent in 2022. The r-star beliefs of the private sector and the central bank start rising again and, in our simulation, are close together at just above one percent at the end of the simulation.

Comparing these outcomes to the data, we see that the model does not capture well the movements in r-star beliefs in 2019 and 2020. BlueChip measures of r-star declined substantially in 2019 without material reductions in output and inflation. By contrast, the measure of Holston et al. (2017) jumped up during the pandemic, rose further and then subsequently declined. Methodological changes made to this model after the pandemic (Holston et al., 2023) make it hard to compare this measure against our model, which predicts a notable decline in the central bank's r-star belief and a swift recovery afterwards. What the model does get right is the sustained increase in private sector r-star expectations in the last two years of the simulation. Also, the simulation does not impose an ELB constraint, and as a result the policy rate in 2020 drops well below zero. In the following year, the policy rate implied by the model rises rapidly, while the actual policy rate was held at the ELB for another year.

Despite these shortcomings, the model reproduces salient fact about the behavior of the yield curve in the aftermath of the pandemic. Figure F.4 compares realized and simulated yield curves side-by-side. The model captures the slow decline in long-term interest rates before the pandemic as well as their rapid rise afterwards. It also captures the flattening of the yield curve in the years after the GFC, as well as its steepening and subsequent inversion following the pandemic.

Figure F.3: Macroeconomic outcomes in a long simulation.

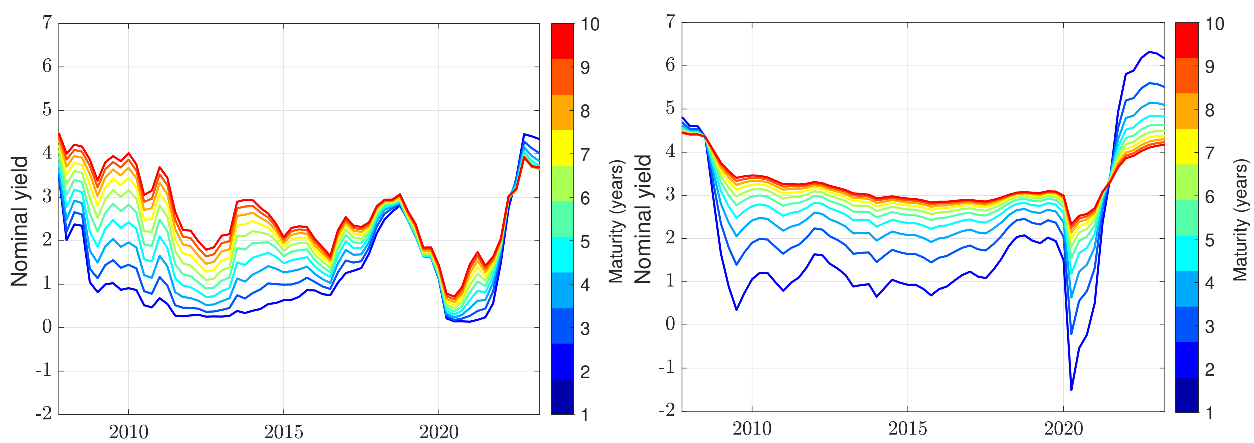


Note: The simulations are based on a sequence of demand and cost-push shock that are constructed so that the paths for the output gap and inflation in the hall-of-mirrors equilibrium replicate the data series shown. Other shocks are set to zero. Parameters used for the calibration are shown in Table 1.

Figure F.4: Yield curves in a long simulation.

(a) Yield curve data

(b) Yield curve simulation



Note: Simulation of yield curves and forward interest rates in the hall-of-mirrors equilibrium, constructed using expected paths of interest rate and assuming that the expectation hypothesis holds. For the construction of monetary policy surprises and associated yield movements see Appendix D. The simulations are based on the same sequence of shocks as those underlying Figure F.3.