Journal of Economic Dynamics & Control 138 (2022) 104372

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Multi-layered rational inattention and time-varying volatility☆

Stephan Hobler

London School of Economics, United Kingdom

ARTICLE INFO

Article history:
Received 28 January 2021
Revised 17 March 2022
Accepted 21 March 2022
Available online 23 March 2022

JEL classification:
D8
E31
E32
E42
E52
E7

Keywords:
Rational inattention
Phillips curve
Monetary policy

ABSTRACT

Standard rational inattention models suppose that agents process noisy signals about otherwise fully revealing data. I show that introducing imperfect data quality yields new insights in settings in which volatility is time-varying. I impose a two-layered signal structure in which agents learn imperfectly about noisy sources. Treating data as only partially revealing of the true fundamental amplifies impulse responses to a second moment shock and, if data quality is sufficiently poor, can change the qualitative direction of the response. I apply my findings to the price-setting problem of firms and find that higher data quality enhances the transmission of monetary policy and reduces macroeconomic volatility. I also show how the empirically documented procyclicality of data quality has non-trivial implications for the Phillips curve.

'Steering the economy has sometimes been likened to driving along a winding road looking only in the rear-view mirror. I wish it were that easy. In practice, the rear window is also a little misted up ... Not only do we not know where we are going, but we have only an imperfect idea of where we have been.'- Charles Bean (2007)

1. Introduction

Data uncertainty has long been recognised as an inherent feature of real-time decision making. Official statistics are revised continuously and often subject to substantial measurement errors (see Aruoba, 2008, Croushore, 2011, Croushore and Stark, 2001, Manski, 2015 and Orphanides, 2001). What is more, data inaccuracies are found to be more pronounced during recessions and periods of heightened uncertainty – precisely at times when vital decisions must be made (Jordá et al., 2020; Swanson and van Dijk, 2006). Croushore (2011) illustrates this point with a notable example: The largest revision from advance to first release data ever recorded for quarterly real GDP occurred in the middle of the financial crisis of 2008/09. Likewise, at the deepest point of the Covid recession US real GDP growth was subject to an above average absolute revision.

Given the empirical evidence that this literature has uncovered the Bank of England started to provide an estimate of data uncertainty in their quarterly Monetary Policy Report (formerly Inflation Report) in August 2007. The Monetary Policy Committee has since published output growth projections for both the past and the future, using fan charts to indicate the uncertainty around those values. Fig. 1 illustrates one of these fan charts from the January 2020 *Monetary Policy Report* and

☆ I am grateful to the editor and two anonymous referees for their extensive feedback and excellent comments. I thank Martin Ellison for his support throughout this project and for his many valuable comments and suggestions.

E-mail address: s.j.hobler@lse.ac.uk

https://doi.org/10.1016/j.jedc.2022.104372
0165-1889/

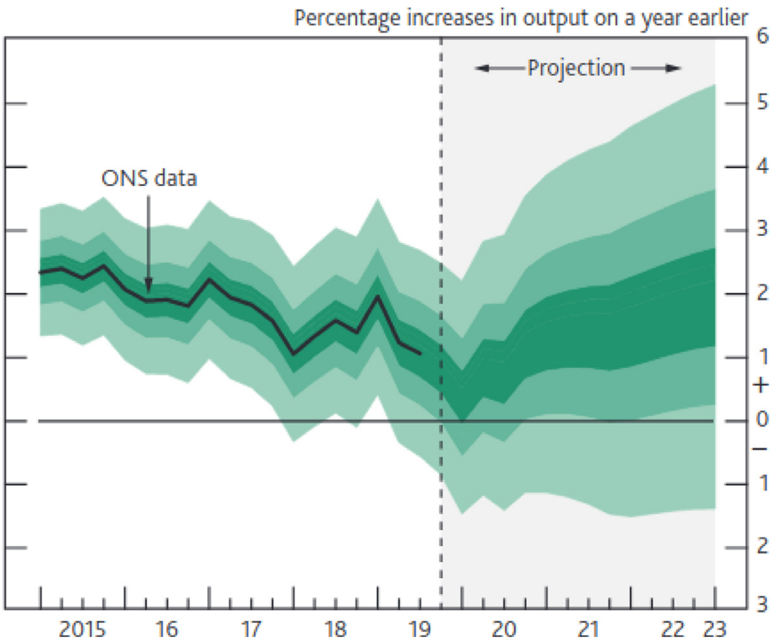

Fig. 1. January 2020 UK GDP Fan Chart. *Source*: Bank of England.

indicates considerable uncertainty around the official ONS data. The Bank of England is not only concerned about the noise contained in past observations but believes imperfect data to be sufficiently vital to economic agents as to communicate it to them.

The Covid crisis has further heightened the attention of policymakers to the need for frequent and reliable real-time statistics as well as highlighted the noisiness of real-time data. As a case in point, the UK's Office for National Statistics (ONS) started in March 2020 to attach a disclaimer to their initial data releases citing governments imposed public health restrictions as well as current economic conditions due to the pandemic as generating additional uncertainty around real-time GDP estimates.

Using the Real-Time Dataset for Macroeconomics created by Croushore and Stark (2001) I confirm the above cited empirical regularities. In particular, I document three stylised facts for revisions to quarterly US inflation, nominal GDP growth and real GDP growth. First, corrections to real-time data can be substantial with large mean absolute revisions (MAR) and a signal to noise ratio mostly ranging between 1 and 3. As an implication we may expect the measurement error in real-time data to play a non-negligible role in the decision-making process of agents. Second, we observe a long-run decline in the size and volatility of revisions over time. This reflects the advances in statistical methods, survey quality and the way data is collected. Third, the data highlights the time-varying nature of revision errors. As a case in point, the noise in initial data releases for the above three macroeconomic variables was twice as large and volatile during the Great Recession relative to the period of the Great Moderation.

The above empirical and anecdotal evidence underscores the practical importance of data revisions for real-time decision making. In the present paper, I explicitly model these considerations and place additional constraints on the set of available information. Agents are restricted to learn about the fundamental from noisy sources such as reading real-time statistics or central bank reports akin to the Monetary Policy Report by the Bank of England. In effect, agents choose their private signal of a noisy public signal. This *'signal of a signal'* structure captures the notion by Myatt and Wallace (2012) that information is subject to both "sender" and "receiver" noise. That essentially introduces a distinction between the set of learnable information available to agents and the probability space spanned by the fundamental. But what do we lose, if anything, by treating data sources as fully revealing of the fundamental? This question lies at the core of the present paper.

The most immediate interpretation sees the irreducible noise in public signals as measurement error of real-time data as in the introductory example. Complementarily, Haldane et al. (2020) use an analogous set-up in which the noise term is a stand-in for imperfect or incomplete central bank communication. This interpretation can be rationalised as follows. As part of a more general central bank communication strategy the Bank of England regularly transmits their views on current and future economic conditions to the public, mainly via the quarterly Monetary Policy Report. Notably, these reports are orthogonal to news about future policy decisions. Instead, the Bank reviews the economic developments it considers most pressing given current conditions, publishes summary tables of real-time data and gives projections on a variety of economic indicators. This constitutes an effort to inform the public about current economic conditions. To the extent that the central bank does not include all or simplifies some data or gives imperfect projections the Monetary Policy Report can be viewed as a noisy public signal about the true state of the economy - not very dissimilar to real-time data.

Though non-standard, the *'signal of a signal'* structure of the type discussed above is not entirely novel in the rational inattention literature (see Myatt and Wallace, 2012 and Fuster et al., 2018).[1] Whereas previously the focus lied on the static choice between information sources with varying levels of accuracy I examine how optimal attention allocations change dynamically in response to volatility shocks, once the notion of noisy data is taken seriously in the model. The dynamic perspective is paramount as, if volatilities were time-invariant, we could always find an isomorphism between the standard rational inattention model and its counterpart with noisy data by adjusting the generally free marginal cost parameter. In a world of time-varying volatility, however, this tight correspondence no longer holds.

I consider a variant of the standard Linear Quadratic Gaussian rational inattention model in which I impose the above *'signal of a signal'* structure. Agents are restricted to processing information about the fundamental only indirectly through noisy public signals while their payoff function remains the squared difference between the optimal and their realised action. I examine two distinct types of volatility shocks: i) a change in the variance of the fundamental and ii) a change in the variability of the irreducible noise contained in the public signal. I call the former a *fundamental* shock and the latter an *information quality* or *irreducible noise* shock.

In a macroeconomic application of the model I place the two-layered *'signal of a signal'* framework into the economic environment of Afrouzi and Yang (2021) to propose an attention-driven Phillips curve that links fundamental uncertainty and data uncertainty to the structure of the economy. I find a set of three results.

Comparative statics show that data quality can have substantive effects on the transmission of monetary policy. More accurate public signals increase attention levels, enhance the transmission of monetary policy and reduce the persistence of inflation and output deviations. Even as agents are more attentive I find that the unconditional variance of inflation and output is lower the higher data quality, highlighting how better data quality can aid the central bank fulfil their mandate.

When variances are time-varying the presence of noisy data carries additional non-trivial implications. I identify a region in the parameter space for the marginal cost of information processing for which in the short run a more dovish monetary policy regime increases attention in a model with irreducible noise and lowers information processing in a model without such information quality concerns. I show that when initial information quality is poor this range is substantial owing to large interaction effects between the underlying variances. This insight can help contextualise some implications of Afrouzi and Yang (2021). Using a dynamic LQG rational inattention model with perfect data the authors find that a shift in monetary policy which increases the variance of the fundamental may explain a temporary flattening of the Phillips curve. I show that this mechanism is weaker in an environment in which observable information is a noisy signal of reality.

I propose a mechanism which may have contributed to a temporary flattening of the Phillips curve during the Great Recession. A fall in information quality induced by the Great Recession could have lowered attention levels in the economy temporarily and thereby disrupted the relationship between inflation and output as well as rendered macroeconomic variables more persistent, ceteris paribus. Even if fundamental volatility increases concurrently as has been observed empirically I show that the class of rational inattention models considered has a tendency for attention to fall in response to a joint increase in the underlying variances. However, this result is sensitive to whether the signal to noise ratio increases and to the initial level of the signal to noise ratio. Nonetheless, incorporating data uncertainty into a rational inattention model may have important quantitative effects. The idea that fluctuations in information quality have real effects follows in spirit Van Nieuwerburgh and Veldkamp (2006) and Fajgelbaum et al. (2017), who find respectively that changes in data quality can account for the asymmetry observed in GDP data and the sluggish recovery after the Great Recession.

The rest of the paper proceeds as follows. Section 2 briefly reviews the traditional rational inattention model and then characterises the extended variant with irreducible noise. Section 3 applies the mechanism to the dynamic pricing decision of producers and derives an attention-driven Phillips curve. Section 4 considers the impulse responses of attention to changes in underlying volatilities in the dynamic setting and Section 5 concludes. The Appendix contains omitted proofs, model derivations and additional results and figures.

2. The model

To fix ideas I first review the traditional rational inattention problem of Sims (2003),(2010). I then introduce imperfect data quality into the set-up to capture its role in the agent's attention problem.

2.1. The traditional RI problem

Let $\theta \in \Theta$ denote the underlying fundamental relevant for an agent's decision making, $y \in \mathcal{Y}$ the action taken by the agent and $\tilde{\omega}$ the marginal cost of processing information. Agents have access to an imperfect signal $s \in \mathcal{S}$. Rationally inattentive agents flexibly choose the distribution of their signal subject to an entropy-based cost. This captures the idea that the types of mistakes that agents make are not arbitrary but subject to choice (Maćkowiak et al., 2018b).

It is common to place the two following restrictions on the model: preferences are linear-quadratic, i.e. $\mathbb{E}_{\theta|s}u(\theta, y) = -\mathbb{E}_{\theta|s}\frac{1}{2}(\theta - y)^2$ and the fundamental follows a Gaussian distribution, i.e. $\theta \sim \mathcal{N}(0, \sigma_\theta^2)$. It is a standard result from information theory that under those assumptions the agent's optimal signal structure will be Gaussian with independent noise

[1] Bloedel and Segal (2021) use an analogous multi-layered signal structure to answer the question of how a principal should communicate with a rationally inattentive agent.

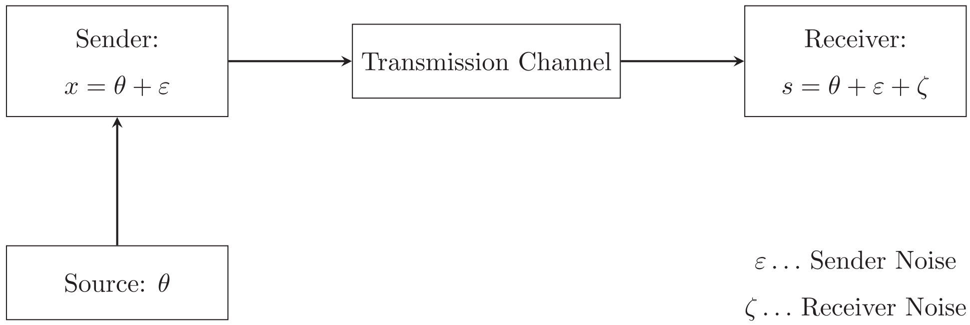

Fig. 2. Two Stage Information Processing About Public Signal.

$\zeta \sim \mathcal{N}(0, \sigma_\zeta^2)$. In such a case, the problem simplifies to choosing the posterior variance conditional on signal s to solve

$$\max_{\mathbb{V}ar(\theta|s)} -\mathbb{V}ar(\theta|s) - \tilde{\omega} \log\left(\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(\theta|s)}\right) \tag{1}$$

subject to

$$0 \le \mathbb{V}ar(\theta|s) \le \mathbb{V}ar(\theta)$$

I refer to (1) as the *traditional RI problem*.

2.2. RI With imperfect data quality

In the traditional RI problem the agent could in principle observe the fundamental perfectly if either their capacity were infinite or information processing costless. In the modified model I follow Myatt and Wallace (2012) and Fuster et al. (2018) and change the signal structure to prevent agents from learning the fundamental perfectly, even under those extreme cases. That effectively places a constraint on the set of available information. Agents have at best access to a public signal with common Gaussian irreducible noise

$$x = \theta + \varepsilon, \quad \varepsilon \sim \mathcal{N}(0, \sigma_\varepsilon^2)$$

Agents now optimally choose a mapping from public signals (and not the fundamental directly) to distributions over private signals $s : \chi \to \Delta(\mathcal{S})$.[2] It is this information flow – from public signal x to private signal s – that is subject to a cost proportional to the Shannon mutual information function denoted as $I(x; s)$. This follows the interpretation in Myatt and Wallace (2012) that information is subject to both ”sender” and ”receiver” noise. While the cost of information transmission is relative to the public signal, the objective function remains the mean squared error between actions and the fundamental. In effect, there is a disconnect between the set of information available to the agent and what they fundamentally care about. This modelling assumption has the additional implication of mechanically increasing the effective cost of learning by constraining the set of available information. This is an issue I will return to at the end of the section and throughout the paper.

Following the above reasoning I model the information transmission process of agents as two separate activities (see Fig. 2). First, information is acquired and collected in a public signal. For example, a statistics agency compiles information about real-time aggregate variables such as inflation or nominal output. The noise σ_ε^2 contained in that best possible signal is a reduced form way to capture that data collection is imperfect even before the information is transmitted to agents. Once the data has been compiled into a public signal it is published, communicated to the agents and subsequently incorporated into their decision making process. A natural candidate to measure that flow of information and their associated costs is the Shannon mutual information.

Given the Linear Quadratic Gaussian nature of the problem agents formally solve

$$\max_{f_{(\mathcal{X},\mathcal{S})}(x,s)} - \mathbb{V}ar(\theta|s) - 2\omega I(x; s)$$

subject to

$$\int_{s\in\mathcal{S}} f_{(\mathcal{X},\mathcal{S})}(x, s) ds = f_{\mathcal{X}}(x) \quad \forall x \in \mathcal{X}$$

[2] Bloedel and Segal (2021) refer to this as *costly communication* as opposed to *delegated information acquisition* as in Lipnowski et al. (2020) since the agent cannot learn about the state itself but only about the public signal.

$$I(x;s) = \int_{x\in\mathcal{X}}\int_{s\in\mathcal{S}} f_{(\mathcal{X},\mathcal{S})}(x,s)\log\left(\frac{f_{(\mathcal{X},\mathcal{S})}(x,s)}{f_{\mathcal{X}}(x)f_{\mathcal{S}}(s)}\right)dxds$$

The following result allows us to greatly simplify the problem.

Proposition 1. *Suppose that preferences are quadratic in θ and $x = \theta + \varepsilon$, where θ and ε are independently Gaussian distributed. Given that agents can only learn the public signal the joint distribution of the public and the private signal $f_{(\mathcal{X},\mathcal{S})}(x,s)$ will be Gaussian.*

Proposition 1 states that we can restrict our focus to signals of the type $s = x + \zeta$, $\zeta \sim \mathcal{N}(0, \sigma_\zeta^2)$. Note that the Gaussianity of the posterior signal follows from optimal agent behaviour. This allows us to rewrite the objective function in a similar fashion as in (1). By an application of the Total Law of Variance agents choose their optimal posterior variance about the public signal to maximise

$$\max_{\mathbb{V}ar(x|s)} -\mathbb{V}ar(\theta|x) - \left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^2\mathbb{V}ar(x|s) - \omega\log\left(\frac{\mathbb{V}ar(x)}{\mathbb{V}ar(x|s)}\right) \tag{2}$$

subject to

$$0 \le \mathbb{V}ar(x|s) \le \mathbb{V}ar(x)$$

Isomorphism result Restricting the learning technology of agents to processing information about the fundamental indirectly through noisy public signals has rendered information processing *effectively* more costly. In fact, learning about the fundamental from a public signal is isomorphic to learning about the public signal in the standard fashion at a higher information processing cost. To see this point, note that (2) has the same solution as

$$\max_{\mathbb{V}ar(x|s)} -\mathbb{V}ar(x|s) - \hat{c}\log\left(\frac{\mathbb{V}ar(x)}{\mathbb{V}ar(x|s)}\right) \tag{3}$$

where the total effective cost is denoted as $\hat{c} := \omega\left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^{-2}$.

The above equivalence relation illustrates that the traditional RI problem and the RI problem with irreducible noise are closely related. The latter introduces an additional term capturing imperfect information quality that amplifies the standard cost parameter ω.[3] This highlights that when the set of available information is constrained to public signals an agent effectively faces two distinct types of costs. They incur the standard disutility from processing a bit of information as well as additional losses from the fact that the rate at which a reduction in uncertainty about the public signal translates into a reduction in uncertainty about the fundamental is below one. In that sense rational inattention with imperfect data quality can be viewed as a variant of traditional rational inattention that endogenises the cost of information processing. The endogenous nature also implies that this isomorphism breaks down in a world of time-varying volatility.

2.3. Attention with time-Varying volatility

Empirical evidence indicates that the distribution of economic shocks is time-varying. Bloom (2009) and Bloom et al. (2018) show that second moment shocks are integral features of business cycles and key drivers of macroeconomic fluctuations. Likewise, Jordá et al. (2020) and Swanson and van Dijk (2006) find that the volatility of revision errors fluctuates with economic conditions.[4] Motivated by these findings I examine how attention levels respond to two distinct types of uncertainty shocks:

(i) to the variance of the underlying fundamental σ_θ^2
(ii) to the variance of the non-fundamental, irreducible noise σ_ε^2

The model at hand allows for the traditional discussion on the effects of shocks to the fundamental volatility but also on how a change in data quality impacts attention choices of agents. Note that the model nests the classic rational inattention model with time-varying volatility by letting $\sigma_\varepsilon^2 \to 0$. This is a useful benchmark case to keep in mind throughout this section and all results presented below carry through to that limiting case.

Definition 1. The attention level in an economy is defined as the weight κ placed on the signal in the linear projection $\mathbb{E}[\theta|s] = \mathbb{E}[\theta] + \kappa(s - \mathbb{E}[s])$. In a dynamic setting the attention level κ_t similarly refers to the Kalman gain of the linear projection $\mathbb{E}[\theta_t|s^t] = \mathbb{E}[\theta_t|s^{t-1}] + \kappa_t(s_t - \mathbb{E}[s_t|s^{t-1}])$.

[3] That learning about a public signal mechanically increases costs can also be seen from the following decomposition of the mutual information $\mathbb{I}(x;s) = \mathbb{I}(q;s) + \mathbb{I}(\varepsilon;s|q)$. The first term is the cost incurred in the standard model while the latter is the additional cost incurred by the constraint on available information.

[4] The empirical literature and data on the cyclicality of data quality is discussed in greater depth in Section 3.1.

By the above definition attention denotes the rate with which agents process incoming signals and respond to contemporaneous shocks. Solving for the optimal attention level of an agent in the static RI model with irreducible noise yields

$$\kappa = \max\left\{\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\varepsilon^2} - \omega\frac{1}{\sigma_\theta^2}, 0\right\} \tag{4}$$

where the solution is interior for sufficiently small effective processing costs $\hat{c} < \sigma_\theta^2 + \sigma_\varepsilon^2$.

Proposition 2. *Attention levels are (i) increasing in the variance of the fundamental, σ_θ^2, but (ii) decreasing in the variance of the irreducible noise, σ_ε^2.*

The first result of Proposition 2 is well understood in the literature. The greater the fundamental's variance the greater potential losses and thus the higher attention levels to guard against them (see for example Maćkowiak and Wiederholt (2009) and Paciello and Wiederholt (2014)). The time and state-dependency of attention also finds support in recent empirical studies by Coibion and Gorodnichenko (2015) and Coibion et al. (2018).

The second result has drawn less attention in the literature. Notable exceptions are Myatt and Wallace (2012) and Fuster et al. (2018), who show that amongst a set of public signals a rationally inattentive agent will elect to attend only to the most accurate information source, i.e. that with the smallest variance of the irreducible noise. These papers primarily focus on the selection of information sources in the belief formation process. The above results further show that conditional on the choice of information sources a model which includes imperfect data quality adds additional insights. For this the dynamic perspective is central. If volatilities are constant there exists an isomorphism linking the two models via the generally free cost parameter $\tilde{\omega}$.[5] In contrast, in a world of time-varying volatility this isomorphism breaks down. When volatility is time-varying a model with imperfect data will necessarily produce distinct responses. Hence, the above specification provides a rationale for why we may not wish to treat the reduced form parameter $\tilde{\omega}$ as constant with respect to time but as endogenous in the economy's underlying volatility regime.

Proposition 2 has important implications for how we may think about 'Bloom type' uncertainty shocks. Whether higher volatility raises or lowers attention levels crucially depends on the type of variance considered. This conclusion calls for caution when using general uncertainty indices as a proxy for 'Bloom type' second moment shocks. While the latter is concerned with changes in the variability of the fundamental, uncertainty indices may in fact capture changes in both volatilities, those of the underlying fundamental and the irreducible noise. A case in point could be the widely used VIX, the volatility of the stock market. One may argue that the source of those fluctuations in volatility is due to changes in σ_θ^2 as well as σ_ε^2. For instance, market volatility could be driven by an increase in the fraction of noise traders for reasons orthogonal to the fundamental. It may therefore be useful to devise indices that capture the relevant, effective uncertainty in the economy and decompose overall subjective volatility into its three components σ_θ^2, σ_ε^2 and σ_ζ^2.

Mechanism Given that losses are convex one may initially think that greater uncertainty of *any* type induces an agent to increase the precision of their prediction and thus their information processing. To see why this is not the case, I discuss the intuition behind the mechanism of the model below.

One way to think about the economic forces that are at play considers the impact of volatility changes on the cost-side of the transformed attention problem in (3). A deterioration (improvement) of information quality raises (lowers) the *effective* cost of information processing and incentivises the agent to lower (raise) their attention level.

There is a complementary explanation that focuses on the effects of volatility changes on the benefit side.[6] In effect, Proposition 2 captures that the agent's ability to influence its objective function fluctuates with changes in the underlying information quality. To see this, by the Total Law of Variance we may rewrite expected losses as

$$\mathbb{V}ar(\theta|s) = \mathbb{V}ar(\theta|x) + \mathbb{V}ar(\mathbb{E}[\theta|x]|s)$$

Now suppose information quality falls either through an increase in the variance of the irreducible noise or a decrease in the variance of the fundamental. There are two immediate effects. The uncertainty about the fundamental given the best possible signal x increases. That exposes the agent to larger losses. At the same time, the usefulness of x in predicting the underlying fundamental θ diminishes. The best prediction of the fundamental given the public signal, $\mathbb{E}[\theta|x]$, is shrunk towards the constant prior (its unconditional expectation $\mathbb{E}[\theta]$). Hence, the variability of $\mathbb{V}ar(\mathbb{E}[\theta|x]|s)$ vanishes. But the volatility of the best possible prediction is exactly the mechanism by which the agent controls the objective function. As this variability shrinks so does the agent's ability to influence their payoff function. Since information processing is costly the agent reduces their attention to the observable signal x as its power to predict the fundamental diminishes and marginal benefits fall.

[5] To obtain equivalence between models we can normalise the cost parameter as

$$\tilde{\omega} = \frac{\sigma_\theta^2\sigma_\varepsilon^2}{\sigma_\theta^2+\sigma_\varepsilon^2} + \omega$$

[6] Fuster et al. (2018) similarly note that net benefits fall as the irreducible noise in data is perceived to be more pronounced but do not explore this point further.

Table 1
Impact of the cross-derivative on the semi-elasticity of attention.

Semi-elasticity	$SNR > 1$	$SNR < 1$
$\Delta\kappa/\%\Delta\sigma_\theta^2$	Reinforcing in σ_ε^2	Dampening in σ_ε^2
$\Delta\kappa/\%\Delta\sigma_\varepsilon^2$	Dampening in σ_θ^2	Reinforcing in σ_θ^2

2.4. Interaction effects

The underlying variances enter the solution for the optimal attention level jointly through the squared correlation between the fundamental and the public signal. Hence, the presence of irreducible noise itself affects the dynamics of attention in response to volatility shocks to the fundamental variable. For the reader's convenience I restate the interior solution to the attention problem.

$$\kappa = \underbrace{\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_\varepsilon^2}}_{\substack{\text{(1) Correlational Effect}\\ =\mathbb{C}orr^2(\theta,x)}} - \underbrace{\omega\frac{1}{\sigma_\theta^2}}_{\substack{\text{(2) Fundamental Variance}\\ \text{Effect}}}$$

I define the relevant measure of volatility shocks to be the change in the natural logarithm of the underlying variances. This specification ensures that a unit shock to uncertainty is of the same magnitude (but clearly not sign) irrespective of whether $\log\sigma_i^2$, $\log 1/\sigma_i^2$ or the signal to noise ratio $\log(\sigma_i^2/\sigma_j^2)$ changes for $i, j \in \{\theta, \varepsilon\}$, $i \neq j$. Intuitively, we can up to a first order approximation interpret the shift in the attention level in response to a change in the log of the underlying volatilities in terms of their respective semi-elasticities.[7]

Corollary 1. *Define the interaction effect as the change in the semi-elasticity of attention when the other variance rises. The direction of the interaction effect is determined by whether the signal to noise ratio* $\lessgtr 1$.

Proof. The cross-derivative is given by

$$\frac{d^2\kappa}{d\log\sigma_\theta^2 d\log\sigma_\varepsilon^2} = \frac{\sigma_\theta^2\sigma_\varepsilon^2}{(\sigma_\theta^2+\sigma_\varepsilon^2)^3}(\sigma_\theta^2-\sigma_\varepsilon^2)$$

□

Corollary 1 states that whether the semi-elasticity of attention is amplified or dampened by larger values of the other variance depends on the initial information quality. The semi-elasticity of attention is maximised precisely at the threshold where the signal to noise ratio equals one.

Table 1 summarises whether the interaction effect is reinforcing or dampening the semi-elasticity of attention.

We can recognise the squared correlation as the sigmoid function $\mathbb{C}orr^2(\theta,x) = \frac{1}{1+\exp(-(z-y))}$ for $z \coloneqq \log\sigma_\theta^2$ and $y \coloneqq \log\sigma_\varepsilon^2$. The upward and downward concave parts of the sigmoid function represent the diminishing returns to greater or lower information quality respectively. Overall, this captures the notion that responses are muted at the extremes and amplified for middle-range values of variances.

Change in the fundamental's variance Let us consider how the response of attention to an increase in the fundamental's variance is impacted by the variance of the irreducible noise. The left panel in 3 plots the two components that determine the attention level against the log of the fundamental's volatility.[8]

A larger (smaller) irreducible noise variance is represented by a horizontal translation to the right (left) of the squared correlation function denoted by the black solid line. When information quality is initially good the agent is in the right tail of the sigmoid function in the high SNR parameter region. A horizontal translation to the right will move the agent out of the flat tail into the steeper part of the sigmoid function. All else equal, a steeper slope of the squared correlation implies a larger semi-elasticity of attention. The reverse holds if initially information quality is poor and the agent starts off in the flat tails in the low SNR parameter region. Combined those two observations confirm that the cross-effects are non-monotone and maximised when the signal to noise ratio equals one.

Change in the irreducible noise variance Let us now consider how the semi-elasticity of attention to an increase in the variance of the irreducible noise is impacted by the fundamental's variance. The left panel in 3 plots the attention level against the log of the irreducible noise volatility.

A larger (smaller) variance of the fundamental is represented by a shift of the attention function to the lower left (upper right). When information quality is initially good (poor) the agent is in the left (right) tail. A shift of the sigmoid function to

[7] Further, as costs are based on the reduction in the entropy of the prior and posterior distribution the natural way to classify volatility is by the log transform.

[8] Note that in the traditional RI model where $\sigma_\varepsilon^2 \to 0$ the black sigmoid function is constant at one and interaction effects are equal to zero.

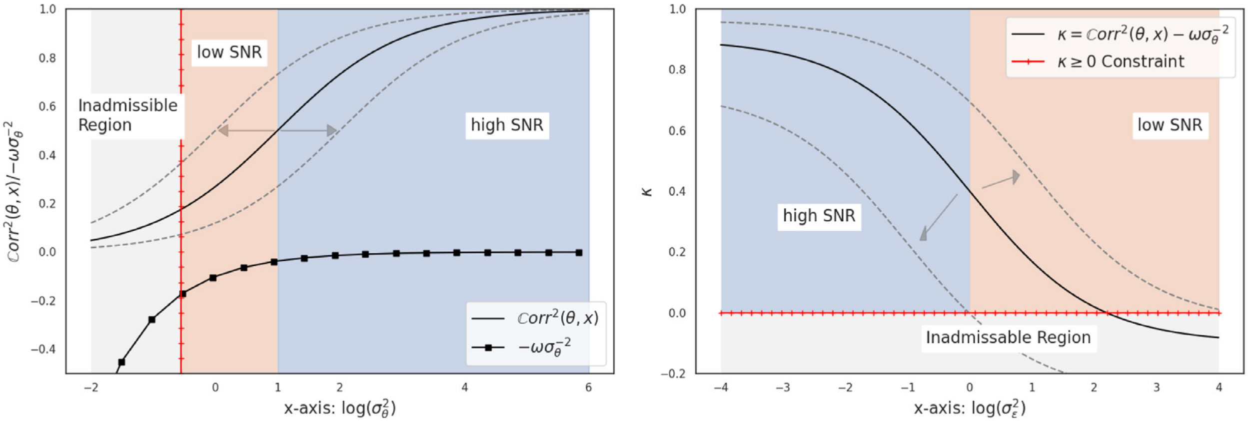

Fig. 3. Attention decomposition. *Left:* Attention level by components as a function of the fundamental variance. Low SNR denotes region where $\sigma_\theta^2 < \sigma_\varepsilon^2$. High SNR is defined analogously. *Right:* Attention As Function of Irreducible Volatility.

the lower left will move the agent into a flatter (steeper) region within the High (Low) SNR parameter space. The response of attention to a change in the volatility of the irreducible noise will thus be dampened (amplified). The converse holds if the fundamental variance falls. Combined those two observations again confirm that the cross-effects are non-monotone and maximised when the signal to noise ratio equals one. All cross-effects operate through a translation of the squared correlation function and represent the local returns to information quality.

Curvature Property The attention function is mostly concave in their own variance. The plots in Fig. 3 provide some intuition for this result. Note that the non-negativity constraint cuts off much of the region for which the square correlation function is most convex. Further, in the left panel of Fig. 3 the curvature is determined by an additional curve (square marker) which is strictly concave in its own variance. Both those effects tend to render the attention function concave in its own variance. As a result, a rational inattention model with irreducible noise suggests that attention is more responsive to volatility shocks that depress information quality than to those that improve the signal to noise ratio. That is information processing changes by more when the fundamental variance falls or the noise variance rises. This generalises from the traditional RI model.[9]

3. Application

I proceed to examine the macroeconomic implications of the rational inattention model with imperfect data quality in the dynamic setting. I first present three stylised facts about the dynamics of data quality over time. I then derive an attention-driven Phillips curve that nests Afrouzi and Yang (2021) as a limiting case and relates both the monetary policy and information quality regime of an economy to the slope of the Phillips curve.

3.1. Data

In this section, I empirically motivate the relevance of data quality for studying the attention strategies of firms and their interactions with the transmission of monetary policy. In particular, I examine the statistical properties of data revisions in the Real-Time Dataset for Macroeconomics by Croushore and Stark (2001) for quarterly inflation, nominal GDP growth and real GDP growth (see Table 2). I find a set of three results.

First, the measurement error in real-time statistics as measured by the difference between initial release and most recent data is substantial. I find that over the past 40 years signal to noise ratios mostly ranged somewhere between 1 and 3 for the period from 1984 to 2020. I use the same methodology to compute signal to noise ratios as in Lorenzoni (2009), who finds an estimate for the signal to noise ratio of quarterly PCE inflation that is even lower than the range above.[10] Overall, low signal to noise ratios indicate that a substantial portion of variation in real-time data is noise. This fact motivates the study of the macroeconomic implications of imperfect data quality.

Second, we observe an evident trend improvement in data quality as measured by the variance and the MAR over the past 55 years. For instance, for all three macroeconomic variables the revision error was approximately twice as large and volatile in the time period between 1965–1984 relative to 1984–2020. This reflects the advances in statistical methods, survey quality and the way data is collected.

[9] In the traditional RI model, the squared correlation term is constant at 1 and the curvature properties follow from the strict concavity of the $-\omega\sigma_\theta^{-2}$ function in the right panel of Fig. 3.

[10] I compute signal to noise ratios by first fitting an AR(2) process to the most recent data series, computing the variance of the residuals and dividing by the variance of revision errors as in Lorenzoni (2009).

Table 2
Revision statistics - from initial to most recent data (quarterly data at annual rates).

		1965–2020	1965–1984	1984–2020	2001–2020*	GM	GR	Recessions	2001–2007
Inflation	Variance	0.74	1.17	0.52	0.48	0.54	1.33	1.27	0.36
	MAR	0.66	0.85	0.56	0.53	0.59	1.04	0.90	0.57
	SNR	1.66	2.02	1.12	1.69	0.75	0.94	1.13	1.73
NGDP Growth	Variance	3.78	6.15	2.40	2.19	2.35	4.20	4.75	2.29
	MAR	1.51	2.01	1.26	1.17	1.28	2.02	1.66	1.15
	SNR	6.96	9.72	2.33	2.73	1.76	4.24	3.91	1.56
RGDP Growth	Variance	4.10	7.28	2.38	1.88	2.44	5.50	4.33	1.70
	MAR	1.55	2.16	1.24	1.10	1.28	1.84	1.68	1.07
	SNR	4.94	6.48	2.25	2.63	1.83	2.68	3.82	2.06

GR: Great Recession, GM: Great Moderation (1984–2007), *: 2001–2020 excludes the GR.

Counter to this long-run trend the empirical evidence suggests that data quality varies with economic activity. For example, the Great Recession saw a substantial increase in the variance of data revisions and the mean absolute revision (MAR). Relative to the preceding period of 2001–2007 the Great Recession saw mean absolute revisions that were nearly twice as large. Similar results hold for the variance of data revisions. It is worth noting that the countercyclicality of revision errors has been documented even before the Great Recession (Swanson and van Dijk, 2006). Heightened data uncertainty does not appear to be unique to the experience of the Great Recession. That is further documented by comparing the variance and MAR during all NBER Recessions and their full sample counterparts and noticing an increase in those two measures during economic downturns. More recently during the Covid crisis, concerns about the quality of real-time statistics have resurfaced highlighting that data quality shocks are not isolated to the experience of the Great Recession.

These findings are in line with an existent literature documenting the countercyclicality of revision errors. Marini and Shrestha (2013) show that the above patterns are also robust across countries and discuss some factors for why data quality might fall during recessions. First, statistics offices could derive their survey data from samples of businesses and households which may no longer be representative during times of rapid churn in the economy. Second, short-term indicators rely on projections based on assumptions and extrapolations derived from past data. During recessions statistics offices may need to depart from these established methods as economic conditions change drastically. Together these factors introduce greater uncertainty into real-time data during an economic downturn. Another theoretic mechanism for why data quality might be pro-cyclical is proposed in Fajgelbaum et al. (2017) and Van Nieuwerburgh and Veldkamp (2006). Economic activity naturally generates information when, for example, firms undertake investments and the return on their projects is made public to other market participants. In absence of investment activity no such inference can be made and less information is being generated naturally.

The combination of empirical evidence from the Real-Time Dataset for Macroeconomics, references from the literature and a discussion of possible theoretical mechanisms for the decline in data quality during economic downturns motivate the study of a macroeconomic application of a rational inattention model with time-varying irreducible noise.

3.2. Model environment

The Economy The economy is standard and I refer the reader to Afrouzi and Yang (2021) for a more detailed exposition. In its reduced form the log-linearised Euler Equation for the representative household and the monetary policy rule are given by

$$i_t = \rho + \mathbb{E}_t^f \Delta q_{t+1} \tag{Euler Equation}$$

$$i_t = \rho + \phi \Delta q_t - \sigma_u u_t, \qquad u_t \overset{iid}{\sim} \mathcal{N}(0, 1) \tag{MP Rule}$$

where $q_t \coloneqq \log Q_t - \log \bar{Q}$ denotes nominal aggregate demand as the log deviation from the steady state, u_t the monetary policy shock, ρ the discount rate of the representative household and $\mathbb{E}_t^f$ the fully frictionless expectations of households. Let $\phi > 1$. Aggregate nominal demand is uniquely pinned down by the history of monetary policy shocks and follows the following random walk process (see Lemma 3.1 in Afrouzi and Yang, 2021).

$$q_t = q_{t-1} + \frac{\sigma_u}{\phi} u_t$$

The above policy rule has the property that the monetary authority directly controls nominal demand. It yields a stochastic process that is characterised by the primitives of the monetary policy rule. Notably, the volatility of nominal demand is decreasing in the stabilising term ϕ and increasing in the variance of monetary policy shocks σ_u^2. Following Afrouzi and Yang (2021) I interpret changes in the underlying variance of nominal aggregate demand as a change in the policy regime. In particular, policy regimes can be classified as follows.

Definition 2. A dovish (hawkish) monetary policy regime is characterised by a high (low) value $\frac{\sigma_u^2}{\phi^2}$.

Note that the results below are not exclusive to monetary policy shocks. In fact, the monetary policy shocks in the model act more generally as a stand in for a wider class of demand shocks. It follows that while the discussion below focuses on the implications of a shift in the monetary policy regime the same model (and its implied results) can be used more widely to address the effects of a volatility shock to nominal demand and its interaction with the information quality regime of an economy.

Firms. There is a unit mass of price-setting firms $i \in [0,1]$, each producing a variety that is sold under monopolistic competition. Prices are perfectly flexible but firms are rationally inattentive. Each firm i learns about nominal demand from a common public signal $x_t = q_t + \varepsilon_t$, $\varepsilon_t \overset{iid}{\sim} \mathcal{N}(0, \sigma_\varepsilon^2)$. Data sources that are informative about nominal demand directly are not available. For instance, firms may choose how much attention to pay to the release of real-time data by national statistics reports or how closely to read the reports produced by their respective central banks. As a straight forward adaptation of Afrouzi and Yang (2021) the rational inattention problem of firm i is given by

$$V\left(s_i^{-1}\right) = \max_{\{p_{i,t}, s_{i,t} \in \mathcal{S}^t\}_{t=0}^{\infty}} \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \left[\mathbb{E}\left[-(\lambda-1)(p_{i,t}-q_t)^2 \middle| s_i^t\right] - 2\omega I\left(x^t; s_i^t \middle| s_i^{t-1}\right)\right] | s_i^{-1}\right] \tag{5}$$

where $s_i^t \coloneqq \{s_{i,\tau}\}_{\tau=0}^t \cup s_i^{-1}$ denotes the entire history of firm i's signals up to period t. Parameter λ^{-1} is a per worker subsidy that corrects for the steady state distortions from monopolistic competition. Costs are specified to be proportional to the Shannon mutual information function $I(x^t; s_i^t | s_i^{t-1})$.

3.3. The full model

Maćkowiak et al. (2018a) and Afrouzi and Yang (2021) show that in the standard LGQ case with Gaussian state variables we can without any loss of generality restrict our attention to Gaussian signals. Proposition 3 shows that these results carry over to the dynamic rational inattention model with irreducible noise.

Proposition 3. *Given that $q^t \to x^t \to s^t$ forms a Markov chain and that q^t and x^t are jointly Gaussian then the optimal sequence of signals $s^t \in \mathcal{S}^t$ solving (5) follows a Gaussian distribution.*

The solution to (5) is complicated by the fact that the flow utility of the firm depends on the entire covariance matrix of past histories of public signals.[11] A priori it is not evident whether agents would find it optimal to learn about past values of public signals. Intuitively, information about past states x^{t-1} may be used to 'average out' the measurement error contained in the current public signal and increase knowledge about the fundamental. To thus render the problem tractable I restrict the learning technology to process information about a fixed, finite past. In particular, I make the assumption that firms only process information about current public signals. That is $x^{t-1} \perp s_t | s^{t-1}, x_t$.

Note that this assumption captures the notion that agents only learn from the most accurate public signal about the fundamental. As past public signals contain additional noise about the Law of Motion of the fundamental firms following this heuristic will only attend to the current public signal.[12] Proposition 9 of Myatt and Wallace (2012) provides some intuition for why we might expect this to be a good approximation. The authors prove that in the corresponding static setting with independent public signals agents optimally choose to ignore all but the most accurate public signal.

With the model preliminaries discussed we can now proceed to characterise the solution to the firm's attention problem. As firms are identical in the model I drop the i subscript unless relevant. To further simplify notation, let $X_t \coloneqq \mathbb{V}ar(x_t|s^{t-1})$ denote the prior uncertainty and $Y_t \coloneqq \mathbb{V}ar(x_t|s^t)$ the posterior uncertainty of a firm. The problem simplifies as follows. Producers choose a sequence of posterior variances $\mathbb{Y} = \{Y_\tau\}_{\tau=0}^{\infty}$ to maximise the discounted value of expected future profits for a given level of initial uncertainty X_0

$$J(X_0, \mathbb{Y}) = \sum_{t=0}^{\infty} \beta^t \left[-(\lambda-1)\mathbb{V}ar(q_t|s^t) - \omega \log\left(\frac{X_t}{Y_t}\right)\right] \tag{6}$$

subject to

$$X_{t+1} = \mathbb{V}ar(q_t|s^t) + \frac{\sigma_u^2}{\phi^2} + \sigma_\varepsilon^2 \tag{a}$$

$$\mathbb{V}ar(q_t|s^t) = \sigma_\varepsilon^2 \delta(X_t) + \delta(X_t)^2 Y_t \tag{b}$$

[11] To see this, note that the Markov chain property exploited in the proof of Proposition 3 allows us to rewrite $\mathbb{V}ar(q_t|s^t) = \mathbb{V}ar(q_t|x^t) + \sum_{i,j \geq 0} \alpha_i \alpha_j \mathbb{C}ov(x_{t-i}, x_{t-j}|s^t)$ for some constants α_i, α_j.

[12] Some intricacies arise when data quality is time-varying as this allows for the possibility of past signals to be most accurate about the fundamental. But while I believe this endogenous adjustment margin of agents to be an interesting pursuit a treatment of this is left for future research.

$$0 \le Y_t \le X_t \tag{c}$$

for all $t = 0, 1, \dots$ and where

$$\delta(X_t) \coloneqq 1 - \frac{\sigma_\varepsilon^2}{X_t} \tag{7}$$

Constraint (a) uses the Law of Motion for nominal demand and establishes the intertemporal link between periods. Constraint (b) is an application of the Total Law of Variance and (c) incorporates both the no-forgetting constraint that the posterior variance can be no larger than the prior belief's variance as well as the non-negativity property of variances.

Role of $\delta(X_t)$. Eq. (7) captures the informativeness of the public signal and formally denotes $\mathbb{C}orr^2(q_t, x_t|s^{t-1})$. The weight $\delta(X_t) \le 1$ plays a central role in the firm's attention problem as it modulates the rate with which

(i) a given posterior uncertainty about the public signal maps into the posterior uncertainty about the fundamental via the Law of Total Variance
(ii) the posterior uncertainty translates into next period's prior uncertainty via the Law of Motion

The difference in the dynamics of the standard model without and the extended model with noisy data follow precisely from the interplay between the underlying variances and the quality of information $\delta(X_t)$.

Recursive Formulation. Firms enter each period with prior uncertainty X_t about the public signal. Due to the LQG nature of the problem choosing the optimal pricing strategy is identical to choosing the optimal posterior variance of the observable and, as a corollary, from the viewpoint of firms the state is fully captured by the prior variance. The recursive structure can be expressed as the dynamic program[13]

$$V(X) = \max_Y -(\lambda - 1)\left(\sigma_\varepsilon^2\delta(X) + \delta(X)^2 Y\right) - \omega \log\left(\frac{X}{Y}\right) + \beta V\Big(\overbrace{\delta(X)^2 Y + \Psi(X)}^{=X'}\Big) \tag{8}$$
$$\text{subject to } 0 \le Y \le X$$

where

$$\Psi(X) \coloneqq \sigma_\varepsilon^2 \delta(X) + \frac{\sigma_u^2}{\phi^2} + \sigma_\varepsilon^2$$

Combining the first order condition with an application of the Benveniste-Scheinkman condition yields optimality relation (9) for the sequence of reservation uncertainties $\{\bar{Y}_\tau\}_{\tau=0}^\infty$. Due to the presence of the no-forgetting constraint the optimal posterior variance is given by $Y_t^* \equiv \min\left\{\bar{Y}_t, X_t\right\}$. That is the firm's prior uncertainty must exceed a reservation uncertainty of $\bar{Y}_t$ for them to process information.

$$0 = \underbrace{-(\lambda-1)\delta(X_t)^2}_{(1)} \underbrace{+\omega\frac{1}{\bar{Y}_t}}_{(2)} \underbrace{-\beta\omega\delta(X_t)^2\left[\frac{1}{X_{t+1}} + \frac{1}{\delta(X_{t+1})^2 Y_{t+1}^*}\frac{\partial \mathbb{V}ar(q_{t+1}|s^{t+1})}{\partial X_{t+1}}\right]}_{(3)=\beta\delta(X_t)^2 V'(X_{t+1})} \tag{9}$$

where

$$\frac{\partial \mathbb{V}ar(q_{t+1}|s^{t+1})}{\partial X_{t+1}} = \frac{\sigma_\varepsilon^2}{X_{t+1}^2}\left[\sigma_\varepsilon^2 + 2\delta(X_{t+1})Y_{t+1}^*\right]$$

The optimality condition can be decomposed into three terms

(1) The first term represents how an additional decrease in the variance about the public signal translates into a reduction in expected *contemporaneous* losses.
(2) This denotes the direct *contemporaneous* marginal cost of a reduction in the uncertainty about the public signal.
(3) The last term denotes the firm's *intertemporal* considerations. The continuation value is discounted by (i) the standard discount factor for time preferences β and (ii) $\delta(X_t)^2$ to account for the fact that a marginal change in the posterior variance does not translate one to one into a reduction in the prior uncertainty the next period.

I define (1) as the marginal benefit of reducing the posterior variance about the public signal and the sum of (2) and (3) as the corresponding net marginal cost. This interpretation follows from the fact that the intertemporal incentives $\beta\delta(X_t)^2 V'(X_{t+1})$ are proportional to the information processing cost parameter ω. This implies that the marginal value of entering a period with lower prior uncertainty is derived from the cost savings as less new information needs to be acquired. As a corollary, by regulating the cost parameter ω one can modulate the importance of intertemporal effects in the firm's decision making, all else equal.

[13] Details on the properties of the solution to the functional equation are provided in Appendix F.

Policy function The solution defines the optimal reservation uncertainty $\bar{Y}_t$ as a function of the firm's prior variance X_t. The policy function exhibits the following properties

P.1 The reservation uncertainty is decreasing in the prior variance – i.e. $\frac{d\bar{Y}_t}{dX_t} < 0$.
P.2 For a given prior variance X_t, larger underlying volatilities σ_u^2/ϕ^2 and σ_ε^2 imply a larger reservation uncertainty.

The policy function differs in important ways from the case of perfect data quality. In the latter the optimal posterior variance is in fact independent of the prior and hence property P.1 does not hold. This fact is central to understanding how data imperfections impact the response of the economy to volatility shocks. For example, in a model with irreducible noise a higher monetary policy shock variance will induce both a shift in the policy function by P.2 and, countervailing that effect, a movement along the policy function by P.1 as the prior variance X_t increases due to the higher volatility regime. Whether the overall effect will be positive or negative depends on the relative magnitude of the shift in the function relative to the movement along it. Further, since P.1 does not hold in the case of perfect data quality the latter effect will not be present in such a scenario and the posterior uncertainty will unambiguously rise in response to a volatility shock.

Dynamic vs. static model As $\beta \to 0$ the FOC (9) converges to that of the static model. For $\beta > 0$ the presence of intertemporal incentives increases attention as more precise signals today carry additional informational value for the following period. These effects operate through two channels. First, a lower prior variance in the next period implies that a given posterior uncertainty about the public signal is mapped into a lower posterior uncertainty about the fundamental. Further, a lower prior variance reduces the cost of attaining a given target posterior uncertainty. The relative optimal posterior variances when the solution to the dynamic problem is interior is[14]

$$\frac{\bar{Y}_t^{stat}}{\bar{Y}_t^{dyn}} = \left(1 + \frac{-\beta V'(X_{t+1})}{\lambda - 1}\right) \min\left\{1, \frac{\delta(X_t)^2(\lambda - 1)}{\omega} X_t\right\} \geq 1 \tag{10}$$

Most notably the presence of intertemporal incentives implies that firms find it optimal to process new information in situations in which a static model would predict them to ignore all new data. That follows from the fact that the slope of the value function can be steep even at the steady state, especially when the marginal cost ω is high, the signal to noise ratio low or total volatility low. The intuition is straightforward. The larger the marginal cost to process information the larger the potential cost savings from entering the next period with a more precise prior. Second, the worse data quality the less informative current signals are about the fundamental and the greater the relative importance of optimising the prior. At the same time a high irreducible noise variance exposes the agent to greater future risk and thereby further increases the value of a more precise prior relative to the static counterpart.[15]

Optimal pricing Given the quadratic loss function firm i optimally sets its price equal to the posterior expectation of current nominal demand. By Proposition 3 and since agents process information about current public signals only we can restrict our attention to signals of the form $s_{i,t} = x_t + \zeta_{i,t}$, $\zeta_{i,t} \sim \mathcal{N}(0, \sigma^2_{\zeta,t}(i))$.[16] Hence,

$$p_{i,t} = p_{i,t-1} + \kappa_t(q_t + \varepsilon_t + \zeta_{i,t} - p_{i,t-1})$$

and the Kalman gain in the economy is given by

$$\kappa_t = \max\left\{\delta(X_t)\left(1 - \frac{\bar{Y}_t}{X_t}\right), 0\right\}$$

Aggregation Suppose all firms start with the same prior uncertainty $X_0 = \mathbb{V}ar(x_0|s^{-1})$. We can distinguish between two cases: when the no-forgetting constraint is binding and when the solution is interior. The dynamic structure of inflation and output is given by the following set of equations respectively.

1. Suppose there exists a period T such that $\bar{Y}_t \geq X_t$ for all $t \leq T$, then for all $t \leq T$

$$\pi_t = 0$$
$$y_t = y_{t-1} + \frac{\sigma_u}{\phi} u_t$$

2. Suppose $\bar{Y}_T < X_T$, then for all $t \geq T+1$

$$\pi_t = \frac{\kappa_t}{\kappa_{t-1}}(1 - \kappa_{t-1})\pi_{t-1} + \kappa_t\left(\frac{\sigma_u}{\phi} u_t + \Delta\varepsilon_t\right)$$
$$y_t = (1 - \kappa_t)y_{t-1} + (1 - \kappa_t)\frac{\sigma_u}{\phi} u_t - \kappa_t\varepsilon_t$$

where $\Delta\varepsilon_t \equiv \varepsilon_t - \varepsilon_{t-1}$.

[14] Note that the difference in the amount of information processed is the log transform of this ratio, i.e. $\mathbb{I}^{dyn} - \mathbb{I}^{stat} = \frac{1}{2}\log(X_t/\bar{Y}_t^{dyn}) - \frac{1}{2}\log(X_t/\bar{Y}_t^{stat}) = \frac{1}{2}\log\frac{\bar{Y}_t^{stat}}{\bar{Y}_t^{dyn}} \geq 0$.

[15] Further details are provided in Appendix C.

[16] Note that any other Gaussian signal $\tilde{s}_{i,t} = \alpha_t x_t + \tilde{\zeta}_{i,t}$ yields the same posterior distribution as $s_{i,t} = x_t + \zeta_{i,t}$ for $\zeta_{i,t} \equiv \tilde{\zeta}_{i,t}/\alpha_t$.

3. The Phillips curve is given by

$$\pi_t = \frac{\kappa_t}{1-\kappa_t}(y_t + \varepsilon_t)$$

Dynamics and transmission of shocks Short-run and long-run dynamics differ in two important ways. First, the no-forgetting constraint gives rise to the possibility of a corner solution in which macroeconomic variables are perfectly persistent and the slope of the Phillips curve zero. Yet, even when the solution is interior the dynamics for inflation depend non-trivially on current and past levels of attention. Whereas in the steady state the persistence of inflation is decreasing in the Kalman gain in the short-run a rise in today's attention levels may even render inflation more persistent. This fact has material implications for the shape of the impulse response of inflation to a monetary policy shock when the underlying volatility is time-varying as we will discuss in Section 4.3. This highlights the importance of taking seriously the full dynamics of the Kalman gain and not reduce them to their steady state behaviour as it is commonly done in the literature.

Impulse responses to the monetary policy shock inherit the standard properties from demand shocks. The non-fundamental shock is isomorphic to a standard cost-push shock contemporaneously, which pushes up inflation while depressing real output. But it differs qualitatively over longer horizons as what is important for inflation dynamics is not the realisation of the shock itself but rather its first difference $\Delta\varepsilon_t$.

3.4. Long-run dynamics

It can be shown that under suitable regularity conditions the dynamic system is well-behaved and converges to a unique steady state. Lemma 1 formalises this result.

Lemma 1. *Steady state properties are as follows*

1. *Let the optimal next period's prior variance be denoted by $g(X)$ and the optimal policy function for the posterior variance by $h(X)$. Suppose that $\beta V''(g(X)) \leq \frac{\omega(1-(\sigma_\varepsilon^2/X)^2)}{[\delta(X)^2 h(X)]^2}$. There exists global convergence to the unique fixed point of the Law of Motion of the prior variance.*
2. *The Kalman gain remains at zero and the economy perfectly persistent at most for a finite number of periods. In the long-run, steady state information processing is positive and given by*

$$\kappa_\infty = \frac{\sigma_u^2/\phi^2}{X_\infty - \sigma_\varepsilon^2} > 0$$

The condition on the convexity of value function in Lemma 1 differs from the second order condition for a local maximum only by a factor $1-\left(\sigma_\varepsilon^2/X\right)^2$. The condition itself is non-trivial as it is endogenous through the policy functions but numerical simulations confirm that it is satisfied.

3.5. Comparative statics

Comparative statics carry over straight forwardly from the static model. It can be numerically verified that the steady state Kalman gain is increasing in σ_u^2/ϕ^2 and decreasing in σ_ε^2 and ω. I further establish an analogue to the isomorphism result in Section 2.2. Panel A in Fig. 4 shows the contour plot for the steady state Kalman gain for a range of different data quality and information cost regimes. For any $(\sigma_\varepsilon^2, \omega)$ pair there exists a corresponding $(0, \tilde{\omega})$ pair with $\tilde{\omega} > \omega$ that attains the same Kalman gain. Hence, in the long-run the cost parameter can fully capture imperfect data quality. However, as in the static case, this tight correspondence may not hold in the short-run if the economy is subject to volatility shocks.

Panels B. and C. respectively plot the impulse responses of inflation and output to a one standard deviation MIT shock to monetary policy. As is intuitive from the above comparative statics poorer data quality or greater information processing costs imply larger and more persistent output deviations. These have important implications for the transmission of monetary policy.

Implications Over the past decades statistics offices and central banks have spent considerable effort to improve data quality and the way they communicate their views on the economic outlook to the public. These efforts can be seen as improving data quality by lowering the irreducible noise in real-time statistics. Those forces raise attention levels, induce a steeper Phillips curve and insofar as monetary policy shocks are a good proxy for monetary policy transmission imply that central banks have a more powerful policy instrument to control inflation.

What is more, a trend increase in data quality can be found to reduce the unconditional variances of inflation and output as given by Corollary 2. With better data quality firms coordinate less on non-fundamental variation even as they pay more attention to these shocks. As a consequence, macroeconomic fluctuations are more muted. A high data quality regime therefore allows the central bank to keep inflation and output closer to target, highlighting a key mechanism in which providing more precise information to the public can help the central bank better meet their mandate.

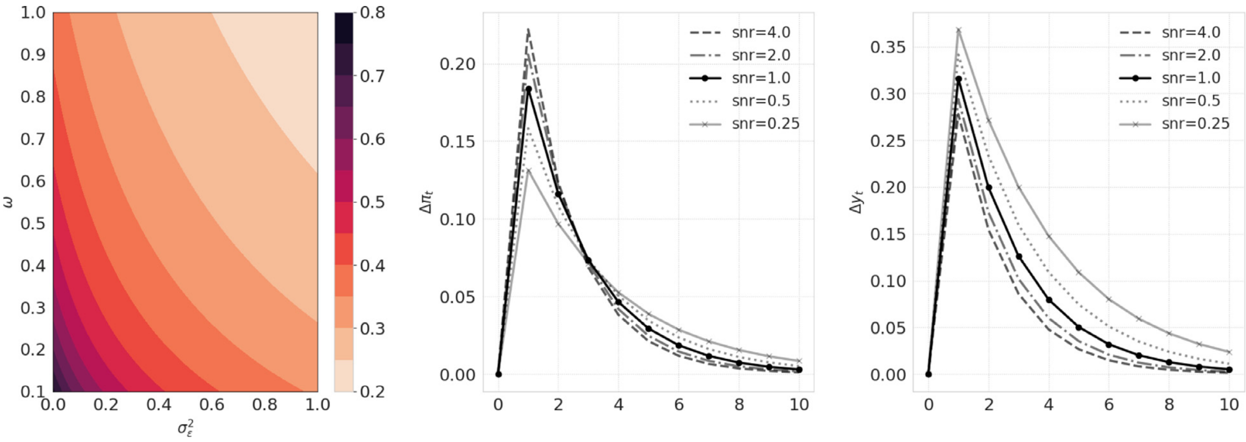

Fig. 4. Data quality and the transmission of monetary policy. *Panel A (Left):* Contour plot of the steady state Kalman gain. *Panel B (Centre):* Impulse response of inflation to monetary policy shock. *Panel C (Right):* Corresponding impulse response of the output gap.

Corollary 2. *The long-run variance of inflation and output are respectively given by*

$$\mathbb{V}ar(\pi) = \frac{\kappa}{2-\kappa}\left(\frac{\sigma_u^2}{\phi} + 2\kappa\sigma_\varepsilon^2\right)$$
$$\mathbb{V}ar(y) = \frac{(1-\kappa)^2}{\kappa(2-\kappa)}\frac{\sigma_u^2}{\phi} + \frac{\kappa}{2-\kappa}\sigma_\varepsilon^2$$

4. Volatility shocks

In the comparative statics exercises below I initialise the economy to start off at its steady state. To reduce the notational burden I set $\lambda = 2$ without any loss of generality.

4.1. More dovish monetary policy

Recall that a switch to a more dovish monetary policy is characterised by an increase in σ_u^2/ϕ^2.

Mechanism. Initially a rise in the volatility of the monetary policy shock raises the firm's prior uncertainty X_t. This improves the informativeness of public signals $\delta(X_t)$ and increases the marginal benefit of learning. At the same time, all else equal, a higher volatility reduces the firm's incentive to learn as today's information is less informative about the next period. This force counteracts the higher *contemporaneous* marginal benefits by shifting the *intertemporal* net marginal cost schedule outwards. The direction of the total change depends on the initial level of data quality and the cost parameter ω. The following Lemma summarises.

Lemma 2. *Suppose the economy is at its steady state initially and σ_u^2/ϕ^2 increases. Both the contemporaneous marginal benefit and the intertemporal net marginal cost of reducing Y_t increase.*

Lemma 2 shows that the response of the posterior variance to an increase in the monetary policy shock variance is ambiguous. For the posterior variance to fall, i.e. $dY_t^* < 0$, a positive $\sigma_\varepsilon^2 > 0$ and ω not too large are jointly necessary while neither is sufficient.

The sources of time-variability in the Kalman gain are two-fold. First, there is a *mechanical* effect in that a firm which arrives at the same posterior uncertainty from a higher prior uncertainty must necessarily process more information as the signal to noise ratio improves. Second, there is an additional *endogenous* effect as firms adjust their posterior uncertainty optimally. In the presence of perfect data quality the latter unambiguously dampens the response of the Kalman gain to the volatility shock. In contrast, if data were imperfect the endogenous adjustment channel may even be a source of amplification. To see this note that the total change in the level of attention is given by

$$\frac{d\kappa_t}{d\sigma_u^2/\phi^2} = \underbrace{\frac{\partial\kappa_t}{\partial X_t}}_{>0}\underbrace{\frac{dX_t}{d\sigma_u^2/\phi^2}}_{=1} + \underbrace{\frac{\partial\kappa_t}{\partial Y_t}}_{<0}\underbrace{\frac{dY_t}{d\sigma_u^2/\phi^2}}_{\lesseqgtr 0} \tag{11}$$

The decomposition highlights that a rise in the posterior variance does not necessarily imply a fall in attention. Only if the rise in the posterior variance is sufficiently large relative to the rise in the prior variance does the Kalman gain fall. That is, the effects operating through the intertemporal net marginal cost channel must be sufficiently strong.

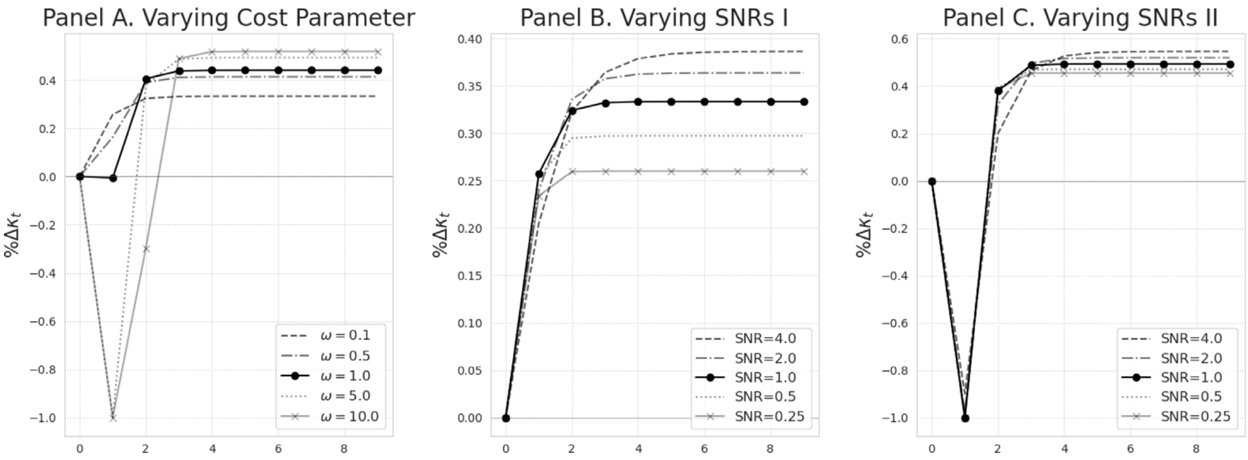

Fig. 5. Effect of fundamental volatility shock on attention. *Panel A.* – $(\sigma_u^2/\phi^2)_{t=0} = 0.25$, $(\sigma_u^2/\phi^2)_{t\geq 1} = 0.25$, $\sigma_\varepsilon^2 = 0.5$. *Panel B.* – $\omega = 0.1$. Panel C. – $\omega = 5$.

Impulse responses I consider a 100% increase in the variance of the monetary policy shock and compute the impulse responses for the Kalman gain for varying cost parameters and signal to noise ratios. Fig. 5 plots the respective impulse responses as percentage deviations from initial steady state values. As discussed in Section 3.5 in the long run steady state attention rises. In the short run the dynamic behaviour crucially depends on the cost parameter ω. Varying signal to noise ratios impact the magnitude of the shifts in attention, but do not qualitatively alter the results.

This result can be reconciled with our earlier finding that information quality affects the firm's optimal attention allocation. The signal to noise ratio can determine whether the posterior uncertainty about the public signal falls or rises in response to an increase in the monetary policy shock variance. However, those effects are quantitatively too small relative to the change in the prior uncertainty to have qualitative implications for the Kalman gain, though they can act to amplify impulse responses.

4.2. Comparison between models

In this section I compare the instantaneous response of the Kalman gain to an increase in the monetary policy shock variance under two alternative models: i) traditional dynamic rational inattention and ii) dynamic rational inattention with imperfect data. In particular, I show that when data sources are noisy (model ii) the observed flattening of the Phillips curve is not likely due to a more dovish monetary policy as traditional dynamic RI (model i) may suggest, underscoring the importance of the interaction effects between volatilities.

Information Cost Adjustment Since the presence of irreducible noise mechanically increases the utility cost of information processing care must be taken when comparing models. I thus adjust the cost parameter in the standard model as to equate the steady state Kalman gains prior to the realisation of the volatility shock. Effectively this implies a larger cost parameter in the standard model. This ensures that the structural features of the initial steady state economies are identical between models. An alternative specification in which I account for total volatility yields identical qualitative implications.

Results The response of attention to a rise in the monetary policy shock variance is non-monotonic. This holds for both the traditional model and the extended model with noisy data. In particular, the size of the cost parameter ω determines whether the Kalman gain rises or falls on impact. Numerical simulations show that for each $(\sigma_u^2/\phi^2, \sigma_\varepsilon^2)$ pair there exists a $\bar{\omega}$ such that for all $\omega \geq \bar{\omega}$ the Kalman gain falls temporarily (possibly to zero) before it gradually converges to a new and higher steady state value.

The threshold $\bar{\omega}$ is plotted in Fig. 6 as a function of σ_u^2/ϕ^2 for three signal to noise ratios: High, Medium and Low with signal to noise ratios of $4, 1, 1/4$, respectively.[17] There is a corresponding threshold for each of the models. For the dynamic model with noisy data $\bar{\omega}$ is denoted by the border between the red and the white shaded region. For the dynamic model with perfect data $\bar{\omega}$ is denoted by the border between the white and the blue shaded region. In the simulation I consider a 100% increase in the monetary policy shock variance.

The white shaded area in Fig. 6 denotes the region for which the Kalman gain is falling in a model without irreducible noise but rising in a model that includes such concerns for imperfect information quality. This parameter space is larger the lower the signal to noise ratio. If noisy data is prevalent, then the choice between the two models matters for the qualitative behaviour of the economy in response to a change in the monetary policy regime. Conversely, for high signal to

[17] This range seems empirically plausible. Lorenzoni (2009) finds an estimate of 0.5 for PCE inflation data. Evidence presented in Section 3.1 finds signal to noise ratios approximately ranging between below 1 to 3.

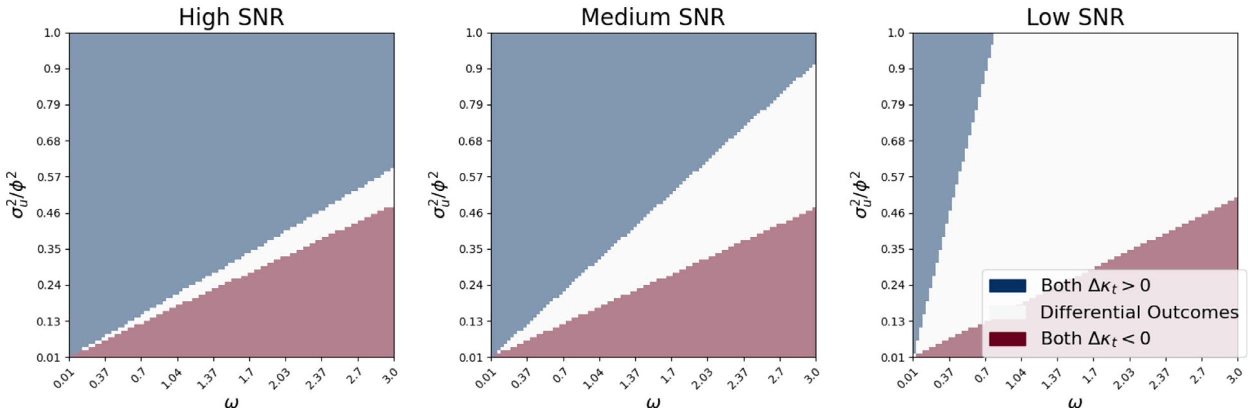

Fig. 6. Relative impulse responses of attention to fundamental variance shock. *Notes:* Adjusted Cost-Structure. 100% increase in initial monetary policy shock variance. 100×100 grid for $\omega \in [0.1, 3]$ and $\frac{\sigma_u^2}{\phi^2} \in [0.01, 1]$. SNR defined as $\frac{\sigma_u^2/\phi^2}{\sigma_\varepsilon^2}$. High SNR = 4, Medium SNR = 1, Low SNR = 0.25.

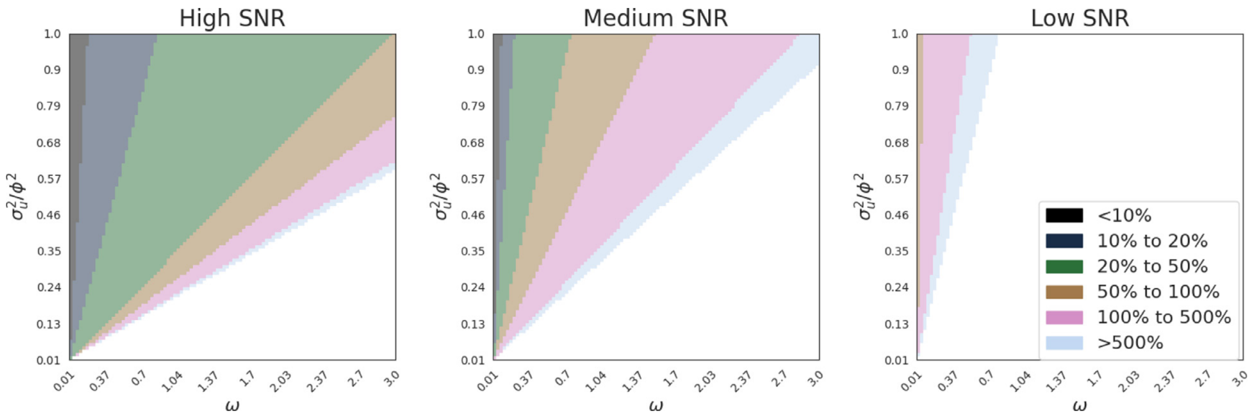

Fig. 7. Amplification of attention response. *Notes:* Limited to region where Kalman gain is rising under both models – i.e. to be read as the change in the Kalman gain is 50% to 100% greater in a model with irreducible noise relative to a model with perfect data. Specifications as in Fig. 6.

noise ratios the difference in the choice between models seems to be largely immaterial for the *qualitative* results for most of the parameter space.

Amplification More generally the introduction of irreducible noise has large *quantitative* effects as it acts to amplify instantaneous responses in the attention level of firms. Fig. 7 shows that the amplification is substantial, mostly ranging between 20% to 100%.

Taking stock The above discussion highlights that the composition of information processing costs has material dynamic implications over and above total effective costs. A fixed cost parameter $\tilde{\omega}$ set to normalise the Kalman gain to the value implied by a model with irreducible noise cannot generate the same dynamic behaviour. When costs from imperfect information quality are substantial the two dynamic rational inattention models are qualitatively different. This conclusion extends to the whole parameter space denoted by the white shaded region in Fig. 6. The above discussion indicates that the mechanism in Afrouzi and Yang (2021) – where the flattening of the Phillips curve is driven by a more dovish monetary policy in the wake of the Great Recession – is weaker in a setting with non-negligible noise in the data. What is more, it may be precisely such periods of stark economic contraction during which we would expect the noisiness of observable data to be more pronounced.

4.3. A fall In data quality

The two-layered *'signal of a signal'* structure provides a framework to examine the impact of information quality shocks on the real economy. I propose that fluctuations in the information content of public signals shed light on an alternative mechanism for why the slope of the Phillips curve can flatten during recessions.

Mechanism As greater irreducible noise in the data lowers the quality of information $\delta(X_t)$ the firm's marginal benefit of learning falls. At the same time net marginal costs increase. The latter is not straightforward as there are two competing

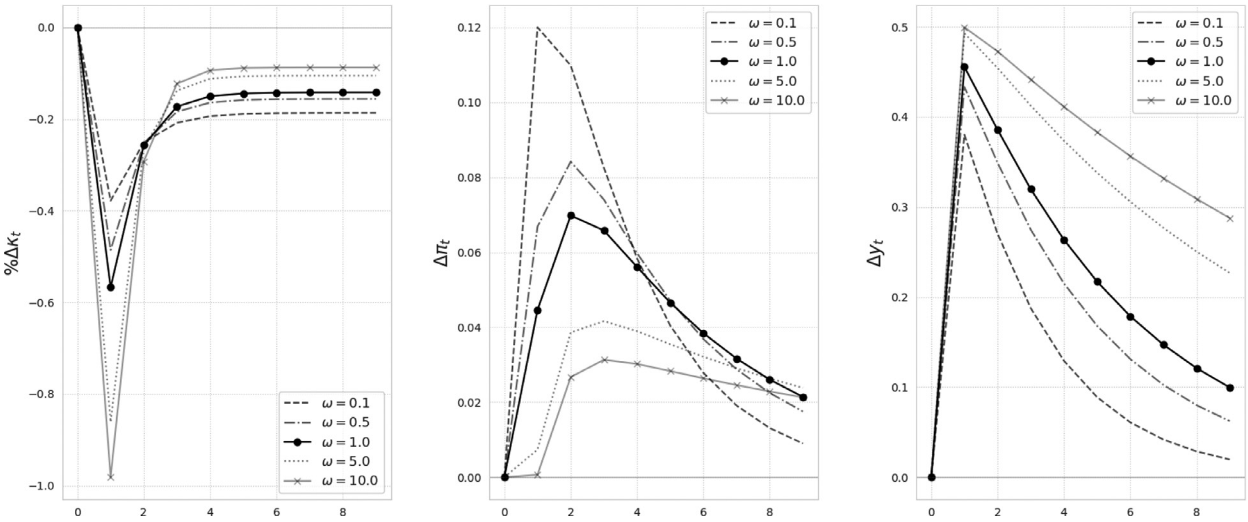

Fig. 8. Effect of fall in data quality on attention and transmission of monetary policy. *Notes:* $\sigma_u^2/\phi^2 = 0.25$, $\sigma_{\varepsilon,t=0}^2 = 0.25$, $\sigma_{\varepsilon,t\geq 1}^2 = 0.5$. *Panel A (Left):* Response to the Kalman gain. *Panel B (Centre):* Response to inflation *Panel C (Right):* Response to the output gap.

effects. On the one hand, greater irreducible noise for a given prior uncertainty exposes the firm to larger losses in the future which they would want to guard against. On the other hand, since today's information contains greater noise it is less useful to predict future states, thus lowering the firm's incentive to process information. Simulations find that the latter effect dominates, net marginal costs rise and attention falls.

Impulse responses Panel A. of Fig. 8 shows the impulse responses of attention to a 100% increase in the variance of the irreducible noise for varying cost parameters and signal to noise ratios. Responses are denoted as percentage deviations from initial attention values. The response is monotonic. A rise in the volatility of the irreducible noise raises the optimal posterior variance and lowers the Kalman gain across all specifications. Attention levels partially recover from their initial precipitous drop but remain at perpetually lower levels. The effect on attention levels is persistent and does not dissipate over time.

This stresses the importance of understanding the source of the flattening of the Phillips curve. A fall in attention due to a more dovish monetary policy stance has starkly different long-run implications relative to a situation when lower information processing is driven by an increase in the volatility of the irreducible noise. In the former case, the Kalman gain eventually recovers to a level higher than initially. In the latter, attention and the slope of the Phillips curve remain depressed over all time horizons.

The transition to a lower data quality regime has substantive implications for the transmission of monetary policy. Panel B. and C. respectively plot the impulse responses of inflation and output to a one standard deviation MIT shock to monetary policy as the economy is transiting to the low data quality steady state. Compared to the impulse responses for the steady state Kalman gain in Fig. 4 the deviations in output are larger and the impulse response of inflation follows a hump-shape. This implies that the monetary policy shock is more persistent with larger deviations from the optimum along the transition path than in steady state.

The hump-shape of the impulse response of inflation is an implication of the fact that the dynamics for inflation depend on the current and the last period's Kalman gain. A model based on steady state values of Kalman gains as is common in the literature (Maćkowiak and Wiederholt, 2009; 2015) misses this dynamic and would underestimate the deviation in output and their persistence.

Endogeneity of cost parameter Analogously to the isomorphism result of Section 3.5 there exists a mapping between shocks to data quality and shocks to the information processing cost ω. In numerical simulations detailed in the Appendix E reducing the signal to noise ratio by a half generates approximately the same impulse response of the Kalman gain as a $20 - 50\%$ increase in the cost parameter. Important differences remain however. For example, the shock to the variance of the irreducible noise is once and permanent. In contrast, the path for the cost shock is increasing and converges to its permanently higher value after a few periods. Overall, the correspondence between shocks stresses the intuitive notion that a reduction in the signal to noise ratio increases the effective costs of information processing.

This tight relationship further motivates the interpretation of imperfect data quality as semi-endogenising the marginal cost of attention which is typically treated as a free parameter and difficult to empirically discipline. Relating it to data quality may therefore make some progress in this regard. Further, the empirical evidence on the time-varying nature of data quality presented in Section 3.1 implies that assuming the information processing cost to be constant may not be an appropriate assumption for some applications.

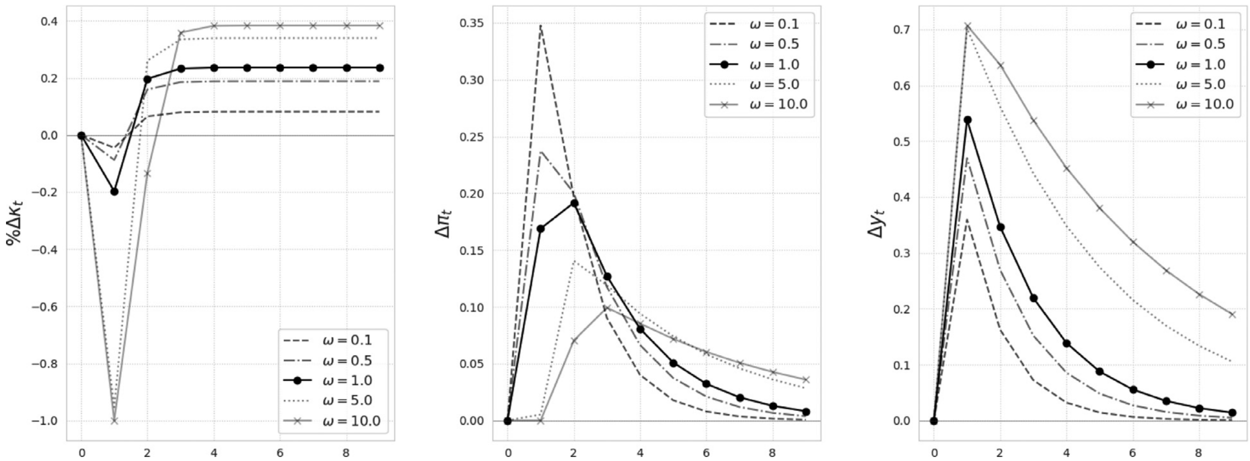

Fig. 9. Transition Path of Kalman Gain from a low to a high volatility regime. *Notes:* $\sigma_u^2/\phi^2 = \sigma_\varepsilon^2 = 0.25$ at $t = 0$ and $\sigma_u^2/\phi^2 = \sigma_\varepsilon^2 = 0.5$ for $t \geq 1$. *Left:* Response to the Kalman gain. *Centre:* Response to inflation *Right:* Response to the output gap.

4.4. Volatility regime shift

The empirical evidence above suggests that in times of an economic downturn both fundamental volatility rises and data quality falls. This concurrent rise in the two underlying variances is the subject of this section. I consider impulse responses of the economy along the transition path from a low to a high volatility regime in which volatilities rise by 100% while holding the signal to noise ratio constant.

The main result shows that there is a tendency for attention to fall in response to higher overall volatility in the short-run and rise over the long-run. In the absence of a rise in the signal to noise ratio current public signals are not more informative about the current fundamental and due to the increase in volatility less informative about the future. In fact, it can be shown that the measure of informativeness of the current public signal $\delta(X_t)$ falls even in the absence of a change in the signal to noise ratio.[18] Hence, a higher volatility regime raises the firms' reservation uncertainty and along the transition path they temporarily process less information until they reach the new steady state in which attentiveness is higher than initially.

Impulse responses shown in Fig. 9 highlight the complications arising from deteriorating data quality for the transmission of monetary policy. Even when the fundamental variance rises the concurrent deterioration in data quality flattens the Phillips curve, renders inflation and output fluctuations more persistent and reduces the on impact effect of monetary policy shocks.

It needs to be cautioned that the above results are sensitive to whether the signal to noise ratio rises and to the initial signal to noise ratio. The greater the improvement in the informativeness of public signals and the higher the initial signal to noise ratio the more likely it is for attention to rise in response to a volatility regime change. Naturally, as the role of the irreducible noise diminishes in the model impulse responses will converge to those of the standard model. Nonetheless, the exercise shows that the empirically observed increase in the variance of the unlearnable error acts to dampen the response in the Kalman gain and disrupt the transmission of monetary policy relative to the standard dynamic rational inattention model.

4.5. Anticipated shocks

In this Section, I consider the effect of releasing news about persistently lower volatilities of the underlying shocks from period $t + J$ onwards. The motivation to study the implications of anticipated volatility shocks is two-fold.

Central banks might announce the return to a hawkish monetary policy regime in the future as part of a forward guidance policy. In the model this is equivalent to an anticipated fall in the variance of the monetary policy shock. This exercise examines the importance of the timing of policy announcements and follows Reis (2011) in nature.

Likewise, the historical experience detailed in Section 3.1 indicates that a fall in data quality due to cyclical fluctuations will be transitory. It is thus important to study the case in which firms upon a deterioration in data quality anticipate to return to a good data regime in the future. Note that I am not imposing a deterministic process on the volatility but rather that firms form some concrete expectation of when data quality will improve. As a case in point, during the Covid crisis

[18] To see this note that we can write $\delta(X_t) = 1 - \left(\frac{\Theta(X_{t-1})}{\sigma_\varepsilon^2} + 1 + \frac{\sigma_u^2/\phi^2}{\sigma_\varepsilon^2}\right)^{-1}$ for some function $\Theta(X_{t-1})$. As $\Theta(X_{t-1})$ is pre-determined with respect to period t and thus exogenous with respect to a change in the irreducible noise variance a rise in total volatility for a given $(\sigma_u^2/\phi^2)/\sigma_\varepsilon^2$ lowers $\delta(X_t)$.

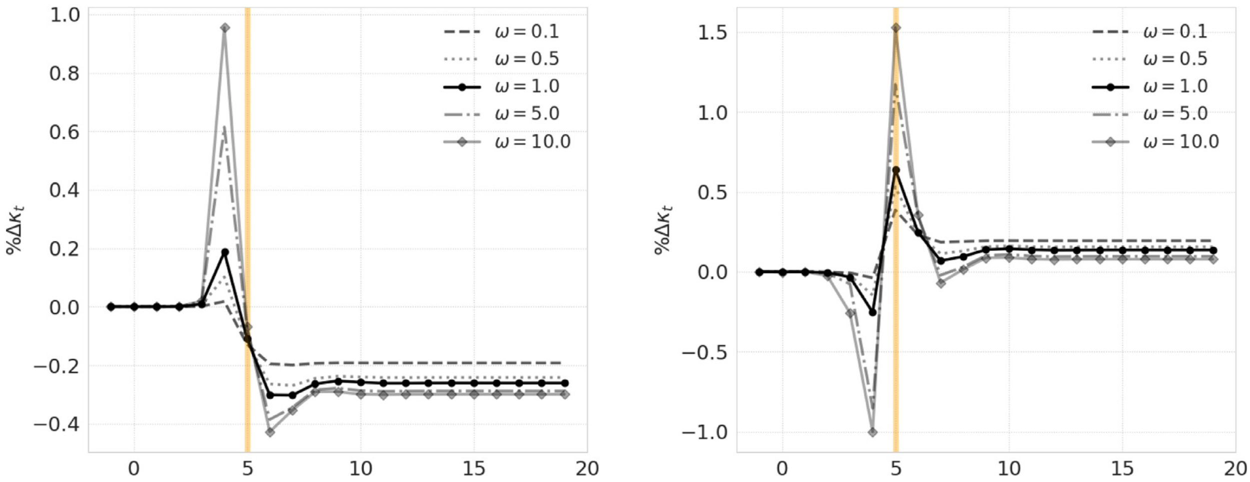

Fig. 10. Attention in response to anticipated volatility shock.*Panel A (Left):* $\frac{\sigma_u^2}{\phi^2} = 0.5$ for $j \leq J-1$ and $\frac{\sigma_u^2}{\phi^2} = 0.25$ for $j \geq J$. *Panel B (Right):* $\sigma_\varepsilon^2 = 0.5$ for $j \leq J-1$ and $\sigma_\varepsilon^2 = 0.25$ for $j \geq J$. Orange band indicates time when volatility shock realises.

lockdowns halted a large fraction of economic activity and induced substantial difficulties for data collection, which most probably depressed data quality. However, as most lockdowns were announced together with an end date there was the expectation that data quality would improve at a (known) future date.

Faced with an anticipated change in future volatility from $(\sigma_u^2/\phi^2, \sigma_\varepsilon^2)$ to $(\tilde{\sigma}_u^2/\tilde{\phi}^2, \tilde{\sigma}_\varepsilon^2)$ the firm the solves the following modified problem

$$\tilde{V}(X_t) = \max_{\{Y_{t+j}, X_{t+j+1}\}_{j=0}^{J-1}} -\sum_{j=0}^{J-1} \beta^j \left(\sigma_\varepsilon^2 \delta(X_{t+j}) + \delta(X_{t+j})^2 Y_{t+j} + \omega \log\left(\frac{X_{t+j}}{Y_{t+j}}\right)\right) + \beta^J V(X_{t+J}; \tilde{\sigma}_u^2/\tilde{\phi}^2, \tilde{\sigma}_\varepsilon^2)$$

subject to

$$X_{t+j+1} = \begin{cases} \sigma_\varepsilon^2 \delta(X_{t+j}) + \delta(X_{t+j})^2 Y_{t+j} + \frac{\sigma_u^2}{\phi^2} + \sigma_\varepsilon^2 & \text{if } j < J-1 \\ \sigma_\varepsilon^2 \delta(X_{t+j}) + \delta(X_{t+j})^2 Y_{t+j} + \frac{\tilde{\sigma}_u^2}{\tilde{\phi}^2} + \tilde{\sigma}_\varepsilon^2 & \text{if } j = J-1 \end{cases}$$

and the sequence of corresponding no forgetting constraints.

The problem is solved in two steps. First, the firm computes their value function $V(X'; \tilde{\sigma}_u^2/\tilde{\phi}^2, \tilde{\sigma}_\varepsilon^2)$ for any future X' and volatility pair $(\tilde{\sigma}_u^2/\tilde{\phi}^2, \tilde{\sigma}_\varepsilon^2)$. Given their initial prior uncertainty X_t the firm chooses the optimal sequence of $\{Y_{t+j}, X_{t+j+1}\}_{j=0}^{J-1}$ as to maximise today's flow payoff and its associated continuation value at a possibly higher or lower level of volatility.

I consider the following simulations in Fig. 10. Let $J = 5$. Panel A. plots the impulse response for the Kalman gain when the fundamental variance decreases by 50% whilst Panel B. plots the impulse response for the Kalman gain when the variance of the unlearnable noise falls by 50%. The shifts in variances are assumed to be made known to firms in period t, that is J periods prior to their realisation.

4.5.1. Forward guidance about return to Hawkish monetary regime

Forward guidance about a hawkish monetary policy has three main effects as shown in Panel A. of Fig. 10. First, the response in attention is concentrated on the short period directly prior to the realisation of the policy change. Hence, the timing of the announcement is not of substantive importance as long as it occurs in advance of the regime switch. However, model implications are materially different for an anticipated compared to an unanticipated volatility shock. In the former firms increase their information processing temporarily to pre-empt the anticipated fall in the informativeness of data. In the latter this pre-emption effect is absent and firms to tend to lower their attention levels immediately. In the long-run attention levels converge to the same steady state regardless of whether the policy change is anticipated or unannounced.

The mechanism at work is the converse of the argument outlined in Lemma 2 that intertemporal effects act to increase the firm's posterior uncertainty in response to a rise in the monetary policy shock variance. As volatilities are fixed for J periods until the volatility shock materialises the contemporaneous channel of the fall in the signal to noise ratio is shut off. At the same time, the reduction in future volatility implies that agents anticipate their prior knowledge to be more informative about the future. Hence, firms utilise the high signal to noise ratio prior to the realisation of the policy shift and acquire more information to attain a more precise prior. Once the policy regime change is completed the fall in the signal to noise ratio dominates and agents reduce their attention.

Combining the results from Section 4.1 and Panel A. highlights that a temporary switch to a dovish monetary policy regime when accompanied by a clear communication strategy can in fact increase attention by firms in the short-run, temporarily steepen the Phillips curve and enhance the transmission of monetary policy.

4.5.2. News about a rise in data quality

News about greater data quality in the future induces the opposite effect. Firms find it optimal to take a *wait-and-see* approach in which they reduce their attention until the improvement in data quality materialises. Intuitively, anticipating an increase in the signal to noise ratio agents optimally substitute attention from the low data quality periods to the high data quality period. Hence, when a rise in the irreducible noise variance as in Section 4.3 is accompanied by the anticipation of a future improvement in data quality the effects of a fall in attention are amplified. As a result, a transitory fall in data quality magnifies the disruptions to the transmission of monetary policy as it depresses attention levels by firms through two compounding channels.

5. Conclusion

I characterise and solve a variant of the traditional Linear Quadratic Gaussian rational inattention model that captures imperfect data quality. I show that a careful distinction between fundamental uncertainty and data uncertainty is of substantive importance when studying the dynamics of attention in a world of time-varying volatility. The presence of noisy data interacts with the fundamental's variance in non-trivial ways. In the dynamic extension of the model these interaction effects have important qualitative implications. They determine whether attention rises or falls in response to a more volatile fundamental. Volatility shocks thus cannot in general be considered in isolation of the initial information quality regime in the economy.

I show how higher data quality, be it from improvements in data collection or from more advanced central bank communication, raises the economy's attention levels, strengthens the transmission of monetary policy and reduces fluctuations in the macroeconomy.

On the issue of time-varying Phillips curves I propose an alternative mechanism that can disrupt the relationship between output and inflation during recessions. A deterioration in information processing by firms could render macroeconomic variables more persistent and the recovery sluggish, relative to the predictions of a standard rational inattention model. What is more, insofar as information quality correlates with economic activity those dynamics may be a general feature of the business cycle.

Appendix A. Model derivations

Ad Section 2: Solution to The Static Attention Problem

The maximisation problem for the RI problem with irreducible noise reads

$$\max_{\mathbb{V}ar(x|s)} -\mathbb{V}ar(\theta|x) - \left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^2 \mathbb{V}ar(x|s) - \omega \log\left(\frac{\mathbb{V}ar(x)}{\mathbb{V}ar(x|s)}\right) \tag{A.1}$$

with the first order condition for an interior solution

$$\left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^2 = \omega \frac{1}{\mathbb{V}ar(x|s)} \tag{A.2}$$

Given that $\kappa = \max\left\{\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\left(1 - \frac{\mathbb{V}ar(x|s)}{\mathbb{V}ar(x)}\right), 0\right\}$,[19] we arrive at the solution stated in the main text.

From the solution to the first order condition and the property of Gaussian variables it follows that

$$\mathbb{V}ar(x|s) = \frac{(\sigma_\theta^2 + \sigma_\varepsilon^2)\sigma_\zeta^2}{\sigma_\theta^2 + \sigma_\varepsilon^2 + \sigma_\zeta^2} = \omega \frac{(\sigma_\theta^2 + \sigma_\varepsilon^2)^2}{\sigma_\theta^4} \tag{A.3}$$

and thus

$$\sigma_\zeta^2 = \max\left\{\frac{\omega(\sigma_\theta^2 + \sigma_\varepsilon^2)^2}{\sigma_\theta^4 - \omega(\sigma_\theta^2 + \sigma_\varepsilon^2)}, 0\right\} \tag{A.4}$$

Comparative statics are given by

$$\frac{d\sigma_\zeta^2}{d\sigma_\theta^2} = -\frac{\omega(\sigma_\theta^2 + \sigma_\varepsilon^2)^2}{[\sigma_\theta^4 - \omega(\sigma_\theta^2 + \sigma_\varepsilon^2)]^2}\left(2\frac{\sigma_\theta^2\sigma_\varepsilon^2}{\sigma_\theta^2 + \sigma_\varepsilon^2} + \omega\right) < 0 \tag{A.5}$$

[19] Note that the Kalman gain is given by $\kappa = \frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(s)}$ and $\mathbb{V}ar(s) = \frac{\mathbb{V}ar^2(x)}{\mathbb{V}ar(x) - \mathbb{V}ar(x|s)}$.

$$\frac{d\sigma_\zeta^2}{d\sigma_\varepsilon^2}=\frac{\omega(\sigma_\theta^2+\sigma_\varepsilon^2)}{[\sigma_\theta^4-\omega(\sigma_\theta^2+\sigma_\varepsilon^2)]^2}\left[2\sigma_\theta^4-\omega(\sigma_\theta^2+\sigma_\varepsilon^2)\right]\geq 0 \tag{A.6}$$

where the non-negativity of the latter derivative is given by the condition for an interior solution.

Ad Section 3.3: Dynamic Model

Objective Function Derivation. Since $q_t \to x_t \to s_t$ forms a Markov chain $f(q_t|x_t,s^{t-1},s_t)=f(q_t|x_t,s^{t-1})$ (see proof of Proposition 1). We can use that fact, the Total Law of Variance and the definition $x_t=q_t+\varepsilon_t$ to write

$$\mathbb{V}ar(q_t|s^t)=\mathbb{E}\Big[\mathbb{V}ar(q_t|x_t,s^t)|s^t\Big]+\mathbb{V}ar\Big[\mathbb{E}[q_t|x_t,s^t]|s^t\Big] \tag{A.7}$$

$$=\mathbb{E}\Big[\mathbb{V}ar(\varepsilon_t|x_t,s^{t-1})|s^t\Big]+\mathbb{V}ar\Big(x_t-\mathbb{E}[\varepsilon_t|x_t,s^{t-1}]|s^t\Big) \tag{A.8}$$

Further, the Gaussianity assumption implies that the joint, marginal and conditional distributions of $\{q_t,x_t,s^t\}$ are Gaussian as well. Then by the properties of the Gaussian distribution and re-expressing the result in neater notation used in the main text yields

$$\mathbb{V}ar(q_t|s^t)=\mathbb{V}ar(\varepsilon_t|x_t,s^{t-1})+\left[1-\frac{\mathbb{V}ar(\varepsilon_t)}{\mathbb{V}ar(x_t|s^{t-1})}\right]^2\mathbb{V}ar\Big(x_t|s^t\Big) \tag{A.9}$$

$$=\mathbb{V}ar(\varepsilon_t)\left[1-\frac{\mathbb{V}ar(\varepsilon_t)}{\mathbb{V}ar(x_t|s^{t-1})}\right]+\left[1-\frac{\mathbb{V}ar(\varepsilon_t)}{\mathbb{V}ar(x_t|s^{t-1})}\right]^2\mathbb{V}ar\Big(x_t|s^t\Big) \tag{A.10}$$

$$=\delta(X_t)\sigma_\varepsilon^2+\delta(X_t)^2Y_t \tag{A.11}$$

First Order Conditions The first order conditions of the dynamic programming problem of (8) yields

$$0=-(\lambda-1)\delta(X)^2+\omega\frac{1}{Y}+\beta\delta(X)^2V'(X')-\mu \tag{A.12}$$

where μ denotes the Lagrange multiplier on the no-forgetting constraint $Y\leq X$. Note that the non-negativity constraint will never bind since $\lim_{Y\to 0}RHS=\infty$.

By the Benveniste-Scheinkman condition,

$$V'(X)=-(\lambda-1)\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X}-\omega\frac{1}{X}+\beta V'(X')\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X} \tag{A.13}$$

$$=-(\lambda-1)\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X}-\omega\frac{1}{X}+\left[(\lambda-1)-\omega\frac{1}{\delta(X)^2Y}+\frac{\mu}{\delta(X)^2}\right]\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X}+\mu \tag{A.14}$$

$$=-\omega\frac{1}{X}-\omega\frac{1}{\delta(X)^2Y}\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X}+\frac{\mu}{\delta(X)^2} \tag{A.15}$$

which at an interior solution $\mu=0$ reduces to[20]

$$V'(X)=-\omega\left(\frac{1}{X}+\frac{1}{\delta(X)^2Y}\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X}\right) \tag{A.16}$$

where the partial derivative expressed at time t is equivalent to $\frac{\partial\mathbb{V}ar(q_t|s^t)}{\partial X_t}$, holding constant Y_t. Iterating the last expression forward and substituting into the FOC yields the expression for an interior solution in the main text.[21]

The explicit expression for the partial derivative follows straightforward from noting that by the Envelope Theorem the derivative is taken, holding Y_t fixed. It is easy to verify that the incremental change in the continuation value at an interior solution is parametrised by ω.

[20] The last line follows from the fact that any function multiplied by μ will only enter evaluated at the constrained value $Y=X$ and $1+\frac{1}{\delta(X)^2}\frac{\partial\left(\sigma_\varepsilon^2\delta(X)+\delta(X)^2Y\right)}{\partial X}_{|Y=X}=\frac{1}{\delta(X)^2}$

Limiting Cases Let $\beta \to 0$, then the interior solution to the first order condition is given by

$$Y_t = \frac{\omega}{(\lambda - 1)\delta(X_t)^2} \tag{A.17}$$

as in the static case.

Let $\sigma_\varepsilon^2 \to 0$, then the first order condition converges to the problem in Afrouzi and Yang (2021)

$$0 = -(\lambda - 1) + \omega \frac{1}{Y_t} - \beta\omega \frac{1}{X_{t+1}} \tag{A.18}$$

Solution to Kalman Gain The Kalman gain for an interior solution is

$$\kappa_t = 1 - \frac{\mathbb{V}ar(q_t|s^t)}{\mathbb{V}ar(q_t|s^{t-1})} \tag{A.19}$$

$$= 1 - \frac{\delta(X_t)^2 Y_t + \delta(X_t)\sigma_\varepsilon^2}{X_t - \sigma_\varepsilon^2} \tag{A.20}$$

$$= 1 - \frac{\delta(X_t)Y_t + \sigma_\varepsilon^2}{X_t} = \delta(X_t)\left(1 - \frac{Y_t}{X_t}\right) \tag{A.21}$$

where I have used the Law of Total Variance and the fact that $\delta(X_t) \equiv \frac{X_t - \sigma_\varepsilon^2}{X_t}$.

Aggregation Aggregation is adapted from Afrouzi and Yang (2021). Suppose the solution is interior and information processing positive. Then, we can aggregate prices as

$$p_t = \int_0^1 p_{i,t-1} + \kappa_t(q_t + \varepsilon_t + \zeta_{i,t} - p_{i,t-1})di \tag{A.22}$$

$$= p_{t-1} + \kappa_t(q_t + \varepsilon_t - p_{t-1}) \tag{A.23}$$

Using the identity that $q_t = y_t + p_t$ and rearranging yields the Phillips curve

$$\pi_t = \frac{\kappa_t}{1 - \kappa_t}\left(y_t + \varepsilon_t\right) \tag{A.24}$$

for an interior solution.

To get the relation for the persistence in the economy note that using the Phillips curve relation

$$\Delta y_t = \Delta\left[(\kappa_t^{-1} - 1)\pi_t - \varepsilon_t\right] \tag{A.25}$$

$$= (\kappa_t^{-1} - 1)\pi_t - (\kappa_{t-1}^{-1} - 1)\pi_{t-1} - \Delta\varepsilon_t \tag{A.26}$$

Then given $\Delta q_t = \frac{\sigma_u}{\phi}u_t = \Delta y_t + \pi_t$,

$$\kappa_t^{-1}\pi_t - (\kappa_{t-1}^{-1} - 1)\pi_{t-1} - \Delta\varepsilon_t = \frac{\sigma_u}{\phi}u_t \tag{A.27}$$

Rearranging yields the result in the text.

To get the dynamic relation for the output gap note that

$$y_t = \left(\kappa_t^{-1} - 1\right)\pi_t - \varepsilon_t \tag{A.28}$$

[21] Since the constraint is occasionally binding for low realisations of the state variable it is instructive to consider the cases in which $\mu > 0$ and/or $\mu' > 0$. The FOC is then given by

$$0 = -(\lambda - 1)\delta(X)^2 + \omega\frac{1}{Y} - \beta\omega\delta(X)^2\left(\frac{1}{X'} + \frac{1}{\delta(X')^2 Y'}\frac{\partial\left(\sigma_\varepsilon^2\delta(X') + \delta(X')^2 Y'\right)}{\partial X'}\right) + \beta\frac{\delta(X)^2}{\delta(X')^2}\mu' - \mu$$

It can be shown that given that the system will eventually have an interior solution after a finite number of periods it holds that $\frac{\mu}{\delta(X)^2} - \beta\frac{\mu'}{\delta(X')^2} \geq 0$, which holds trivially if $\mu' = 0$. Hence, the optimality condition can be expressed more succinctly as

$$-(\lambda - 1)\delta(X)^2 + \omega\frac{1}{Y} - \beta\omega\delta(X)^2\left(\frac{1}{X'} + \frac{1}{\delta(X')^2 Y'}\frac{\partial\left(\sigma_\varepsilon^2\delta(X') + \delta(X')^2 Y'\right)}{\partial X'}\right) \geq 0$$

$$= \left(\kappa_t^{-1} - 1\right)\left(\frac{\sigma_u}{\phi} u_t - \Delta y_t\right) - \varepsilon_t \tag{A.29}$$

Rearranging yields the result in the main text.

Appendix B. Omitted proofs

Proof of Proposition 1

Proof. As the variables are defined, we can form the following Markov Chain $\theta \to x \to s$. In particular, $f(s|x,\theta) = f(s|x)$. Thus,

$$f(\theta|x,s) = \frac{f(s|x,\theta)f(x|\theta)f(\theta)}{f(s|x)f(x)} \tag{B.1}$$

$$= \frac{f(s|x)f(x|\theta)f(\theta)}{f(s|x)f(x)} = f(\theta|x) \tag{B.2}$$

By the Law of Total Variances and since $f(\theta|x,s) = f(\theta|x)$

$$\mathbb{V}ar(\theta|s) = \mathbb{E}[\mathbb{V}ar(\theta|x,s)|s] + \mathbb{V}ar(\mathbb{E}[\theta|x,s]|s) \tag{B.3}$$

$$= \mathbb{V}ar(\theta|x) + \left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^2 \mathbb{V}ar(x|s) \tag{B.4}$$

where the final line follows from the Gaussian assumption. Thus, the loss function can be re-expressed isomorphic to the traditional RI problem. Standard results apply and Gaussianity is optimal in this setting. To see this,

$$f^*(x,s) = \underset{f(x,s)}{argmax} - \mathbb{V}ar(\theta|x) - \left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^2 \mathbb{V}ar(x|s) - 2\omega I(x;s) \tag{B.5}$$

$$= \underset{f(x,s)}{\arg\max} -\mathbb{V}ar(x|s) - 2\omega\left[\frac{\mathbb{V}ar(\theta)}{\mathbb{V}ar(x)}\right]^{-2}\left(H(x) - H(x|s)\right) \tag{B.6}$$

where the second line follows from the fact that $\mathbb{V}ar(\theta|x)$ is exogenous from the point of view of the agent and the definition of mutual information as the difference of the entropy of the prior and the posterior distribution. The optimality of the Gaussian distribution follows from the property of maximum entropy for the Gaussian distribution for a given variance $\mathbb{V}ar(x|s)$ (see Cover and Thomas 2012, Chapter 12). By duality minimising a variance subject to an entropy constraint is equivalent to maximising entropy subject to a variance (and mean) constraint. □

Proof of Proposition 2.

Proof. It is straightforward to show that

$$\frac{d\kappa}{d\sigma_\theta^2} = \frac{\sigma_\varepsilon^2}{(\sigma_\theta^2 + \sigma_\varepsilon^2)^2} + \omega\frac{1}{\sigma_\theta^4} > 0 \tag{B.7}$$

and

$$\frac{d\kappa}{d\sigma_\varepsilon^2} = -\frac{\sigma_\theta^2}{(\sigma_\theta^2 + \sigma_\varepsilon^2)^2} < 0 \tag{B.8}$$

□

Proof of Proposition 3

Proof. This proof closely follows the proof of Lemma 3 in Afrouzi and Yang (2021). To obtain a contradiction, suppose that the optimal sequence of signals s^t is not Gaussian. Then we can construct a Gaussian process that yields the same sequence of payoffs but costs less in terms of the entropy cost. To see this, let us first rewrite the payoff function.

Since the random vectors of histories $q^t \to x^t \to s^t$ form a Markov chain $p(q^t|x^t,s^t) = p(q^t|x^t)$. Then, by the Total Law of Variance,

$$\mathbb{V}ar(q^t|s^t) = \mathbb{E}\left[\mathbb{V}ar(q^t|x^t,s^t)|s^t\right] + \mathbb{V}ar\left(\mathbb{E}[q^t|x^t,s^t]|s^t\right) \tag{B.9}$$

$$= \mathbb{E}\left[\mathbb{V}ar(q^t|x^t)|s^t\right] + \mathbb{V}ar\left(\mathbb{E}[q^t|x^t]|s^t\right) \tag{B.10}$$

Given that q^t and x^t are jointly Gaussian, then

$$\mathbb{V}ar(q^t|s^t) = \mathbb{V}ar(q^t|x^t) + \mathbb{V}ar\Big(\mathbb{C}ov(q^t, x^t)\mathbb{V}ar(x^t)^{-1}x^t|s^t\Big) \tag{B.11}$$

$$= \mathbb{V}ar(q^t|x^t) + \mathbb{C}ov(q^t, x^t)\mathbb{V}ar(x^t)^{-1}\mathbb{V}ar\Big(x^t|s^t\Big)\Big(\mathbb{C}ov(q^t, x^t)\mathbb{V}ar(x^t)^{-1}\Big)' \tag{B.12}$$

Let us now define a sequence of Gaussian variables $\{\hat{s}_t\}_{t\geq 0}$ such that

$$\mathbb{V}ar(x^t|\hat{s}^t) = \mathbb{E}[\mathbb{V}ar(x^t|s^t)|s^{-1}]$$

Using the result in Eq. (B.12) and the fact that the covariance matrix of jointly Gaussian variables is non-stochastic we see that both sequences of signals induce the same payoff stream by construction. The rest of the proof is a straight forward adaptation of the proof in Afrouzi and Yang (2021) and relies on the maximum entropy property of the Gaussian distribution among the random variables with the same expected covariance matrix. □

Proof Lemma 1

Proof.

1. Note that eventually the solution to the system will be interior and once interior the dynamics will be such that all future solutions remain interior. It thus suffices to investigate the dynamics for an interior solution. Denote by $g(X_t)$ next period's prior variance at the optimal policy and by $h(X_t)$ the optimal policy function for the posterior variance. Then, by the Law of Motion

$$g(X_t) = \delta(X_t)^2 h(X_t) + \delta(X_t)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \frac{\sigma_u^2}{\phi^2} \tag{B.13}$$

By the First Order Condition

$$\frac{\omega}{g(X_t) - [\delta(X_t)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \frac{\sigma_u^2}{\phi^2}]} = 1 - \beta V'(g(X_t)) \tag{B.14}$$

By the Implicit Function Theorem

$$\frac{\omega}{[g(X_t) - [\delta(X_t)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \frac{\sigma_u^2}{\phi^2}]]^2}\Big(g'(X_t) + \delta'(X_t)\sigma_\varepsilon^2\Big) = \beta V''(g(X_t))g'(X_t) \tag{B.15}$$

and rearrange and denote $g(X_t) - [\delta(X_t)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \frac{\sigma_u^2}{\phi^2}] = A(X_t)$

$$\left(\frac{\omega}{[A(X_t)]^2} - \beta V''(g(X_t))\right)g'(X_t) = \frac{\omega}{[A(X_t)]^2}\delta'(X_t)\sigma_\varepsilon^2 \tag{B.16}$$

As is shown in Appendix F at any local maximum $\frac{\omega}{[A(X_t)]^2} > \beta V''(g(X_t))$. Hence, $g'(X_t) > 0$. Further, supposing $\beta V''(g(X_t)) < \left[1 - \left(\frac{\sigma_\varepsilon^2}{X_t}\right)^2\right]\frac{\omega}{[A(X_t)]^2}$ we can verify that $g'(X_t) < 1$. It follows that $|X_{t+1} - X_\infty| = |g(X_t) - g(X_\infty)| = |\int_{X_\infty}^{X_t} g'(t)dt| < |X_t - X_\infty|$ since the derivative is bounded between zero and one. The Law of Motion is thus a contraction and the system converges to a unique steady state.
2. The last point follows from the properties of the stochastic process of nominal aggregate demand as a random walk. As its variance grows linearly with time, the state variable would increase without bounds with a Kalman gain perpetually at zero; i.e. $\lim_{\tau\to\infty} X_\tau = \infty$. From (9) it is clear that as the state variable approaches infinity the reservation utility cannot. To see this, suppose that $\bar{Y}_\tau > X_\tau$ for all $\tau \geq t$ while $\lim_{\tau\to\infty} X_\tau = \infty$. Then the RHS approaches $-(\lambda - 1) < 0$ and the firm would reduce their posterior uncertainty. A contradiction.To derive the expression for the steady state Kalman gain note that by the law of motion for $X_t = X_{t+1} = X_\infty$

$$X_\infty = \delta(X_\infty)^2 Y_\infty + \delta(X_\infty)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \frac{\sigma_u^2}{\phi^2} \tag{B.17}$$

which gives, noting that $X_\infty - \delta(X_\infty)\sigma_\varepsilon^2 - \sigma_\varepsilon^2 = \frac{X_\infty^2 - 2X_\infty\sigma_\varepsilon^2 + \sigma_\varepsilon^4}{X_\infty} = \delta(X_\infty)^2 X_\infty$,

$$Y_\infty = X_\infty - \frac{1}{\delta(X_\infty)^2}\frac{\sigma_u^2}{\phi^2} > 0 \tag{B.18}$$

Then,

$$\kappa_\infty = \delta(X)\left(1 - \frac{X_\infty - \frac{1}{\delta(X_\infty)^2}\frac{\sigma_u^2}{\phi^2}}{X_\infty}\right) = \frac{\sigma_u^2/\phi^2}{X_\infty - \sigma_\varepsilon^2}$$

as was to be shown.

□

Proof Corollary 2

Proof. The unconditional variance can be derived using the law of motion for inflation and output at the steady state Kalman gain.

$$\begin{aligned}\mathbb{V}ar(\pi_t) &= (1-\kappa)^2\mathbb{V}ar(\pi_{t-1}) + \kappa^2\left(\frac{\sigma_u^2}{\phi^2} + 2\sigma_\varepsilon^2\right) + 2\mathbb{C}ov\Big((1-\kappa)\pi_{t-1}, \kappa\left(\frac{\sigma_u}{\phi}u_t + \Delta\varepsilon_t\right)\Big)\\ &= (1-\kappa)^2\mathbb{V}ar(\pi_{t-1}) + \kappa^2\left(\frac{\sigma_u^2}{\phi^2} + 2\sigma_\varepsilon^2\right) - 2(1-\kappa)\kappa^2\sigma_\varepsilon^2\\ &= \frac{\kappa}{2-\kappa}\left(\frac{\sigma_u^2}{\phi^2} + 2\kappa\sigma_\varepsilon^2\right)\end{aligned}$$

and for output

$$\begin{aligned}\mathbb{V}ar(y_t) &= (1-\kappa)^2\mathbb{V}ar(y_{t-1}) + (1-\kappa)^2\frac{\sigma_u^2}{\phi^2} + \kappa^2\sigma_\varepsilon^2\\ &= \frac{(1-\kappa)^2}{1-(1-\kappa)^2}\frac{\sigma_u^2}{\phi^2} + \frac{\kappa}{2-\kappa}\sigma_\varepsilon^2\end{aligned}$$

□

Proof Lemma 2

Proof. Rewrite the optimality relation in (9) as

$$0 = -(\lambda - 1) + \omega\frac{1}{\delta(X_t)^2\bar{Y}_t} + \beta V'(X_{t+1}) \tag{B.19}$$

The contemporaneous effect of an increase in the monetary policy shock variance operates through increasing the prior variance X_t in $\delta(X_t)$. In particular, $\frac{d}{d\sigma_u^2/\phi^2} = \frac{d\delta(X_t)}{dX_t}\frac{dX_t}{d\sigma_u^2/\phi^2} = \frac{\sigma_\varepsilon^2}{X_t^2} > 0$. All else equal it is straightforward to see that this lowers the posterior variance $\bar{Y}_t$.

To compute the intertemporal effect of an increase in the monetary policy shock variance we need to compute how $V'(X_{t+1})$ changes with σ_u^2/ϕ^2. To this effect, let us expand the state space to include the monetary policy shock variance. Note that since we assume that the agent treats the underlying variances as fixed over time we obtain a deterministic law of motion for the state variable where $z_{t+1} = z_t$ where $z \equiv \sigma_u^2/\phi^2$. Hence, the Bellman equation is given by

$$V\big(X;\sigma_u^2/\phi^2\big) = \max_Y -(\lambda-1)\big(\sigma_\varepsilon^2\delta(X) + \delta(X)^2Y\big) - \omega\log\left(\tfrac{X}{Y}\right) + \beta V\left(\delta(X)^2Y + \sigma_\varepsilon^2\delta(X) + \tfrac{\sigma_u^2}{\phi^2} + \sigma_\varepsilon^2;\sigma_u^2/\phi^2\right) \quad \text{subject to } 0 \le Y \le X \tag{B.20}$$

We will next compute the cross-derivative $\frac{d}{d\sigma_u^2/\phi^2}\left(\frac{\partial}{\partial X}V(X',\sigma_u^2/\phi^2)\right)$ evaluated at the steady state. To be precise we will compute the total derivative with respect to σ_u^2/ϕ^2 of the partial derivative with respect to the first argument of the value function evaluated at X'. It is important to note that $X' = \delta(X)^2Y + \sigma_\varepsilon^2\delta(X) + \frac{\sigma_u^2}{\phi^2} + \sigma_\varepsilon^2$ such that $\frac{\partial X'}{\partial\sigma_u^2/\phi^2} = 1$ for a fixed Y and X. Likewise, $\frac{\partial X}{\partial\sigma_u^2/\phi^2} = 1$. Then,

$$\frac{d}{d\sigma_u^2/\phi^2}\left(\frac{\partial}{\partial X}V(X',\sigma_u^2/\phi^2)\right) = \frac{\partial^2}{\partial X\partial\sigma_u^2/\phi^2}V(X';\sigma_u^2/\phi^2) + \frac{\partial^2}{\partial X^2}V(X';\sigma_u^2/\phi^2)\frac{dX'}{d\sigma_u^2/\phi^2} \tag{B.21}$$

where we note that the second term is positive since the value function is convex and $\frac{dX'}{d\sigma_u^2/\phi^2} > 0$. To make progress at computing the first term, we apply the Envelope condition

$$\frac{\partial}{\partial\sigma_u^2/\phi^2}V(X;\sigma_u^2/\phi^2) = \beta\frac{\partial}{\partial X}V(X';\sigma_u^2/\phi^2)\frac{\partial X'}{\partial\sigma_u^2/\phi^2} + \beta\frac{\partial}{\partial\sigma_u^2/\phi^2}V(X';\sigma_u^2/\phi^2) \tag{B.22}$$

$$= \beta \frac{\partial}{\partial X} V(X'; \sigma_u^2/\phi^2) + \beta \frac{\partial}{\partial \sigma_u^2/\phi^2} V(X'; \sigma_u^2/\phi^2) \tag{B.23}$$

Then,

$$\frac{\partial^2}{\partial X \partial \sigma_u^2/\phi^2} V(X; \sigma_u^2/\phi^2) = \beta \frac{\partial^2}{\partial X^2} V(X'; \sigma_u^2/\phi^2) \frac{dX'}{dX} + \beta \frac{\partial^2}{\partial X \partial \sigma_u^2/\phi^2} V(X'; \sigma_u^2/\phi^2) \underbrace{\frac{dX'}{dX}}_{>0;<1} \tag{B.24}$$

where $0 < \frac{dX'}{dX} < 1$ by Lemma 1.

Evaluating this expression at the steady state $X' = X = X^*$ and rearranging yields

$$\frac{\partial^2}{\partial X \partial \sigma_u^2/\phi^2} V(X^*; \sigma_u^2/\phi^2) = \frac{\beta \frac{dX'}{dX}_{|X'=X^*}}{1 - \beta \frac{dX'}{dX}_{|X'=X^*}} \frac{\partial^2}{\partial X^2} V(X^*; \sigma_u^2/\phi^2) > 0 \tag{B.25}$$

which is positive since the value function has been found to be convex in X and Lemma (1) showed that $0 < \frac{dX'}{dX} < 1$. Hence, we see from (B.19) that all else equal an increase in the monetary policy shock variance increases the posterior variance $\bar{Y}_t$ in a neighbourhood around the steady state.

In the special case in which $\sigma_\varepsilon^2 \to 0$ the contemporaneous effect is zero and thus the intertemporal effect dominates, inducing an instantaneous increase in the posterior variance. It can be shown that as $\sigma_\varepsilon^2 \to 0$

$$\frac{dY_t}{d\sigma_u^2/\phi^2} = \beta \frac{Y_t^2}{X_{t+1}^2 - \beta Y_t^2} > 0 \tag{B.26}$$

It follows that the presence of irreducible noise is a necessary condition for the posterior variance to fall when the monetary policy shock variance rises. Further, as $\omega \to \infty$ the relative weighting of the increase in the contemporaneous marginal benefit in the firm's decision making goes to zero, all else equal. Thus, if the posterior variance falls it must be in a setting where σ_ε^2 is sufficiently large and ω sufficiently small. These identify necessary but not sufficient conditions. □

Appendix C. Dynamic vs static model

Fig. C.1 plots the absolute value of the slope of the value function at the respective steady states for various levels of $\sigma_u^2/\phi^2, \sigma_\varepsilon^2$ and ω. The top panel plots $\frac{\bar{Y}_t^{stat}}{\bar{Y}_t^{dyn}}$. The bottom panel plots $|V'(X)|$. In the simulations the posterior variance in the static model is up to twice as large as in the dynamic model. The bottom panels show that the intertemporal forces are even stronger in fact, implying that the biggest difference between models is the fact that agents process information in the dynamic model in circumstances under which they would ignore all new data in the static counterpart.

Appendix D. Implications for macroeconomic fluctuations

Fig. D.1 plots the countour plots for the unconditional variances of inflation and output for different signal to noise ratios and marginal costs of attention. They show the main result from Section 3.5 that the unconditional variances are increasing in σ_ε^2 indicating that poorer data quality implies larger macroeconomic fluctuations.

Appendix E. Equivalent cost shock series

Fig. E.1 shows the series of shocks to the marginal cost of attention ω that replicates the impulse response of the Kalman gain to a 100% increase in the variance of the irreducible noise σ_ε^2 as computed in Fig. 8.

Appendix F. Value function

Note that $J(X_0, \mathbb{Y})$ as defined in Equation6 is well defined and continuous and is bounded above by zero. The constraint set and the law of motion for the prior variance is clearly non-empty. Assumptions 4.1 and 4.2 in Stokey et al. (1989) are therefore satisfied. Bellman's principle of optimality then shows that $V(X) = \sup_{\mathbb{Y}} J(X, \mathbb{Y})$. Since $J(X_0, \mathbb{Y})$ is continuous and the constraint set well-defined the Theorem of the Maximum states that $V(X)$ will be continuous also. The fact that $V(X)$ exists and is unique follows from the fact that $V(X)$ is the fixed point of a contraction mapping of $T: C(X) \to C(X)$ where $C(X)$ denotes the space of bounded continuous functions $f: X \to \mathbb{R}^-$, for $\beta \in (0, 1)$. The proof that Blackwell's sufficiency conditions are satisfied is standard.

Without further assumptions it is, however, not possible to characterise the policy function more sharply. In particular, the usual assumptions of the flow utility as strictly concave and strictly increasing is not satisfied. In fact, $V(X)$ is convex for most of the parameter space and decreasing. This follows naturally from the specification of the cost function as entropy-based but complicates the problem as monotonicity and uniqueness of the maximising policy correspondence may no longer

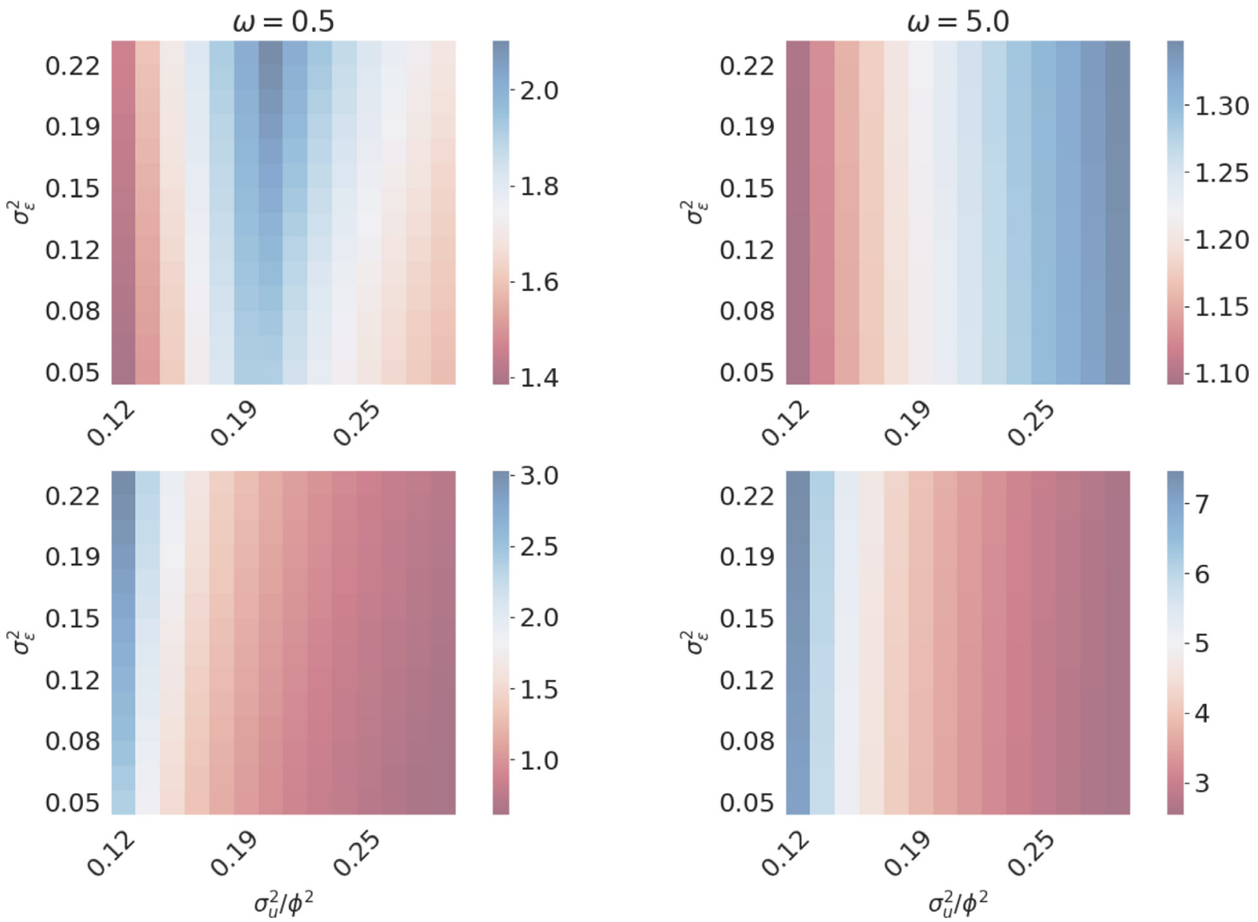

Fig. C.1. 2×2 grid. Top panels show relative posterior variances at the steady state. Bottom panels shows the slope of the value function at the steady state. σ_u^2/ϕ^2 on the x-axis. σ_ε^2 on the y-axis. Left panels $\omega = 0.5$. Right panels $\omega = 5$.

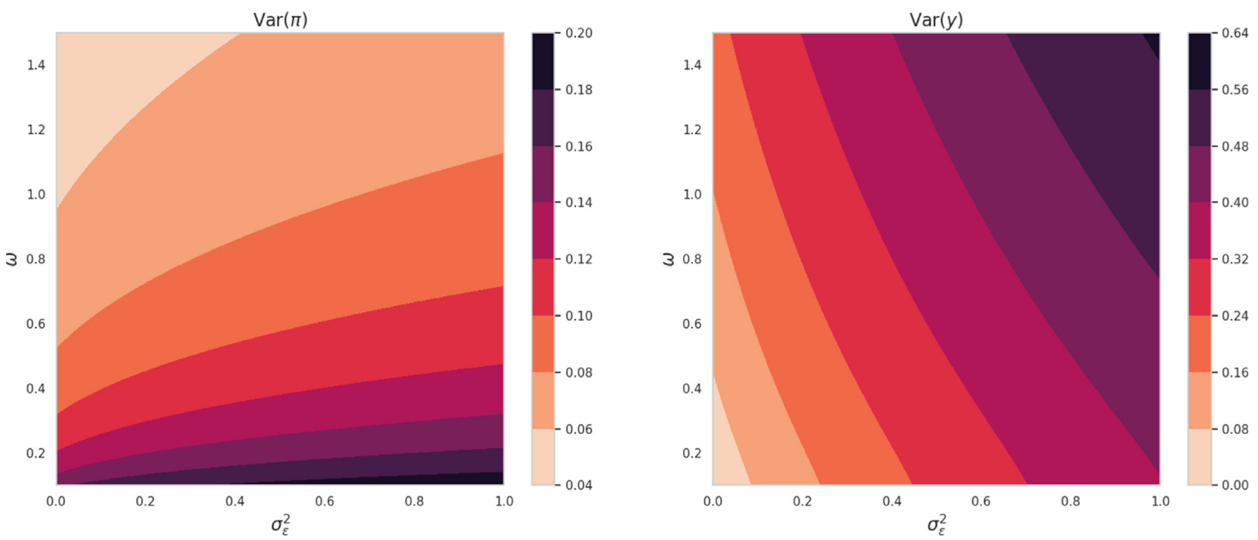

Fig. D.1. Unconditional variances of inflation and output.

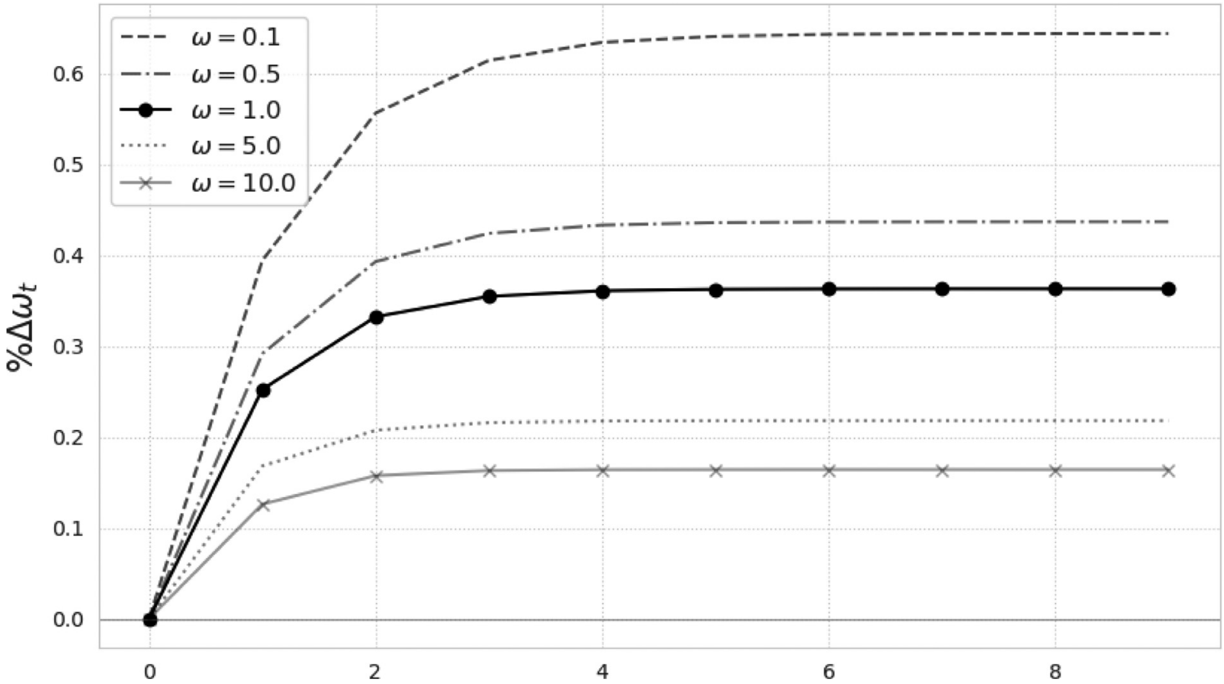

Fig. E.1. Equivalent shock series to ω.

be guaranteed. Below I derive conditions for which the problem is well-defined when the value function is continuously differentiable and find that those are easily satisfied for the examples considered in the main text.

Limiting Case $\sigma_\varepsilon^2 \to 0$.

It can be shown that as the irreducible noise vanishes the problem is well-defined. The FOC for an interior solution simplifies to

$$0 = -(\lambda - 1) + \omega \frac{1}{Y} - \beta\omega \frac{1}{X'} \tag{F.27}$$

with a second order condition for a maximum of

$$-\omega \frac{1}{Y^2} + \beta\omega \frac{1}{X'^2} \leq 0 \tag{F.28}$$

which is satisfied since $Y \leq X$ and $\beta \in (0, 1)$.

General Case Let us restrict our attention to the interior solution. Recall that the First Order Condition is in general given by

$$0 = -(\lambda - 1) + \omega \frac{1}{\delta(X)^2 Y} + \beta V'\left(\delta(X)^2 Y + \delta(X)\sigma_\varepsilon^2 + \sigma_\varepsilon^2 + \frac{\sigma_u^2}{\phi^2}\right) \tag{F.29}$$

while the sufficient condition for a maximum reads

$$-\omega \frac{\delta(X)^2}{[\delta(X)^2 Y]^2} + \beta V''\left(X'\right)\delta(X)^2 \leq 0 \tag{F.30}$$

which shows that the problem is well-defined as long as the value function is not "too convex". That is,

$$\beta V''\left(X'\right) \leq \omega \frac{1}{[\delta(X)^2 Y]^2} \tag{F.31}$$

To assess how reasonable this condition is I plot the function

$$-\left(\sigma_\varepsilon^2 \delta(X) + \delta(X)^2 Y\right) - \omega \log\left(\frac{X}{Y}\right) + \beta V(X') \tag{F.32}$$

for plausible values of prior variances X and X' around the steady state. Fig. F.1 verifies numerically that for plausible prior variances (i.e. not too far from steady state) the problem is concave and has a unique maximum.

Negative Slope of Policy Function The policy function has two parts. An upward-sloping region for low realisations of the state in which Y takes their corner solution of $Y = X$ and a downward-sloping portion. The condition for a downward-sloping policy function is slightly stricter than the condition for concavity. Below I will derive such condition which I find to be satisfied in all specifications considered in the main text.

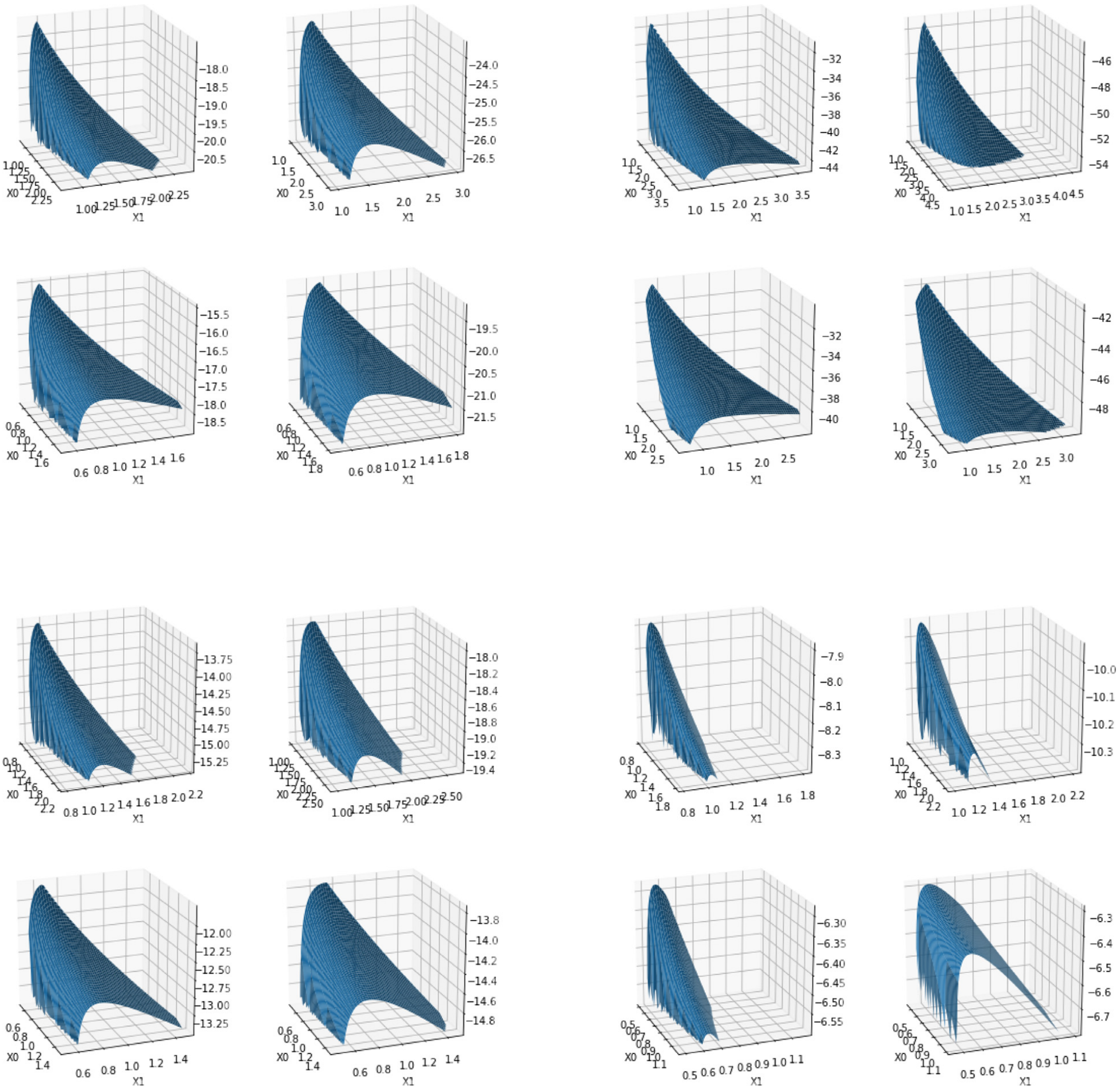

Fig. F.1. Four 2 × 2 grids for each $\omega \in \{1, 5, 0.5, 0.1\}$ respectively. Within each grid the different signal to noise ratios are low, medium, medium, high from left to right respectively.

Let us denote the optimal policy function for current period's posterior variance by $h(X)$ and for next period's prior variance by $g(X) = \delta(X)^2 h(X) + \delta(X)\sigma_\varepsilon^2 + \varepsilon + \frac{\sigma_u^2}{\phi^2}$. By the Implicit Function Theorem on the First Order Condition and denoting $A(X) := \delta(X)^2 h(X)$

$$\left(\frac{\omega}{[A(X)]^2} - \beta V''(g(X))\right)\left[2\delta(X)\delta'(X)h(X) + \delta(X)^2 h'(X)\right] = \beta V''(g(X))\delta'(X)\sigma_\varepsilon^2 \tag{F.33}$$

such that

$$h'(X) = \frac{\delta'(X)}{\delta(X)}\left[-2h(X) + \frac{\beta V''(g(X))}{\frac{\omega}{[A(X)]^2} - \beta V''(g(X))}\frac{\sigma_\varepsilon^2}{\delta(X)}\right] \tag{F.34}$$

and hence $h'(X) \le 0$ iff

$$\beta V''(g(X)) \le \left[\frac{\sigma_\varepsilon^2}{2\delta(X)h(X)} + 1\right]^{-1}\frac{\omega}{[A(X)]^2} \le \frac{\omega}{[A(X)]^2} \tag{F.35}$$

It follows that the second order condition for $h(X)$ as the maximiser is necessary but not sufficient for a downward-sloping policy function. Also note that as $\sigma_\varepsilon^2 \to 0$, $\delta'(X) \to 0$ and $h'(X) \to 0$. That is, the optimal posterior variance is independent of its prior variance.

References

Afrouzi, H., Yang, C., 2021. Dynamic Rational Inattention and the Phillips Curve. CESifo Working Paper Series. CESifo. 8840
Aruoba, S.B., 2008. Data revisions are not well behaved. J. Money Credit Banking 40 (2–3), 319–340. doi:10.1111/j.1538-4616.2008.00115.x. https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4616.2008.00115.x
Bloedel, A., Segal, I., 2021. Persuading a Rationally Inattentive Agent. Working Paper.
Bloom, N., 2009. The impact of uncertainty shocks. Econometrica 77 (3), 623–685. https://ideas.repec.org/a/ecm/emetrp/v77y2009i3p623-685.html
Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., Terry, S.J., 2018. Really uncertain business cycles. Econometrica 86 (3), 1031–1065. doi:10.3982/ECTA10927. https://ideas.repec.org/a/wly/emetrp/v86y2018i3p1031-1065.html
Charles Bean, 2007. Risk, Uncertainty and Monetary Policy. Speech given to Dow Jones, at City Club, Old Broad Street https://www.bankofengland.co.uk/-/media/boe/files/speech/2007/risk-uncertainty-and-monetary-policy.pdf.
Coibion, O., Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: a simple framework and new facts. American Economic Review 105 (8), 2644–2678. doi:10.1257/aer.20110306. https://www.aeaweb.org/articles?id=10.1257/aer.20110306
Coibion, O., Gorodnichenko, Y., Kumar, S., 2018. How do firms form their expectations? new survey evidence. American Economic Review 108 (9), 2671–2713. doi:10.1257/aer.20151299. https://www.aeaweb.org/articles?id=10.1257/aer.20151299
Cover, T., Thomas, J., 2012. Elements of information theory. Wiley. https://books.google.co.uk/books?id=VWq5GG6ycxMC
Croushore, D., 2011. Frontiers of real-Time data analysis. J. Econ. Lit. 49 (1), 72–100. https://ideas.repec.org/a/aea/jeclit/v49y2011i1p72-100.html
Croushore, D., Stark, T., 2001. A real-time data set for macroeconomists. J. Econom. 105 (1), 111–130. https://EconPapers.repec.org/RePEc:eee:econom:v:105:y:2001:i:1:p:111-130
Fajgelbaum, P.D., Schaal, E., Taschereau-Dumouchel, M., 2017. Uncertainty traps. Q. J. Econ. 132 (4), 1641–1692. https://ideas.repec.org/a/oup/qjecon/v132y2017i4p1641-1692..html
Fuster, A., Perez-Truglia, R., Wiederholt, M., Zafar, B., 2018. Expectations with endogenous information acquisition: an experimental investigation. NBER Working Papers 24767. National Bureau of Economic Research, Inc. https://ideas.repec.org/p/nbr/nberwo/24767.html
Haldane, A., Macaulay, A., McMahon, M., 2020. The 3 E's of central bank communication with the public. Bank of England working papers 847. Bank of England. https://ideas.repec.org/p/boe/boeewp/0847.html
Jordá, O., Kouchekinia, N., Merrill, C., Sekhposyan, T., 2020. The fog of numbers. FRBSF Econ. Lett. 2020 (20), 1–5. https://ideas.repec.org/a/fip/fedfel/88397.html
Lipnowski, E., Mathevet, L., Wei, D., 2020. Attention management. American Economic Review: Insights 2 (1), 17–32. doi:10.1257/aeri.20190165. https://www.aeaweb.org/articles?id=10.1257/aeri.20190165
Lorenzoni, G., 2009. A theory of demand shocks. Am. Econ. Rev. 99 (5), 2050–2084. doi:10.1257/aer.99.5.2050. https://www.aeaweb.org/articles?id=10.1257/aer.99.5.2050
Manski, C.F., 2015. Communicating uncertainty in official economic statistics: an appraisal fifty years after morgenstern. J. Econ. Lit. 53 (3), 631–653. doi:10.1257/jel.53.3.631. https://www.aeaweb.org/articles?id=10.1257/jel.53.3.631
Marini, M., Shrestha, M.L., 2013. Quarterly GDP Revisions in G-20 Countries: Evidence from the 2008 Financial Crisis. IMF Working Papers 2013/060. International Monetary Fund. https://ideas.repec.org/p/imf/imfwpa/2013-060.html
Maćkowiak, B., Matějka, F., Wiederholt, M., 2018. Dynamic rational inattention: analytical results. J. Econ. Theory 176 (C), 650–692. doi:10.1016/j.jet.2018.05.001. https://ideas.repec.org/a/eee/jetheo/v176y2018icp650-692.html
Maćkowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. Am. Econ. Rev. 99 (3), 769–803. https://ideas.repec.org/a/aea/aecrev/v99y2009i3p769-803.html
Maćkowiak, B., Wiederholt, M., 2015. Business cycle dynamics under rational inattention. Rev. Econ. Stud. 82 (4), 1502–1532. https://ideas.repec.org/a/oup/restud/v82y2015i4p1502-1532..html
Maćkowiak, B.A., Matejka, F., Wiederholt, M., 2018. Survey: Rational Inattention, a Disciplined Behavioral Model. CEPR Discussion Papers 13243. C.E.P.R. Discussion Papers. https://ideas.repec.org/p/cpr/ceprdp/13243.html
Myatt, D.P., Wallace, C., 2012. Endogenous information acquisition in coordination games. Rev. Econ. Stud. 79 (1), 340–374. doi:10.1093/restud/rdr018. https://academic.oup.com/restud/article-pdf/79/1/340/18404289/rdr018.pdf
Orphanides, A., 2001. Monetary policy rules based on real-time data. Am. Econ. Rev. 91 (4), 964–985. doi:10.1257/aer.91.4.964. https://www.aeaweb.org/articles?id=10.1257/aer.91.4.964
Paciello, L., Wiederholt, M., 2014. Exogenous information, endogenous information, and optimal monetary policy. Rev. Econ. Stud. 81 (1), 356–388. https://ideas.repec.org/a/oup/restud/v81y2014i1p356-388.html
Reis, R., 2011. When Should Policymakers Make Announcements? 2011 Meeting Papers 122. Society for Economic Dynamics. https://EconPapers.repec.org/RePEc:red:sed011:122
Sims, C.A., 2003. Implications of rational inattention. J. Monet. Econ. 50 (3), 665–690. https://ideas.repec.org/a/eee/moneco/v50y2003i3p665-690.html
Sims, C.A., 2010. Rational inattention and monetary economics. In: Friedman, B.M., Woodford, M. (Eds.), Handbook of Monetary Economics. In: Handbook of Monetary Economics, chapter 4, Vol. 3. Elsevier, pp. 155–181. https://ideas.repec.org/h/eee/monchp/3-04.html
Stokey, N., Lucas, R., Prescott, E., 1989. Recursive methods in economic dynamics. Harvard University Press. https://books.google.co.uk/books?id=tWYo0QolyLAC
Swanson, N., van Dijk, D., 2006. Are statistical reporting agencies getting it right? Data rationality and business cycle asymmetry. J. Bus. Econ. Stat. 24, 24–42. https://EconPapers.repec.org/RePEc:bes:jnlbes:v:24:y:2006:p:24-42
Van Nieuwerburgh, S., Veldkamp, L., 2006. Learning asymmetries in real business cycles. J. Monet. Econ. 53 (4), 753–772. https://ideas.repec.org/a/eee/moneco/v53y2006i4p753-772.html