Essays in Monetary and Information Economics

by

Bogdan Marcu

Thesis submitted for the degree of

Doctor of Philosophy in Economics

Department of Economics, University of Warwick

March 2022

Contents

List of Figures

Acknowledgements

Firstly, a special thanks to my supervisor Herakles Polemarchakis for having guided me throughout these years. The thesis would not have been possible without his advice and encouragement. Further, I thank Giovanni Ricco for his comments and for his support towards the end of my studies. I would also like to extend my gratitude to Costas Cavounidis, Roger Farmer, Peter Hammond, Thijs van Rens and the Warwick Macro Group for their feedback and comments.

Further, I would like to thank my extended family (especially my parents and my uncle Bogdan), as well as my girlfriend and my friends for their personal support. I am very fortunate to have you in my life.

Financial assistance from the Department of Economics, University of Warwick and the Economic and Social Research Council is gratefully acknowledged.

Declaration

I submit this thesis to the University of Warwick in accordance with the requirements of the degree of Doctor of Philosophy in Economics. I declare that it has not been submitted for a degree at another university.

March 2022

Abstract

Broadly, the thesis concerns information aggregation in large populations of agents, each of whom has his own "window on the world" (i.e. some imperfect and private information regarding the state of the world). The thesis explores how a policymaker may seek to use the information revealed by the agents' actions to improve social welfare via information disclosures, or via the setting of a policy instrument. In particular, it studies how the policymaker's actions may distort the agents' incentives to acquire (or process) information and, in turn, how this influences the information revealed by the agents' actions and hence the policymaker's ability to improve social welfare in the first place. The three chapters are summarised below.

In chapter 1, I analyse the social value of public information in the context of a static prediction model, where each agent has access to two information sources: a (costly) private signal of the fundamental and a (free) public signal of the average prediction of the other agents. I argue that providing more precise public information (about the average prediction) is welfare-improving only up to a certain threshold, above which it crowds out the acquisition of private information and no longer affects welfare (i.e. if public information is already sufficiently precise such that some agents do not acquire the private signal in equilibrium, then further increasing the precision of public information will not affect welfare).

In chapter 2, I study optimal monetary policy in a setting where firms are rationally inattentive and the central bank learns about fundamentals by observing market prices. I argue that optimal policy minimizes the central bank's own information precision about fundamentals, and I discuss the implications of this tension. The model provides an informational rationale for increased policy activism at times of high aggregate volatility (for instance, during recessions) and for a higher degree of monetary neutrality during such times.

In chapter 3, I again analyse the relationship between optimal policy intervention and price informativeness, but I also account for parameter uncertainty in the spirit of Brainard (1967), i.e. uncertainty regarding the transmission of policy itself. I argue that under parameter uncertainty, the central bank cannot perfectly disentangle the effects of its policy from fundamental shocks moving prices, so policy intervention necessarily crowds out some of the information (concerning fundamentals) which is contained in prices. In terms of central bank learning, this leads to a similar trade-off as in Balvers and Cosimano (1994).

1 A Note on Static Social Learning when Private Information is Costly

1.1 Introduction

In many economic situations, agents make decisions under imperfect and dispersed information concerning payoff-relevant fundamentals. In such contexts, there is scope for social learning, in the sense that agents may benefit by learning from others (Vives (1996); Vives (1997); Burguet and Vives (2000); Bru and Vives (2002)). This paper analyses social learning in the context of a static prediction model featuring costly information acquisition. More specifically, it studies a "herding prediction model with a rational-expectations flavor" as in Bru and Vives (2002) — where each agent out of a continuum seeks to predict a random variable after observing a private signal of its realization, as well as a noisy public signal of the average prediction of the other agents — but in extension to Bru and Vives (2002), it considers the possibility that the agents' private signals are costly.[1] It argues that, in the presence of information acquisition costs, the provision of more precise public information[2] is no longer welfare-improving beyond a certain threshold, while in the absence of information acquisition costs, more precise public information is always welfare-improving (the latter point follows directly from Bru and Vives (2002)).

Similarly to Bru and Vives (2002), there are information externalities at play, as agents do not internalize the effects of their decisions[3] on the informativeness of the public signal of the average action.[4] More specifically, in the same way as in Bru and Vives (2002), informed agents do not take into account how their reaction to their private information affects the informativeness of public information.[5] Additionally, agents do not internalize how their *acquisition of private information* affects the informativeness of the public signal — this latter mechanism[6] is responsible for the main result (namely, that increasing the precision with which agents

[1]In the sense that each agent needs to pay a fixed cost to observe the private signal.

[2]Referring to a higher precision (or lower noise) associated with the public signal of the average prediction.

[3]Referring to both the decision of how to respond to private information, and to the decision whether to acquire the private signal or not, as detailed below.

[4]Throughout the paper, I use the terms "average action" and "average prediction" interchangeably — as will become clear in the context of the model, each agent's action will also be his prediction regarding the fundamental.

[5]And welfare could be improved if informed agents responded more strongly to their private signals, thereby generating more precise public information.

[6]Which is not captured in Bru and Vives (2002), where private information is free.

observe the average prediction is no longer welfare-improving beyond a threshold). When public information is sufficiently precise, only a fraction of agents acquire the private signal in equilibrium — in all such equilibria, holding constant the fraction of agents who are informed, increasing the precision with which agents observe the average prediction improves the informativeness of the public signal about the fundamental (and expected welfare), in a similar fashion to Bru and Vives (2002); but it decreases the returns to acquiring information and leads less agents to become informed (which deteriorates the informativeness of the public signal). In equilibrium, the two effects cancel each other out, such that the informativeness of public information (and thus expected welfare) remain constant in response to changes in the precision with which the average action is observed. In other words, providing a more precise signal of the average action is no longer welfare-improving beyond a certain threshold, whereby all agents acquire the private signal below the threshold, but only a fraction of agents acquire the private signal above the threshold. Thus, if the public signal of the average action conveys information from informed to uninformed agents, then increasing its precision will not improve welfare.

This is in contrast to Bru and Vives (2002), where increasing the precision with which agents observe the average action is always welfare-improving. Hence, the model illustrates that there is an inherent conflict between the incentives to acquire private information and the ability of average actions to provide informative public signals. This resembles Grossman and Stiglitz (1980), where "there is a fundamental conflict between the efficiency with which markets spread information and the incentives to acquire information". In their model, the informativeness of the price system is independent of the amount of noise in the supply of the asset (which is the only source of noise in Grossman and Stiglitz (1980), preventing the asset price from being perfectly informative). This paper presents a parallel of their result in the context of social learning.[7] I also discuss how the result could be interpreted in the context of the literature analysing central bank transparency in endogenous information settings, where the policymaker learns about the state of nature from market prices while simultaneously disclosing information to firms.

Before proceeding to the main text, I provide an illustrative example.[8] Consider a unit mass of firms, each of whom chooses its price p_i to maximize profit subject to unknown market conditions θ. Suppose that firm i's profit is given by $\pi_i = -(p_i - \theta)^2$

[7]This is discussed in more detail in section 1.3.

[8]More examples are provided below, following Bru and Vives (2002).

which can be derived from a log-quadratic approximation of the profit function where $\theta \sim N(0, \sigma_\theta^2)$ is the target price (assumed here to be common across firms and normally distributed). Further, suppose that each firm observes a private signal of the target price $x_i = \theta + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_x^2)$ is i.i.d. across firms (such that private information is noisy and dispersed). It is easy to note that in the absence of any other information source, each firm sets its price according to $p_i = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_x^2} x_i$ and, by virtue of the law of large numbers, the average price reveals the target price $p = \int_0^1 p_i di = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_x^2}\theta$.

A natural question which arises in such a context is whether providing information to firms of the form "average price" plus noise is welfare-improving.[9] For instance, a government agency could compile aggregate price statistics, but these may be subject to measurement error — the policymaker may wonder whether a lower measurement error helps firms make better pricing decisions. The answer provided by Bru and Vives (2002) is that this is indeed true. In other words, making the average price observable with less noise always helps firms better match the target price and achieve higher expected profits[10] (despite the fact that firms do not make socially optimal use of information in the presence of the public signal of the average price, i.e. there is an information externality at play)[11] — for a textbook discussion of this, see Vives (2010), chapter 6.6.

Note: *A similar example is provided in Bru and Vives (2002): "competitive firms decide about investment with macroeconomic uncertainty represented by the random variable θ which determines average profitability. Firms invest taking into account that the profits of investment will depend on the realization of θ. To predict θ each firm has access to a private signal as well as to public information, which is formed by aggregate investment figures compiled by a government agency. Data on aggregate investment incorporates measurement error". For more examples, see Bru and Vives (2002), section 2, or Vives (2010), chapter 6.6.1.*

This paper analyses the social benefit of more precise public information (in the sense described above) when firms need to pay a fixed cost to observe the private signal x_i. It argues that in this case, increasing the precision with which firms ob-

[9]More specifically, I am now referring to a setting where each firm observes its private signal x_i, as well as a noisy signal of the average price when deciding on its own price. The question is whether making the average price observable with higher precision helps firms better match the target price.

[10]See also Corollary 1.2.

[11]More details are provided in Remark 1.1.

serve the average price is welfare-improving only up to a certain threshold — if the precision increases above this threshold, only a fraction of firms acquire the private signal x_i in equilibrium;[12] the main result states that the informativeness of public information (and expected welfare) are the same in all such equilibria.

1.2 Related literature

As previously mentioned, the model closest to the one presented here is Bru and Vives (2002), which I extend by considering the possibility that the agents' private signals are costly. As such, the paper relates more broadly to the literature on social learning, herding and information cascades — see Vives (2010), chapter 6, and the references therein for a general discussion of these topics, and chapter 6.6 for a textbook treatment of Bru and Vives (2002). However, I only analyse the equilibrium behavior of agents, and do not address their socially optimal behavior — in the language of Vives (2010), I analyse only the "market solution" here and do not address the "team efficient solution".[13]

The interaction between social learning and costly information acquisition has been explored in a similar context by Burguet and Vives (2000) — for a textbook treatment, see Vives (2010), chapter 6.3.2. The main difference between their model and the one presented here is that they analyse a dynamic (sequential) prediction problem — where each period, a new generation of short-lived agents predict a random variable after observing a private signal, as well as noisy public signals of the average predictions of agents in past generations — whereas here, I study a static prediction problem, with a "rational-expectations flavour". Another important difference lies in the assumptions concerning costly information acquisition — in Burguet and Vives (2000), each agent can pay a higher cost to increase the precision of his private signal, whereas here each agent faces a binary choice whether to acquire the private signal or not (and the precision of the private signal is exogenous). This assumption concerning costly information acquisition allows me to draw a parallel with Grossman and Stiglitz (1980), as well as with Bernanke and Woodford (1997), as I detail in sections 1.3 and 1.3.1 respectively.[14]

[12] Because if all agents acquired the private signal x_i, then the noisy signal of the average price would be sufficiently informative about the fundamental such that no firm would want to acquire the private signal in the first place. The intuition for this relates to Grossman and Stiglitz (1980) and Bernanke and Woodford (1997) — I expand on this in the main text.

[13] The main result here (i.e. Corollary 1.4) concerns the *equilibrium* acquisition and use of information. Clearly, if the agents' acquisition and use of information was *efficient* (i.e. socially optimal), increasing the precision with which they observe the average action would always be welfare-improving.

[14] Because both papers employ similar assumptions regarding the acquisition of private information. In Grossman and Stiglitz (1980), traders face a binary choice whether to pay a fixed cost

Interestingly, the paper also relates to models analysing central bank transparency in endogenous information settings, where the policymaker learns about the state of nature from market prices (Morris and Shin (2005); Baeriswyl (2011); Baeriswyl et al. (2020)). In these models, firms set prices under imperfect information and the central bank observes a noisy signal of the average price, while simultaneously disclosing (with noise) what it learns about the fundamental (from the noisy signal of the average price) back to the firms. In such a model, the central bank practically mediates the firms' learning from the noisy signal of the price level, whereas here agents learn directly by observing the noisy signal of the average action.[15] The relation with this strand of the literature is explored in more detail in section 1.3.1.

1.3 Model

There is a continuum of agents (indexed by i on the unit interval) who face a linear-quadratic Gaussian problem. Each agent chooses an action $a_i \in \mathbb{R}$ to minimize the distance from a normally distributed fundamental $\theta \sim N(0, 1/\tau_\theta)$.

Information structure: Each agent can pay a fixed cost $c \geq 0$ to observe a private signal of the fundamental with an exogenous precision τ_x:

$$x_i = \theta + \varepsilon_i \tag{1.1}$$

Where $\varepsilon_i \sim N(0, 1/\tau_x)$ is i.i.d. across agents.[16] Additionally, there is a noisy public signal of the average action:

$$y = \overline{a} + \varepsilon_y \tag{1.2}$$

Where $\overline{a} = \int_0^1 a_i di$ denotes the average action, and where $\varepsilon_y \sim N(0, 1/\tau_y)$ is a white noise term, distributed independently of all other random variables. I will investigate the welfare effects of changes in τ_y.

to become informed or not. In Bernanke and Woodford (1997), section 1, forecasters also face a similar binary choice (when costly information acquisition is taken into account).

[15] Unlike Baeriswyl (2011), or Baeriswyl et al. (2020), I abstract from strategic complementarities in price-setting. I also assume that the fundamental is drawn from a proper, Gaussian prior.

[16] The convention is made that a version of the law of large numbers for a continuum of random variables holds, such that errors on average cancel out almost surely (which implies that the average signal across agents reveals the fundamental). For a discussion of the measure-theoretic issues related to using a law of large numbers when integrating across a continuum of random variables, see Judd (1985); Uhlig (1996). The convention is maintained throughout the thesis.

Each agent's payoff is given by:

$$u_i = -(a_i - \theta)^2 - c\,\mathbb{1}_{i \in I} \tag{1.3}$$

Where a_i denotes agent i's action and $\mathbb{1}_{i \in I}$ indicates whether agent i is informed or not (i.e. takes value 1 if agent i observes the private signal x_i and 0 otherwise).

Timing: This is a two-stage game where:

1) firstly, each agent decides whether to buy the private signal (x_i) or not

2) secondly, each agent receives his information (i.e. informed agents observe x_i and y, while uninformed agents observe y) and takes his action a_i

Equilibrium definition: An equilibrium specifies:

i) an information acquisition strategy for each player (specifying whether he buys the private signal x_i or not)

ii) action rules for informed and uninformed agents (mapping signal realizations to actions)

Such that:

i) each agent acts optimally given his information (i.e. $a_i^{I*} = E[\theta|x_i, y]$, $a_i^{U*} = E[\theta|y]$)

ii) each agent's information acquisition strategy is optimal (given the other agents' information acquisition strategies)

When choosing whether to buy the private signal or not, each agent takes as given the other agents' information acquisition strategies. Also, when choosing how to act conditional on the information received, each agent takes as given the other agents' action rules. As is standard in the literature, given that uncertainty is Gaussian and payoffs are quadratic, I restrict attention to linear equilibria — i.e. equilibria in which each agent's action rule is linear in his signals.[17] I also restrict attention

[17] Because uncertainty is Gaussian and payoffs are quadratic, each agent's best-response is linear in his signals when the other agents' action rules are linear in their signals.

to symmetric equilibria in which each type of agent uses the same action rule.[18]

As such, an equilibrium specifies a tuple:

$$(\lambda^*, \phi_1^{I*}, \phi_2^{I*}, \phi^{U*})$$

Such that:

$$\begin{cases} a_i^{I*} = E[\theta|x_i, y] = \phi_1^{I*}x_i + \phi_2^{I*}y \\ a_i^{U*} = E[\theta|y] = \phi^{U*}y \end{cases} \tag{1.4}$$

$$\begin{cases} \lambda^* = 1 & \text{if} \quad L_i/L_u < 1 \\ \lambda^* = 0 & \text{if} \quad L_i/L_u > 1 \\ \lambda^* \in [0,1] & \text{if} \quad L_i/L_u = 1 \end{cases} \tag{1.5}$$

Where

$$L_i = E[(\phi_1^{I*}x_i + \phi_2^{I*}y - \theta)^2] + c \tag{1.6}$$

And

$$L_u = E[(\phi^{U*}y - \theta)^2] \tag{1.7}$$

$\lambda^* \in [0,1]$ denotes the fraction of agents who become informed in equilibrium, ϕ_1^{I*} and ϕ_2^{I*} denote the responsiveness of informed agents' actions to the private and public signal respectively, and ϕ^{U*} denotes the responsiveness of uninformed agents' actions to the public signal.

(1.4) states that both informed and uninformed agents act optimally conditional on their information (to maximize individual utility (1.3)) — in other words, this is an equilibrium condition for the second step of the game. Given that agents act optimally conditional on their information, the expected losses of informed and uninformed agents are given by (1.6) and (1.7) respectively. An equilibrium in the first step of the game then requires (1.5).

Using a utilitarian welfare criterion which places an equal weight on all agents, the welfare loss is defined as:

$$L_W = -\int_0^1 E(u_i)di \tag{1.8}$$

[18]This is without loss of generality.

And, in equilibrium, this can be expressed as:

$$L_W^* = \lambda^* L_i + (1 - \lambda^*) L_u \tag{1.9}$$

I derive the equilibrium by backward induction. Hence, I first determine the equilibrium action rules for a given λ (step two in the timing of events). Then, I determine the equilibrium λ (step one in the timing of events). These are characterised in Propositions 1.1 and 1.2 respectively. Unless mentioned otherwise, I restrict attention to strictly positive, finite τ_x, τ_y, τ_θ.

Proposition 1.1. *Suppose a fraction $\lambda \in (0, 1]$ of agents are informed. Then, there is a unique equilibrium in the second step of the game, in which ϕ_1^{I*} solves:*[19]

$$\phi_1^{I*} = \tau_x / \left[\tau_x + \tau_\theta + \tau_y (\lambda \phi_1^{I*})^2\right] \tag{1.10}$$

Then, ϕ_2^{I*} and ϕ^{U*} can be computed as functions of ϕ_1^{I*} (see equations (A.8) and (A.9)).

Proof: See Appendix A.1. □

Note: *This is similar to Proposition 1 in Bru and Vives (2002) — more specifically, they analyse the case in which all agents are informed (i.e. when $\lambda = 1$).*

From the perspective of each informed agent, the optimal reaction to his private signal x_i depends on the reaction of the other informed agents to their private signals (because this affects the informativeness of the public signal). More specifically, denote by $\overline{\phi}_1^I$ the responsiveness of all informed agents $j \neq i$ to their private signals (i.e. suppose all informed agents $j \neq i$ act according to $a_j^I = \overline{\phi}_1^I x_j + \overline{\phi}_2^I y$). Then, the (individually) optimal responsiveness of agent i's action to the private signal x_i (denoted below by $\phi_{1,i}^{I*}$) can be expressed as:

$$\phi_{1,i}^{I*} = \tau_x / \left[\tau_x + \tau_\theta + \tau_y (\lambda \overline{\phi}_1^I)^2\right] \tag{1.11}$$

This depends on the precision of the private signal (τ_x), the precision of the prior (τ_θ), and the informativeness of the public signal about the fundamental[20] (which is captured by the last term in the denominator on the R.H.S. of (1.11); for more

[19] An analytical expression for ϕ_1^{I*} is provided in Appendix A.1 — see equation (A.10).

[20] Which is defined as the precision of the unbiased signal of the fundamental contained in the public signal of the average action — for more details, see (A.4) in Appendix A.1.

details, see Appendix A.1). The informativeness of the public signal itself depends on the precision with which agents observe the average action (τ_y), the fraction of agents who are informed (λ) and the responsiveness of other informed agents' actions to their private signals ($\overline{\phi}_1^I$). If other informed agents react more strongly to their private information when taking their actions (i.e. a higher $\overline{\phi}_1^I$), then the public signal becomes more informative about the fundamental, so each informed agent finds it optimal to react less strongly to his private signal and more strongly to the public signal (i.e. a lower $\phi_{1,i}^{I*}$). As equilibria in the second step of the game are necessarily symmetric, equation (1.10) follows from (1.11). For more details, see Appendix A.1.

Remark 1.1. *Informed agents do not make socially optimal use of their private information. More specifically, expected welfare would be higher if all informed agents responded more strongly to their private information (thereby generating more precise public information and leading to lower expected losses for both informed and uninformed agents). As previously mentioned, I focus the discussion in this paper on the equilibrium use of information (and do not address the efficient use of information), thus I do not solve the "team problem" here.*[21] *For a discussion of the efficient use of information in this setting when all agents are informed ($\lambda = 1$), see Bru and Vives (2002), or Vives (2010).*

Before proceeding, it is helpful to discuss some properties of the equilibrium in the second step of the game, which will be useful to establish the existence and uniqueness of the equilibrium in the first step of the game. Firstly, we will note (in Corollary 1.1) that, holding everything else constant, increasing the fraction of agents who are informed (i.e. a higher λ), or increasing the precision with which agents observe the average action (i.e. a higher τ_y), increases the informativeness of the public signal about the fundamental (in the equilibrium reached in the second step of the game). Secondly, we will note that a higher informativeness of the public signal benefits uninformed agents relatively more than informed agents (because the marginal benefit of additional information is higher for uninformed agents).

To formalise these remarks, it is useful to introduce some notation. Let $\tau_s(\lambda, \tau_y)$ denote the informativeness of the public signal about the fundamental (in equilib-

[21] Indeed, the main result of the paper concerns the equilibrium acquisition and use of information (clearly, if the agents' acquisition and use of information were socially optimal, increasing the precision with which agents observe the average action would always be welfare-improving).

rium),[22] as a function of λ and τ_y (we will vary λ to characterise the equilibrium and τ_y to study comparative statics). In other words, after fixing λ and τ_y, one can use Proposition 1.1 to compute the equilibrium action rules in the second step of the game:

$$(\phi_1^{I*}(\lambda, \tau_y); \phi_2^{I*}(\lambda, \tau_y); \phi^{U*}(\lambda, \tau_y))$$

Then, the informativeness of the public signal about the fundamental writes as:[23]

$$\tau_s(\lambda, \tau_y) = \tau_y[\lambda \phi_1^{I*}(\lambda, \tau_y)]^2 \tag{1.12}$$

In the same spirit, denote by $\gamma(\lambda, \tau_y)$ the ratio of the expected loss of informed agents to the expected loss of uninformed agents (in the equilibrium of the second step of the game, as a function of λ and τ_y). Taking into account that agents make (individually) optimal use of their information we can express the expected losses of informed and uninformed agents as:

$$L_i(\lambda, \tau_y) = \frac{1}{\tau_x + \tau_\theta + \tau_s(\lambda, \tau_y)} + c \tag{1.13}$$

$$L_u(\lambda, \tau_y) = \frac{1}{\tau_\theta + \tau_s(\lambda, \tau_y)} \tag{1.14}$$

Hence, $\gamma(\lambda, \tau_y)$ writes as:

$$\gamma(\lambda, \tau_y) = \frac{L_i(\lambda, \tau_y)}{L_u(\lambda, \tau_y)} = \left[\frac{1}{\tau_x + \tau_\theta + \tau_s(\lambda, \tau_y)} + c\right] / \left[\frac{1}{\tau_\theta + \tau_s(\lambda, \tau_y)}\right] \tag{1.15}$$

Corollary 1.1. *Holding everything else constant, increasing the fraction of agents who are informed (λ), or increasing τ_y, improves the informativeness of the public signal about the fundamental:*

$$\frac{\partial \tau_s(\lambda, \tau_y)}{\partial \tau_y} > 0; \frac{\partial \tau_s(\lambda, \tau_y)}{\partial \lambda} > 0$$

Proof: See Appendix A.2. □

Corollary 1.2. *Suppose private information is free ($c = 0$). Increasing τ_y is welfare-improving:*

$$\frac{\partial L_W^*}{\partial \tau_y}_{|c=0} < 0$$

[22] Or, equivalently, the precision of the unbiased signal of the fundamental contained in y.
[23] See the discussion above, or equation (A.4).

Proof: See Appendix A.3. □

Note: *Corollary 1.2 is also proved in Bru and Vives (2002).*

Corollary 1.3. *Holding everything else constant, increasing the fraction of agents who are informed (λ), or increasing τ_y benefits uninformed agents relatively more than informed agents:*

$$\frac{\partial\gamma(\lambda,\tau_y)}{\partial\lambda} > 0; \frac{\partial\gamma(\lambda,\tau_y)}{\partial\tau_y} > 0$$

Proof: See Appendix A.4. □

Remark 1.2. *In the limit, as the average action becomes observable with almost perfect precision (i.e. $\tau_y \to \infty$), the public signal becomes almost perfectly informative about the fundamental (i.e. $\tau_s(\lambda,\tau_y) \to \infty$),*[24] *so $L_u(\lambda,\tau_y) \to 0$, while $L_i(\lambda,\tau_y) \to c$, see (1.13) and (1.14).*

On the other hand, if the average action is observed perfectly (i.e. $y = \overline{a}$), there is no equilibrium in the second step of the game (for any $\lambda > 0$) — it is straightforward to argue this by contradiction:

- *Suppose that informed agents respond to their private signals ($\phi_1^* \neq 0$). Then, the public signal perfectly reveals the fundamental, and each informed agent finds it optimal to not respond to his (noisy) private signal ($\phi_1^* = 0$), hence a contradiction.*

- *Suppose that informed agents do not respond to their private signals ($\phi_1^* = 0$). Then, the average action is uninformative about the fundamental, so each informed agent finds it optimal to put a strictly positive weight on his private signal x_i (i.e. $\phi_1^{I*} > 0$), hence a contradiction.*

In what follows, I characterise the equilibrium in the first step of the game. Before doing so, I assume that the cost of acquiring the private signal is sufficiently low, such that in the absence of any public information ($\tau_y = 0$), agents acquire the private signal:

$$\frac{1}{\tau_\theta + \tau_x} + c < \frac{1}{\tau_\theta} \Leftrightarrow c < \frac{\tau_x}{\tau_\theta(\tau_x + \tau_\theta)}$$

Otherwise, the unique equilibrium in the first step of the game specifies $\lambda^* = 0$ for any τ_y (i.e. agents never become informed). This assumption is maintained through-

[24] This is easily proved by noting that $\lim_{\tau_y \to \infty} \phi_1^{I*} = 0$ which implies that $\tau_s(\lambda,\tau_y) \to \infty$ as $\tau_y \to \infty$. See also Bru and Vives (2002).

out the rest of the paper. Proposition 1.2 characterises the equilibrium in the first step of the game (for different values of τ_y) and Corollary 1.4 discusses the welfare effects of more precise public information provision.

Proposition 1.2. *Let $\hat{\tau}_y$ denote the solution to $\gamma(1, \hat{\tau}_y) = 1$.*[25] *For $\tau_y \leq \hat{\tau}_y$ the unique equilibrium specifies $\lambda^* = 1$. For $\tau_y > \hat{\tau}_y$, the fraction of agents who become informed in equilibrium ($\lambda^* \in (0, 1)$) solves $\gamma(\lambda^*, \tau_y) = 1$.*[26] *There is always a unique equilibrium for any parameter values.*

Corollary 1.4. *If $\tau_y \leq \hat{\tau}_y$, increasing τ_y is welfare-improving: $\frac{\partial L_W^*}{\partial \tau_y}_{|\tau_y \leq \hat{\tau}_y} < 0$. On the other hand, if $\tau_y > \hat{\tau}_y$ changing τ_y does not affect the welfare loss: $\frac{\partial L_W^*}{\partial \tau_y}_{|\tau_y > \hat{\tau}_y} = 0$.*

Proof: See Appendix A.5 and Appendix A.6. □

Taken together, Proposition 1.2 and Corollary 1.4 state that there is a threshold level of public information $\hat{\tau}_y$, whereby below the threshold ($\tau_y \leq \hat{\tau}_y$) all agents become informed and more precise public information is welfare-improving, whereas above the threshold ($\tau_y > \hat{\tau}_y$) only a fraction of agents become informed and more precise public information is no longer welfare-improving. In other words, if there are at least some agents who do not acquire the private signal in equilibrium, then increasing the precision with which agents observe the average action will not improve expected welfare.

Equilibria where public information is sufficiently imprecise, such that all agents acquire the private signal ($\tau_y \leq \hat{\tau}_y$), are similar to the ones in Bru and Vives (2002), and their properties have already been discussed above. For instance, the first part of Corollary 1.4 is analogous to Corollary 1.2.

The main result, namely the second part of Corollary 1.4, follows readily after observing that, for sufficiently high precision of public information ($\tau_y > \hat{\tau}_y$), a fraction less than one of agents become informed in equilibrium ($\lambda^* < 1$),[27] and that in such equilibria, the expected loss of informed agents is necessarily equal to the expected

[25] We will note in the proof that there is a unique $\hat{\tau}_y$ which solves $\gamma(1, \hat{\tau}_y) = 1$.

[26] We will also note in the proof that there is a unique λ^* which solves $\gamma(\lambda^*, \tau_y) = 1$ for $\tau_y > \hat{\tau}_y$. An analytical expression for λ^* is provided in Appendix A.6 — see equation (A.22).

[27] For details, see Appendix A.5.

loss of uninformed agents — i.e. $\gamma(\lambda^*, \tau_y) = 1$, or equivalently:

$$\underbrace{\frac{1}{\tau_\theta + \tau_s(\lambda^*, \tau_y)}}_{L_u(\lambda^*, \tau_y)} = \underbrace{\frac{1}{\tau_x + \tau_\theta + \tau_s(\lambda^*, \tau_y)} + c}_{L_i(\lambda^*, \tau_y)} \tag{1.16}$$

In turn, this requires that the informativeness of the public signal in any such equilibrium (i.e. $\tau_s(\lambda^*, \tau_y)$) is equal to a constant function of parameters, which is independent of the precision with which the average action is observed — in order for agents (in the first step of the game) to be indifferent between acquiring the private signal and not acquiring it, it must be the case that the informativeness of the public signal is given by:[28]

$$\tau_s(\lambda^*, \tau_y) = \left[\sqrt{\frac{\tau_x(4 + \tau_x c)}{c}} - \tau_x\right]/2 - \tau_\theta \tag{1.17}$$

Note that this is independent of τ_y. It follows from (1.16) and (1.17) that the welfare loss in any equilibrium in which $\tau_y > \hat{\tau_y}$ can be expressed as:

$$L_W^* = \lambda^* L_i(\lambda^*, \tau_y) + (1 - \lambda^*) L_u(\lambda^*, \tau_y) = \frac{2}{\sqrt{\frac{\tau_x(4+\tau_x c)}{c}} - \tau_x} \tag{1.18}$$

Which is also independent of τ_y. Also, note that this is independent of the precision of the prior from which the fundamental is drawn (τ_θ). In fact, expected welfare in all equilibria where public information is sufficiently precise ($\tau_y > \hat{\tau_y}$) depends only on the precision of the private signal (τ_x), and the cost of acquiring it (c). For more details, see Appendix A.6. There, I also argue that, because the informativeness of public information is constant in any equilibrium in which $\tau_y > \hat{\tau_y}$ (see (1.17)), the responsiveness of informed agents' actions to their private signals is also constant in any such equilibrium. It follows that any increase in the precision with which agents observe the average action above $\hat{\tau_y}$ is accompanied by a fall in the fraction of agents who acquire the private signal (i.e. a lower λ^*), such that the informativeness of the public signal remains unchanged (for details, see Appendix A.6).

This echoes Grossman and Stiglitz (1980),[29] where the informativeness of the price system is independent of the amount of noise in the asset's supply (because any change in the amount of supply noise is accompanied by a change in the fraction of

[28] i.e. if the informativeness of public information was higher or lower, then agents would no longer be indifferent between buying the private signal and not buying it.

[29] For more details, see part 4 in section H and Theorem 4 in their paper.

traders who acquire private information, such that the informativeness of the price system remains unchanged). To better understand the parallel, remark that in the model presented here, it is the noise in the observation of the average action which prevents agents from perfectly learning the fundamental by observing the public signal, while in Grossman and Stiglitz (1980) it is the noisy supply of the asset which prevents the asset price from being perfectly informative.

"An increase in noise increases the proportion of informed traders. At any given λ,[30] *an increase in noise reduces the informativeness of the price system; but it increases the returns to information and leads more individuals to become informed; the remarkable result obtained above establishes that* ***the two effects exactly offset each other*** *so that the equilibrium informativeness of the price system is unchanged."* (Grossman and Stiglitz (1980))

The intuition for the result in Grossman and Stiglitz (1980) is similar to the one presented above — namely, in order for traders in their model to be indifferent between becoming informed and not becoming informed, the net benefit of acquiring private information must be zero, and this only happens when the informativeness of the price system is a constant function of parameters which is independent of the amount of noise in the asset's supply.

Translated to the context of social learning, the result reads as follows:

"An increase in the observation noise of the average action (lower τ_y) increases the proportion of informed agents. At any given λ, an increase in noise (lower τ_y) reduces the informativeness of the public signal (see Corollary 1.1); but it increases the returns to information and leads more agents to become informed; the result obtained above establishes that ***the two effects exactly offset each other*** *so that the equilibrium informativeness of the public signal is unchanged. (see (1.17))"*

As previously mentioned, in any equilibrium in which each agent is indifferent between acquiring the private signal and not acquiring it, it must be the case that the informativeness of the public signal is constant, such that the net benefit of becoming informed is zero. In turn, this implies that expected welfare is necessarily the same in all equilibria in which at least some agents are uninformed. Hence, if the public signal of the average action conveys information from informed to uninformed

[30]In their model $\lambda \in [0, 1]$ also denotes the fraction of informed agents.

agents, then increasing its precision will no longer be welfare-improving.

Remark 1.3. *The parallel with Grossman and Stiglitz (1980) emerges because of the assumptions concerning costly information acquisition, namely that, in the first step of the game, agents face a binary choice whether to pay a fixed cost to acquire private information or not.*[31] *However, note that there is no impossibility result in the model presented here. One could construct such a result (and a more direct parallel with Grossman and Stiglitz (1980)) by considering a model where the fundamental is the sum of two components: one which is perfectly learnable at a cost, and one which is not — the main result would be similar (namely, in all equilibria in which some agents are informed while others are not, expected welfare is the same).*

1.3.1 Relation to models of endogenous central bank information

In this section, I discuss the relation between the model and the literature analysing central bank transparency in endogenous information settings, where the central bank learns about the state of nature by observing market prices (Morris and Shin (2005); Baeriswyl (2011); Baeriswyl et al. (2020)). Hence, in this section only, I assume that the noisy signal of the average action is no longer observable to agents — instead, there is a policymaker who observes the noisy signal $y = \overline{a} + \varepsilon_y$. The policymaker observes the signal y, constructs an unbiased signal of the fundamental (s) and simultaneously discloses a public signal $z = s + \varepsilon_z$, where $\varepsilon_z \sim N(0, 1/\tau_z)$ is distributed independently of all random variables and τ_z measures the policymaker's degree of transparency.[32,33] I provide more details about the setup in Appendix A.7. In such a setting, it is straightforward to prove the following using similar arguments as in the previous section:

Suppose that private information is free (c=0). Increasing the degree of central bank transparency (i.e. increasing τ_z)) improves the informativeness of the public

[31] The results would be different, for instance, if we allowed agents to pay a higher cost in order to increase the precision of their private signals.

[32] The setup is similar to Baeriswyl (2011) — the main difference being that the model here does not feature strategic complementarities in price-setting (but does feature costly information acquisition). Also, in Baeriswyl (2011), the central bank announces a semi-public signal, whereas here it is purely public (i.e. the error term ε_z is common across firms — modeling public information disclosures as in Baeriswyl (2011) would not make a difference concerning the main result).

[33] Note that if the policymaker is perfectly transparent ($\tau_z \to \infty$), the setup is equivalent to the one from before, when firms directly observed the signal y. If the policymaker is perfectly opaque ($\tau_z = 0$), then there is no public signal.

signal about the fundamental, although there are two opposing effects at play — on one hand, the central bank discloses its information more precisely (which increases the informativeness of the public signal about the fundamental); on the other hand, because firms react to the central bank's disclosure, the noisy signal of the price level becomes less informative about the fundamental, so the central bank's information about the fundamental becomes less precise.[34] *Nonetheless, the latter effect never dominates the former, so a higher degree of transparency always leads the central bank's disclosure to be more informative about the fundamental (hence improving expected welfare).*[35]

Conversely, suppose that private information is costly ($c > 0$). Further, suppose that the central bank observes the average action with sufficiently high precision ($\tau_y > \hat{\tau}_y$, where $\hat{\tau}_y$ is defined in Proposition 1.2). Then, there is a threshold $\hat{\tau}_z$, above which changes in the degree of transparency no longer affect the informativeness of the central bank's disclosure about the fundamental. For $\tau_z > \hat{\tau}_z$, further increasing the degree of central bank transparency prompts less firms to become informed in equilibrium (i.e. leads to a lower λ^*)— in turn, because less firms become informed in equilibrium, the central bank is more poorly informed about the fundamental (because the noisy signal of the average action is less informative). The main result of the paper states that in this case, the direct and indirect effects cancel each other out, such that the informativeness of the central bank's disclosure about the fundamental (and expected welfare) are constant in all equilibria in which $\tau_z > \hat{\tau}_z$ (or equivalently $\lambda^* \in (0,1)$).

Relation to Bernanke and Woodford (1997)

Lastly, I consider a setting where the policymaker perfectly observes the average action (i.e. $y = \overline{a}$), while simultaneously disclosing information to agents. This will allow me to draw a parallel with Bernanke and Woodford (1997), section 1, as I explain below. In this case, it is easy to note that the policymaker perfectly learns the fundamental if: i) at least a fraction of agents are informed ($\lambda^* > 0$), ii) informed agents respond to their private signals ($\phi_1^{I*} \neq 0$).[36] In this setting, it is straightforward to prove the following:

[34] Essentially, the central bank is disclosing with higher precision a signal of the fundamental which is less precise.

[35] Thus, full transparency would be optimal in this context. Note that this is no longer true in richer settings, for instance featuring strategic complementarities (Baeriswyl (2011)).

[36] Because then the average action reveals the fundamental, as it aggregates the information dispersed across informed agents.

Proposition 1.3. *Suppose private information is free ($c = 0$). The policymaker can provide almost perfectly precise public information in equilibrium and achieve welfare losses arbitrarily close to zero.*

Conversely, suppose private information is costly ($c > 0$). There is a lower bound on the welfare loss in any equilibrium:

$$L_W \geq \frac{2}{\sqrt{\frac{\tau_x(4+\tau_x c)}{c}} - \tau_x}. \tag{1.19}$$

The policymaker achieves the minimum welfare loss by disclosing the fundamental with precision:

$$\tau_z^* = \left[\sqrt{\frac{\tau_x(4+\tau_x c)}{c}} - \tau_x\right]/2 - \tau_\theta. \tag{1.20}$$

If the policymaker is sufficiently transparent ($t_z > t_z^*$), there is no equilibrium.

Sketch of proof: For the case when private information is free ($c = 0$), there is no equilibrium in the second step of the game if the central bank is perfectly transparent. The argument is similar to the one in Remark 1.2 which proves that there is no equilibrium if the average action is observed without any noise.[37] Nonetheless, the central bank can always decrease the expected welfare loss by increasing its degree of transparency, and it can achieve welfare losses arbitrarily close to zero by revealing the fundamental with almost perfect precision.

For the case when private information is costly ($c > 0$), note that:

- If $\tau_z < \tau_z^*$, all agents acquire the private signal in equilibrium. The policymaker perfectly learns the fundamental and the welfare loss is strictly decreasing in τ_z

- If $\tau_z = \tau_z^*$, each agent is indifferent between acquiring the private signal and not acquiring it. In equilibrium, any fraction $\lambda^* \in (0, 1]$ of agents become informed and the policymaker perfectly learns the fundamental. More specifically, the equilibrium is:

[37] More precisely, if other informed agents respond to their private signals (i.e. $\phi_1^{I*} \neq 0$), each informed agent finds it optimal to not respond to his private signal (because the policymaker learns the fundamental and the public signal is perfectly informative). If informed agents do not respond to their private signals (i.e. $\phi_1^{I*} = 0$), then each informed agent finds it optimal to respond to his private signal (because the public signal is uninformative).

- $\lambda^* \in (0,1]$
- $\phi_1^{I*} = \frac{\tau_x}{\tau_x+\tau_\theta+\tau_z^*}$
- $\phi_2^{I*} = \frac{\tau_z^*}{\tau_x+\tau_\theta+\tau_z^*}$
- $\phi^{U*} = \frac{\tau_z^*}{\tau_\theta+\tau_z^*}$

- If $\tau_z > \tau_z^*$,[38] then there is no equilibrium in the first step of the game — this can be argued by contradiction:
  - Suppose a fraction $\lambda^* > 0$ of agents become informed. Then, there is a unique equilibrium in the second step of the game (in which the average action reveals the fundamental). Because the policymaker learns the fundamental and discloses it with precision $\tau_z > \tau_z^*$, the expected loss of uninformed agents is lower than the expected loss of informed agents, so no agent would want to acquire the private signal in the first step of the game (thus $\lambda^* = 0$), hence a contradiction.
  - Suppose no agent becomes informed ($\lambda^* = 0$). Then, there is a unique equilibrium in the second step of the game in which $a_i = 0 \forall i$. But then the expected loss of informed agents is lower than the expected loss of uninformed agents,[39] so each agent would want to acquire the private signal in the first step of the game ($\lambda^* > 0$), hence a contradiction. □

Remark 1.4: *The impossibility result above*[40] *emerges because the informativeness of public information (and hence $\gamma(\lambda)$) are discontinuous in λ at zero (which is not the case when the policymaker observes the average action with noise, or when agents directly observe the average action with noise).*

Intuitively, the impossibility result above — whereby there is no equilibrium in the first step of the game for $\tau_z > \tau_z^*$ — emerges because if any fraction of agents become informed, then the informativeness of the policymaker's disclosure will be high enough to discourage private information acquisition in the first place.

This is reminiscent of Bernanke and Woodford (1997) where the central bank's policy responds to private sector agents' inflation forecasts. In a similar fashion, they argue that if the forecasters' information is free, then the central bank can "make the variance of inflation arbitrarily close to its minimum value", whereas if

[38]But τ_z is finite, such that the policymaker is not perfectly transparent (otherwise, there is no equilibrium in the second step of the game).

[39]Because we assumed that $c < \frac{\tau_x}{\tau_\theta(\tau_x+\tau_\theta)}$.

[40]Referring to the non-existence of equilibrium when the policymaker is too transparent.

the forecasters' information is costly, "there is a limit to [...] the degree to which the variability of inflation can be reduced, without eliminating the incentive of the forecasters to gather information" (Bernanke and Woodford (1997)). Proposition 1.3 provides a similar result in a setting where private information is noisy and dispersed, and where the central bank discloses information instead of setting policy.[41]

1.4 Conclusion

This paper analyses the social value of public information in the context of a static herding model (in the spirit of Bru and Vives (2002)) featuring costly acquisition of private information. It argues that there is a threshold level of public information, whereby below the threshold all agents acquire private information, and more precise public information is welfare-improving, whereas above the threshold only a fraction of agents acquire the private signal, and more precise public information is no longer welfare-improving. In other words, if the public signal of the average prediction conveys information from informed to uninformed agents, then increasing its precision will no longer be socially beneficial. I relate the result to Grossman and Stiglitz (1980) and discuss its implications in the context of the literature on endogenous central bank information.

[41] Note that we could easily extend the model by allowing the central bank to control a payoff-relevant policy instrument.

2 Optimal Monetary Policy and the Signal Value of Prices under Rational Inattention

2.1 Introduction

By conducting monetary policy, the central bank influences market prices in line with its objectives. Concurrently, market prices reveal information to the central bank about the shifting fundamentals of the economy. The first point should be uncontroversial, given that most central banks have an explicit inflation target. The second point has a long tradition within economics, tracing back to Hayek (1945), who put forward the idea that prices play an important role as aggregators of dispersed information.

Hence, by conveying information regarding fundamentals, market prices may inform the setting of optimal monetary policy, and a feedback loop may arise whereby prices and monetary policy mutually influence each other. This feedback loop has been called the "reflection problem" in deference to Samuelson (1994):

"Monetary policy relies on market prices, and yet monetary policy influences market prices. This two-way flow introduces a potential channel of circularity whereby market outcomes reflect central bank actions, which in turn reflect market outcomes. Paul Samuelson (1994) famously compared this potential circularity with the reactions of a monkey seeing its reflection in the mirror for the first time. The monkey reacts to its own reflection in the mirror, unaware that it is seeing its own reflection." (Morris and Shin (2018))

I analyse this feedback loop and its implications for optimal monetary policy in a setup where price-setters are rationally inattentive, in the spirit of Sims (2003). I study a setting where the central bank extracts information from market prices, while simultaneously influencing them via the conduct of its policy. Firms can only process a finite amount of information, so prices are set under imperfect common knowledge — as is already known this gives rise to real effects of nominal shocks (Woodford (2001)) and provides scope for monetary policy intervention (Adam (2007)). From the perspective of firms, monetary policy is part of the environment, so changes in the central bank's reaction function affect what firms pay attention to, which in turn has implications for what information is revealed by their prices in equilibrium.

Hence the central bank's information is endogenous, in the sense that it depends on its policy reaction function (which shapes the equilibrium allocation of attention and the responsiveness of prices to shocks).

I show that the optimal monetary policy reaction function minimizes the information revealed by prices.[1] Because the central bank learns from prices, this implies that the optimal reaction function must minimize the central bank's own information precision about fundamentals (the mechanism behind this will be discussed in what follows). Consequently, the central bank faces an inherent tension between stabilizing shocks to fundamentals[2] and extracting information about fundamentals from market prices.[3] I also discuss the implications of this tension for optimal policy.

The framework: Besides the informational constraints placed on the central bank (to be detailed below), I study a somewhat standard static monetary economy. There is a household that supplies labour and consumes. Production takes place within a unit mass of monopolistically competitive firms who set nominal prices before markets open. Each firm faces an idiosyncratic productivity shock. Additionally, there is a shock to the household's preferences over consumption and labour which shifts the efficient level of output in the economy — I refer to this (policy-relevant) aggregate shock as the fundamental in what follows.

Alongside the private-sector agents, there is a consolidated monetary and fiscal authority. The fiscal authority subsidises production to control for the distortion caused by monopolistic competition, while the monetary authority directly controls nominal demand. Following most of the literature, I abstract from imperfect information on the side of the household.

Information frictions: As aforementioned, firms are rationally inattentive a la Sims (2003). As such, they choose what they pay attention to subject to a constraint on information flow. Following Maćkowiak and Wiederholt (2009), I assume that paying attention to aggregate and idiosyncratic shocks are separate activities[4]

[1] This statement (and the conditions under which it is true) will be made precise in the main text (Corollary 2.1 in section 2.5).

[2] In order to maximize social welfare (defined in the main text as the expected utility of the representative household).

[3] Note that extracting information from prices is not an explicit goal of the central bank — more precise information (about the fundamental) is beneficial because it allows the central bank to better tailor monetary policy (and achieve lower welfare losses).

[4] Such that firms cannot receive any signal which is informative about both aggregate and idiosyncratic shocks.

— in our context the idiosyncratic shock is firm-specific productivity, while aggregate shocks comprise nominal demand and the fundamental (i.e. the shock to the efficient level of output). Besides the fixed capacity to process information and the independence assumption in the spirit of Maćkowiak and Wiederholt (2009), the firms' acquisition of information is flexible, in the sense that firms have access to an unrestricted set of available signals.

The central bank is fully attentive, but I assume that it cannot directly observe the efficient level of output (which will be relevant for the optimal setting of the monetary instrument). Instead, the policymaker learns about the fundamental by observing the market outcome. More specifically, I follow Baeriswyl et al. (2020) in assuming that the central bank's information source is a noisy signal of the price level. We will also note that observing a noisy signal of real output is informationally equivalent to observing a noisy signal of the price level.[5]

Remark 2.1. *Both the policymaker's and the firms' information is imperfect and endogenous, but for different reasons. The policymaker's information is imperfect because he cannot directly observe the fundamental (and his only information source is a noisy signal of the price level); his information is endogenous because the policy reaction function affects price informativeness.[6] The firms' information is imperfect because they only have a finite capacity to process information; their information is endogenous because it depends on what they choose to pay attention to.*

These informational assumptions seek to capture some of Hayek's view of knowledge in society. The contrast between the observability of information to agents — namely, that firms (and the household) can directly observe shocks, while the central bank cannot — seeks to capture the idea that private sector agents have knowledge of the "particular circumstances of time and place", while the policymaker does not (which is of central importance in Hayek's original argument). Because firms are rationally inattentive, they make idiosyncratic errors when processing information — the presence of these errors "implies that firms' decision-makers know neither the precise values of [shocks], nor exactly what other firms know. Information processing limitations thus reflect Hayek's view that information exists only in the form of dispersed bits of incomplete and frequently contradictory knowledge" (Adam (2007)).

[5] So the policy rule will be similar to a standard Taylor rule.

[6] By price informativeness I refer to the information (concerning the efficient level of output) contained in the noisy signal of the price level. I formally define this in the main text.

Tension between stabilization and learning: This tension occurs when characterising optimal policy in the setup described above and is best understood by noting some properties of optimal policy and price informativeness:

i) Optimal policy seeks to accommodate shocks to the fundamental — in this regard, more precise central bank information (about the fundamental) is beneficial as it allows the central bank to better tailor monetary policy (and achieve lower welfare losses);[7]

ii) Changes in policy affect price informativeness if and only if they prompt firms to pay more or less attention to aggregate shocks. If firms pay more attention to aggregate shocks, price informativeness increases (and vice-versa);[8]

iii) Because optimal policy accommodates shocks to the fundamental, it must also minimize the attention paid by firms to aggregate shocks.

In light of the above, the inherent tension between stabilization and learning should hopefully be clear — from i), ii) and iii) it follows that optimal policy minimizes price informativeness (and thereby minimizes the central bank's own information precision about the fundamental). If the central bank manages to better stabilize/accommodate shocks to the fundamental, then firms pay less attention to aggregate shocks and price informativeness deteriorates — but then the central bank is more poorly informed about the fundamental and is less able to accommodate it (in this sense, the CB's ability to accommodate shocks is partly self-defeating when it learns from prices).

Policy implications: Because the firms' equilibrium allocation of attention depends on the parameters of the environment, so does price informativeness (and the central bank's information precision about the fundamental). It follows that the optimal degree of policy activism also depends on the parameters of the environment. I characterise this analytically and argue that policy should respond more strongly to perceived changes in the price level whenever firms pay more attention to aggregate shocks in equilibrium (because prices are then more informative about the fundamental). Hence, the model provides an informational rationale for increased policy activism during times of high aggregate volatility (for instance, during recessions).

[7] See Claim 2.3.

[8] See (2.33).

Furthermore, it provides a rationale for the increased flexibility of the price level — and the consequently muted ability of monetary policy to affect real output[9] — during such times. When the volatility of the aggregate fundamental is high, firms pay more attention to aggregate shocks in equilibrium, so prices respond more strongly to (both fundamental and) nominal demand shocks — hence the degree of monetary non-neutrality is lower. This second rationale is also discussed in Song and Stern (2022). Also, the theory I propose here is consistent with their empirical evidence documenting firm attention to aggregate shocks (I expand on this below).

2.2 Related literature

Endogenous central bank information: The main questions I seek to answer in this paper are:

What information do prices reveal? What exactly determines price informativeness? If the central bank plays a role in this and it learns about the fundamental by observing prices, how should optimal policy account for the endogeneity of the central bank's information?

These are common questions in the literature on endogenous central bank information. Bernanke and Woodford (1997) discuss the existence and uniqueness of rational expectations equilibria when monetary policy responds to inflation forecasts (which are endogenous to policy). In Aoki (2003) the central bank extracts information about a policy-relevant state from noisy indicators of inflation and output, while influencing these noisy indicators via its policy actions. Nimark (2008) discusses a setting where the central bank learns about the state of the economy from bond market prices (which are influenced by the setting of the short-term nominal interest rate). In Bond and Goldstein (2015), the government extracts information about a firm's fundamental from its stock market price (which is endogenous to policy intervention). Morris and Shin (2005) study how public information released by the central bank affects price informativeness and, in turn, how this influences the central bank's and the firms' information precision about fundamentals over time. Baeriswyl (2011) analyses the optimal degree of transparency of central bank communication in a static (micro-founded) setup where the CB's information source is a noisy signal of the price level. Baeriswyl et al. (2020) expand on Baeriswyl (2011) and analyse the effect of policy intervention in a similar setting — this is the most

[9]Equivalently, a higher degree of monetary neutrality.

closely related paper to the model I present here.[10]

A shared characteristic of the models mentioned above is that private-sector agents are exogenously informed (in the sense that they take their information as given)[11] — hence policy intervention affects price informativeness by changing the way in which the private-sector agents' information about fundamentals maps into market prices. Because in my model firms are rationally inattentive, they choose what information to process and what information to neglect (so the private-sector agents' information is endogenous and reacts to changes in policy). In other words, in a rational inattention setting there is also a mapping from the environment to the private-sector agents' information structure — I thus build on this literature by proposing a novel mechanism whereby changes in policy prompt changes in price informativeness by affecting the signal structure of (or the information processed by) private-sector agents,[12] and by analysing the implications of this mechanism for optimal policy.

Remark 2.2. *In contrast, consider an example where the central bank's information is imperfect but exogenous. Iovino et al. (2021) emphasize the importance of modeling incomplete central bank information,[13] so they assume that the central bank observes a noisy, private signal of the state (TFP) each period. The point stressed by the endogenous central bank information literature is that such a (direct) signal may not exist[14] (i.e. TFP may not be directly observable and may need to be estimated from other observable data which may be influenced by the central bank's actions): "it is not possible to directly observe a productivity shock. Instead, economists estimate productivity shocks with models that link output to capital used and hours worked. These inputs are themselves the results of market interactions, which are influenced by, among other things, the conduct of monetary policy" (Baeriswyl et al. (2020)). "Monetary policy operates in an uncertain environment where some state variables are only observed with error and delay and some variables, like productivity and thus potential output, are not observed at all. Variables that are not observable but relevant for monetary policy have to be inferred from variables that are observ-*

[10] The main difference is that firms here are rationally inattentive, whereas in their paper the firms' information is exogenous. Also, there are no idiosyncratic shocks in their model.

[11] The only exception being Bernanke and Woodford (1997) who consider the possibility that the forecaster's private information is costly. The way in which information is acquired by firms here is substantially different.

[12] Alongside affecting the response of private-sector agents to their information.

[13] On the grounds of realism, and because this provides a coherent microfoundation for monetary policy shocks (Iovino et al. (2021)).

[14] Or that the precision of the (indirect) signal is endogenous to policy.

able" (Nimark (2008)).

Optimal monetary policy and rational inattention: Conceptually, the model relates to papers studying optimal monetary policy when price setters have imperfect information (see for instance Adam (2007), Ball et al. (2005), Paciello and Wiederholt (2013), Iovino et al. (2021) among others). The microfoundations of the model are closely related to Adam (2007)[15]. The main difference is that the central bank in his model is perfectly informed, while the central bank here is imperfectly informed and learns from prices. I also extend the model of Adam (2007) by introducing idiosyncratic productivity shocks and argue that this modelling assumption has important implications.[16] The firms' problem is similar to that in Maćkowiak and Wiederholt (2009).[17]

Note that the central bank's information is exogenous in the models mentioned above. As far as I know, the model I present here is the only one where firms are rationally inattentive and where the central bank learns by observing the market outcome — the paper thus provides a link between the literature on optimal monetary policy when firms are rationally inattentive and the literature on endogenous central bank information.

More recent contributions (within the macro rational inattention literature) provide both conceptual and empirical support to the theory I propose here.

Afrouzi and Yang (2021) develop an attention-driven theory of pricing where the slope of the Phillips curve and the degree of monetary non-neutrality respond endogenously to the conduct of monetary policy (and they argue that more hawkish policy flattens the Phillips curve and leads to higher monetary non-neutrality). I present a similar result in a static setting[18] and argue that price informativeness, the degree of monetary non-neutrality, and the slope of the Phillips curve are all innately linked.[19]

Song and Stern (2022) develop a text-based measure of firm attention to macroeconomic news and provide evidence that firm attention is countercyclical, in the sense

[15]As well as Paciello and Wiederholt (2013).

[16]See section 2.4.

[17]In fact, exactly equivalent following Proposition 2.1.

[18]In my framework stabilization policy also flattens the Phillips curve and increases the degree of monetary non-neutrality — see section 2.4.1.

[19]See Remark 2.4.

that the average amount of attention paid by firms to aggregate shocks is higher during recessions. They argue that this is the case because firms face higher aggregate uncertainty during recessions (and thus find it optimal to pay more attention to macroeconomic variables). They also propose this as a mechanism behind the state dependency of monetary policy as documented by Tenreyro and Thwaites (2016) who estimate weaker responses of real output to monetary policy in recessions than in expansions. A similar mechanism occurs in the model I present here — furthermore, the optimal degree of policy activism is also state-dependent for the same rationale (because firms pay more attention to aggregate shocks in recessions, prices are more informative so the central bank finds it optimal to respond more strongly to perceived changes in the price level).

Paper contribution: The main contribution of the paper is that it provides a link between the literature on endogenous central bank information and the literature studying optimal monetary policy when firms are rationally inattentive.

In doing so, it also resolves an issue arising in Baeriswyl et al. (2020). In their model, if the setting of the monetary policy instrument is observable (which they believe to be "realistic"), changes in the central bank's reaction function have no implications for either welfare or price informativeness — see section 4 in their paper.[20] The model I present here effectively builds on their signaling action framework, but relaxes the assumption that the firms' information is exogenous (instead assuming that firms are rationally inattentive). It is also worth noting that besides taking a stance on why there is imperfect information in the economy, I do not make any different assumptions regarding the observability of information relative to Baeriswyl et al. (2020) — more specifically, I also assume that relevant fundamentals are observable to firms (and the household), but not to the central bank.

Although firms can also observe the setting of the policy instrument here, they cannot pay perfect attention to it (thus policy does not have the same signaling role as in Baeriswyl et al. (2020), section 4). Also, changes in policy prompt changes

[20] "Whereas the central bank finds it optimal to take an action without disclosing any information [. . .], one may wonder whether it is realistic to keep an action secret from the public. This section considers a more realistic operational framework where taking an action signals what the central bank believes about the state of the economy [. . .]" Baeriswyl et al. (2020). The assumption that the setting of the monetary instrument is observable to firms also concords with Maćkowiak and Wiederholt (2009) who "think that [in] a convincing model of real effects of monetary policy due to imperfect information [. . .] information concerning the current state of monetary policy must be publicly available".

in the firms' allocation of attention — hence changes in the CB's reaction function have implications for both welfare and price informativeness here. Furthermore, the mechanism I discuss — whereby changes in policy affect price informativeness by influencing the firms' allocation of attention — is broadly consistent with insights from the macro rational inattention literature, although the model I present here is static and stylised.

Outline: The rest of the paper is organized as follows. In section 2.3, I introduce the framework. In section 2.4, I analyse the role of policy intervention in a setting where the central bank's information is exogenous (i.e. the central bank does not learn from prices); then, I define price informativeness and study its determinants. In section 2.5 I analyse optimal policy in a setting where the central bank's information is endogenous (i.e. the central bank learns from prices). Section 2.6 discusses extensions and section 2.7 concludes. All proofs are collected in section 2.8.

2.3 Model setup

The economy is populated by a representative household, a unit mass of firms, a fiscal authority and a central bank.

Household: The household's preferences over consumption and labour are given by:

$$U(C, L) = \frac{C^{1-\gamma} - 1}{1 - \gamma} - V\frac{L^{1+\psi}}{1 + \psi} \tag{2.1}$$

Where L denotes the household's labour supply, the parameters γ and ψ regulate the curvature of the utility function with respect to consumption and labour, and V is a stochastic labour supply shock, whose log is mean-zero normally distributed $ln(V) \sim N(0, \sigma_v^2)$. Variations in V will shift the efficient level of output in the economy.

C is a composite good defined by the Dixit-Stiglitz aggregator:

$$C = (\int_0^1 C_i^{\frac{1}{1+\Lambda}} di)^{1+\Lambda} \tag{2.2}$$

Where C_i denotes consumption of good i and $(1 + \frac{1}{\Lambda})$ is the elasticity of substitution between different goods. The representative household's budget constraint is given by:

$$\int_0^1 P_i C_i = WL + \Pi - T \tag{2.3}$$

Where P_i is the price of good i, W is the wage rate, Π are aggregate profits made by the firms and T is a tax levied by the fiscal authority. Denote by P the price index which solves $PC = \int_0^1 P_i C_i di$, i.e.:

$$P = (\int_0^1 P_i^{-\frac{1}{\Lambda}} di)^{-\Lambda} \tag{2.4}$$

The household knows V and it takes as given the prices of all goods, as well as the wage rate, the lump sum tax and the firms' profits. It chooses a consumption vector (specifying how much it consumes of each variety) and how much labour to supply, in order to maximize utility (2.1) subject to the budget constraint (2.3).

Firms: Firm i produces good i according to the production function:

$$Y_i = A_i L_i^{\alpha} \tag{2.5}$$

Where Y_i is output, L_i is labour input and A_i is firm-specific productivity. The log of each firm's productivity is independently distributed according to $ln(A_i) \sim N(0, \sigma_A^2)$.

Each firm sets its price and commits to producing any quantity demanded at that particular price. Its nominal profits are given by:

$$\Pi_i = (1 + t_s) P_i Y_i - W L_i \tag{2.6}$$

Where t_s is a per unit production subsidy paid by the fiscal authority. Each firm sets its price to maximize expected profit conditional on its information set — this is presented in more detail in what follows.

Central bank: The central bank directly controls nominal spending:

$$M = PC \tag{2.7}$$

Where M denotes the central bank's policy instrument (which should be interpreted as nominal demand or the money supply).[21] The central bank (henceforth, CB) is imperfectly informed (about V) and commits to a policy rule specifying how it sets the money supply conditional on its private information — this will be described in

[21] Note that this can be microfounded by introducing a (binding) cash-in-advance constraint for the household.

more detail in what follows. The CB will choose the policy rule which maximizes the expected utility of the household.

Fiscal authority: The sole purpose of the fiscal authority is to control for the distortion caused by monopolistic competition. By taxing the household and appropriately subsidising production (setting $t_s = \Lambda$), it ensures that under perfect information firms do not charge a markup above marginal cost,[22] such that the equilibrium is first-best. This is standard in the literature (see, for instance, Adam (2007), or Paciello and Wiederholt (2013)). In this model, Λ is constant, so there are no mark-up shocks.[23]

Deterministic equilibrium: In what follows we will work with an approximation of the model around a deterministic equilibrium (in which $M = 1, V = 1$ and $A_i = 1 \forall i$). More specifically, we will work with log-quadratic approximations of the agents' payoff functions and log-linear approximations of the equilibrium conditions. Note that this is common practice in the literature on optimal monetary policy — see, Adam (2007), Ball et al. (2005), Paciello and Wiederholt (2013) among others.

In the deterministic equilibrium, all firms set the same price $\overline{P_i} = \overline{P} = \frac{1}{\overline{C}}$, and the household consumes all goods in equal amounts:

$$\overline{C_i} = \overline{C} = \alpha^{1/(\gamma + \frac{1+\psi}{\alpha} - 1)} \tag{2.8}$$

For reference, this is derived in Appendix 2.8.1. In what follows, lower-case Latin letters denote variables expressed in terms of log-deviations from the deterministic equilibrium — i.e. $c = ln\left(\frac{C}{\overline{C}}\right), p_i = ln\left(\frac{P_i}{\overline{P}}\right), p = ln\left(\frac{P}{\overline{P}}\right)$ etc. I next derive log-quadratic approximations of the CB's and the firm's objective functions around this deterministic equilibrium (and describe the information structure).

Welfare (CB's objective): Our welfare criterion will be the household's expected utility — as such, we think of the optimal allocation as the one which maximizes the household's utility given the economy's technological constraints. Noting that at any feasible allocation the household needs to supply the labour required to produce

[22]Under imperfect information, it implies that firms set prices equal to their expectation of marginal cost.

[23]In the language of Paciello and Wiederholt (2013), this will imply that there are no shocks which cause inefficient fluctuations under perfect information.

the consumption goods:

$$L = \int_0^1 \left(\frac{C_i}{A_i}\right)^{(1/\alpha)} di \tag{2.9}$$

We can define the optimal allocation as the one which solves:

$$\max_{C_i \forall i} \frac{C^{1-\gamma}-1}{1-\gamma} - \frac{V}{1+\psi}\left[\int_0^1 \left(\frac{C_i}{A_i}\right)^{(1/\alpha)} di\right]^{1+\psi} \tag{2.10}$$

Maximizing the expression above yields that optimal consumption of the composite good depends on the realization of V:

$$C^* = \left(\frac{\alpha}{V}\right)^{1/(\gamma+\frac{1+\psi}{\alpha}-1)} x \underbrace{constant}_{\approx 1 (up\ to\ FOA)} \tag{2.11}$$

While the optimal consumption mix depends on the realization of firm-specific productivity shocks:

$$\hat{C}_I^* := \frac{C_i^*}{C^*} = \left[\frac{A_i^{1/(1+\Lambda-\alpha)}}{\int_0^1 A_i^{1/(1+\Lambda-\alpha)} di}\right]^{1+\Lambda} \tag{2.12}$$

Aggregate output and consumption should be higher whenever the disutility of labour (V) is lower.[24] Also, the household should optimally consume more of the goods produced by firms with above-average productivity (as they require relatively less labour in production).

Note that consumption is efficient in the deterministic equilibrium. Expressing (2.11) and (2.12) in terms of log-deviations from the deterministic equilibrium yields:

$$c^* = -\frac{1}{\gamma + \frac{1+\psi}{\alpha} - 1} v \tag{2.13}$$

$$\hat{c_i}^* = \frac{1+\Lambda}{1+\Lambda-\alpha} a_i \tag{2.14}$$

Our measure of welfare is a log-quadratic approximation of the utility of the household around the deterministic equilibrium. I show in appendix 2.8.1 that the welfare loss is increasing in the distance between the optimal consumption vector (characterised by c^* and $\hat{c_i}^*$ for all i) and the actual consumption vector — more precisely, utility at any feasible allocation $(c, \hat{c_i})$ can be approximated as:

[24]This is the same as in Adam (2007) or Baeriswyl et al. (2020).

$$\widetilde{U}(c,\hat{c}_i) = \widetilde{U}(c^*,\hat{c}_i{}^*) - \eta(c-c^*)^2 - \zeta\int_0^1 (\hat{c}_i - \hat{c}_i{}^*)^2 di \tag{2.15}$$

$$\text{Where } \eta = \frac{1}{2}\overline{C}^{1-\gamma}\left(\frac{1+\psi}{\alpha} - 1 + \gamma\right), \zeta = \frac{1}{2}\overline{C}^{1-\gamma}\left[\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)}\right]$$

And where $\widetilde{U}(c^*,\hat{c}_i{}^*)$ denotes (approximate) utility evaluated at the optimal consumption vector. The first negative term captures losses in utility due to a suboptimal consumption level (i.e. the household consumes too much or too little of the composite good, relative to the optimum), while the second term captures losses in utility due to a suboptimal consumption mix (among the different goods produced by firms). As previously mentioned, the optimal consumption level depends on the realization of v, while the optimal consumption mix depends on the realization of the $a_i's$ (as per equations (2.13) and (2.14)).

Central bank's information: I first solve the model in a setup where the CB's information is exogenous (i.e. the CB does not learn from prices) — in this setting, I analyse optimal monetary policy and its implications for price informativeness. After characterising the equilibrium under exogenous CB information, I analyse the reflection problem in section 2.5.

As such, for now I assume that the CB learns about the efficient level of output by observing an exogenous signal $x = c^* + \varepsilon_x$, where $\varepsilon_x \sim N(0,\sigma_x^2)$ is independent of all other random variables. The CB commits to a policy rule which maps signal realizations to the money supply:

$$m = \rho x \tag{2.16}$$

Where ρ denotes the policy coefficient, which is chosen by the CB to maximize expected welfare. Allowing the CB to commit to a policy rule such as (2.16) is common practice in the literature — see, for instance, Adam (2007), James and Lawler (2011), Baeriswyl et al. (2020).

Firms' objective: It is standard to show[25] that each firm's problem of choosing

[25] See Appendix 2.8.1.

its price (p_i) to maximize expected profits is equivalent to:

$$\begin{aligned} &\min_{p_i} \frac{1}{2} E\left[(p_i - \hat{p}_i)^2 \middle| S_i \right] \\ \text{Where: } &\hat{p}_i = r(m - c^*) + (1-r)p - z_i \\ &p = \int_0^1 p_i di \\ &z_i = \phi a_i \end{aligned} \tag{2.17}$$

And where $r = \left(\gamma - 1 + \frac{\psi+1}{\alpha}\right) / \left(\frac{1+\Lambda-\alpha}{\alpha\Lambda}\right)$ and $\phi = \frac{\Lambda}{1+\Lambda-\alpha}$.

Maximizing a log-quadratic approximation of the profit function around the deterministic equilibrium (with respect to the price p_i) is equivalent to minimizing the objective in (2.17). S_i denotes firm i's information set (to be introduced in the next subsection).

Each firm incurs quadratic profit losses whenever its price (p_i) differs from the profit-maximizing price ($\hat{p}_i$), which depends on nominal demand (m), the fundamental (c^*), the aggregate price level (p) and the firm's productivity (a_i). The sensitivity of firm i's optimal price to nominal demand and the fundamental is captured by r, while the sensitivity of firm i's optimal price to its productivity is given by ϕ. Noting that $p = \int_0^1 p_i di$, we can also interpret $(1 - r)$ as the degree of strategic complementarity in price-setting. In line with the literature, I assume $r \in (0, 1)$, such that prices are strategic complements.

Firms' information: Firstly, let us address what is observable to firms when setting prices. In this regard, I depart from the macro literature on rational inattention[26] in assuming that the price level is not observable to firms. I only make this departure to facilitate comparison with the literature on the signal value of prices, where firms do not observe the price level (or each other's prices).[27] Relaxing this (i.e. allowing firms to observe the price level) does not affect the results in any way.

[26]For instance, Maćkowiak and Wiederholt (2009), Paciello and Wiederholt (2013).

[27]In Baeriswyl (2011), Baeriswyl et al. (2020), only the policymaker gets to learn by observing a noisy signal of the price level — in fact, if the noisy signal of the price level were to be made public, the central bank would not have any (private) information which is not already known by firms, so there would be no role for the policymaker in the model (he could not reduce the welfare loss). Note that a similar issue arises in Morris and Shin (2005). This is not the case here, as firms are rationally inattentive (and it turns out to be irrelevant whether we assume the price level (or a noisy signal of it) to be observable or unobservable to firms).

With this in mind, when choosing its price, each firm can observe the realization of its productivity (a_i), the efficient level of output (c^*), as well as the setting of policy (m). We will analyse price-setting (conditional on an information structure) as a beauty contest game in the spirit of Morris and Shin (2002).

Information processing: As previously mentioned, firms are rationally inattentive a la Sims (2003). As such, we treat them as finite capacity information channels which can only process a certain amount of information. Information processing is quantified as reduction in uncertainty, where uncertainty is measured by entropy.[28]

For the case of Gaussian random variables, entropy has a simple representation. If Q is an n-dimensional random vector $Q = (Q_1, Q_2, ..., Q_n)^T$ whose distribution is multivariate normal with covariance matrix Ω_Q, its entropy is given by

$$H(Q) = \frac{1}{2} log_2[(2\pi e)^n det\Omega_Q]$$

And only depends on the number of random variables and their covariance matrix. Similarly, conditional uncertainty is measured by conditional entropy — if the vector $Z = (Z_1, Z_2, ..., Z_m)^T$ also has a multivariate normal distribution, the conditional entropy of Q given Z is:

$$H(Q|Z) = \frac{1}{2} log_2[(2\pi e)^n det\Omega_{Q|Z}]$$

Where $\Omega_{Q|Z}$ denotes the conditional covariance matrix of Q given Z.

As such, reduction in uncertainty is measured as the difference between unconditional and conditional entropy — this is also known as the mutual information between Q and Z:

$$I(Q, Z) = H(Q) - H(Q|Z)$$

The information processing constraint then places an upper bound on the mutual information between the vector of variables which the decision-maker is interested in learning about, and the vector of signals that the agent receives (to learn about the variables of interest).

[28]Formally, the entropy of a continuous random variable X with density function $p(x)$ is defined as: $H(X) = -\int_{-\infty}^{\infty} p(x) log_2(p(x))dx$, where the convention is to take $p(x)log(p(x)) = 0$ when $p(x) = 0$.

Consider firm i and suppose that firms process information before setting prices[29] — the random variables which firm i is interested in learning about are: the efficient level of output (c^*), the money supply (m) and its firm-specific productivity (a_i).[30] Define the vector $X_i = (c^*, m, a_i)^T$ and suppose that firm i's information set is generated by observing a signal vector $S_i = (s_{i,1}, s_{i,2}, ..., s_{i,K})^T$. The information processing constraint restricts the mutual information between X_i and S_i:

$$I(X_i, S_i) \leq \kappa \Leftrightarrow H(X_i) - H(X_i|S_i) \leq \kappa$$

Where $\kappa > 0$ denotes the firms' information processing capacity (assumed to be the same for all i). Intuitively, the information processing constraint places a limit on how much uncertainty about the random vector X_i can be reduced through the observation of the signal vector S_i. If $\kappa = 0$, firms cannot receive any extra information about the variables of interest, so their posteriors about c^*, m and a_i are the same as the priors. If $\kappa \to \infty$, firms can perfectly learn the realizations of c^*, m and a_i. For positive, finite κ (which I assume here), firms imperfectly learn about the realizations of the variables of interest.

In a similar fashion to Maćkowiak et al. (2018), I assume that each signal in the signal vector S_i can be about a different linear combination of shocks. Further, I impose some additional conditions following Maćkowiak and Wiederholt (2009):

1) All signals received by firms are Gaussian;

2) All noise in signals is idiosyncratic;

3) The signal vector S_i can be partitioned into one subvector that only contains information about aggregate shocks and another subvector that only contains information about idiosyncratic shocks, i.e. we can write $S_i = (S_i^A, S_i^I)^T$, where $\{S_i^A, m, c^*\}$ and $\{S_i^I, a_i\}$ are independent (this is referred to as the independence assumption in what follows).

[29] Such that the price level is not observable to firms, as previously discussed.

[30] Technically, this is only true for $\rho \neq 0$. If the CB does not intervene (i.e. if it sets $\rho = 0$), the money supply is deterministic, so firms only track the efficient level of output (c^*) and firm-specific productivity (a_i).

Note that all conditions above are static counterparts of the assumptions in Maćkowiak and Wiederholt (2009). Condition 1 is just for ease of exposition — Gaussian signals are optimal because the firms' loss function is quadratic and the variables being tracked are also Gaussian.[31] Condition 2 reflects the idea that the information friction here is the firms' limited attention rather than the availability of information. Condition 3 captures the idea that paying attention to aggregate shocks and paying attention to idiosyncratic shocks are separate activities. This is not without loss of generality — the signal structure chosen by firms in equilibrium in the absence of the independence assumption will be different (I relax this assumption and discuss its implications in section 2.6).

Each firm chooses the properties of its signal vector S_i subject to the information processing constraint and the conditions above. As each signal can be about a different linear combination of shocks, we can represent the firm's signal vector as:

$$S_i = F_i X_i + \varepsilon_i$$

Where F_i is a K_i x 3 matrix of coefficients and $\varepsilon_i = (\varepsilon_{1,i}, \varepsilon_{2,i}, ..., \varepsilon_{K,i})^T$ is a Gaussian white noise random vector (independent of X_i) with covariance matrix $\Omega_{\varepsilon i}$. As previously mentioned, $X_i = (c^*, m, a_i)^T$. Choosing the properties of the signal vector S_i entails specifying the number of signals to observe (K_i), the content of these signals (F_i) and the covariance matrix of noise in the signals ($\Omega_{\varepsilon i}$).

Taking into account that each firm sets its price optimally conditional on the information it receives (i.e. it chooses p_i to solve (2.17)), we can write firm i's problem when choosing its allocation of attention as:

$$\begin{aligned} &\min_{K_i, F_i, \Omega_{\varepsilon i}} \frac{1}{2} var[r(m-c^*) + (1-r)p - z_i|S_i] \\ \text{Subject to: } & S_i = F_i X_i + \varepsilon_i, \\ & \varepsilon_i \sim N(0, \Omega_{\varepsilon i}), \\ & I(S_i, X_i) \le \kappa, \\ & \text{Independence assumption.} \end{aligned}$$

For a more precise formulation of the firm's information choice problem see Appendix 2.8.2.[32] When choosing its allocation of attention, each firm takes as given

[31] See Maćkowiak and Wiederholt (2007), technical appendix for a proof.

[32] As will become clear, the signal vectors chosen in equilibrium have a simple representation (see Proposition 2.1), so I leave some details regarding the determination of the equilibrium allocation

the CB's reaction function and the other firms' allocations of attention. It anticipates how it will optimally set its price conditional on the information it receives[33] and takes into account how its decision regarding the allocation of attention will affect its profit losses.

Timing:

1. CB chooses its reaction function (ρ)
2. Firms choose their allocations of attention ($K_i, F_i, \Omega_{\varepsilon i}$)
3. Shocks (v and the a_i's) are realized
4. CB receives its information (x) and sets policy (m) according to the policy rule ($m = \rho x$)
5. Each firm receives its information (S_i) and sets its price (p_i)

Equilibrium definition: An equilibrium specifies: the CB's reaction function (ρ^*), an allocation of attention for each firm ($K_i^*, F_i^*, \Omega_{\varepsilon i}^*$), and a pricing rule for each firm (mapping signal realizations to the firm's price), such that ρ^* maximizes expected welfare, each firm's price is equal to its conditional expectation of the profit-maximizing price ($p_i^* = E[\hat{p}_i|S_i]$) and each firm's allocation of attention minimizes its expected profit losses subject to the information processing constraint and the independence assumption. As previously mentioned, when choosing its allocation of attention, each firm takes as given the other firms' allocations of attention and the CB's reaction function, and it anticipates how prices will be set in equilibrium (in step 5) — this allows us to get an expression for firm i's expected profit losses as a function of its allocation of attention (in step 2). Following the literature, I analyse symmetric equilibria in which all firms choose the same allocation of attention:

$$(K_i^*, F_i^*, \Omega_{\varepsilon i}^*) = (K^*, F^*, \Omega_{\varepsilon}^*) \forall i.$$

Equilibrium content of signals: Before proceeding to derive the equilibrium, it is useful to simplify the analysis regarding the equilibrium allocation of attention. Generally, if a rationally inattentive agent with a quadratic loss function tracks a Gaussian optimal action, it is without loss of generality to consider signals of the

of attention to the Appendix.

[33]It also anticipates how other firms will set prices (conditional on their information) in equilibrium. We will also note that given a symmetric allocation of attention for all other firms $j \neq i$, there is a unique price-setting equilibrium, so each individual firm i can compute the price level as a linear combination of m and c^* when choosing its allocation of attention.

form "optimal action plus noise". Given that firms are constrained to learn about aggregate and idiosyncratic shocks separately (by the independence assumption), it should be without loss of generality to restrict our attention to two-dimensional signal vectors (as in Maćkowiak and Wiederholt (2009)) — one signal of the form "optimal response of price to aggregate shocks plus noise" and one signal of the form "optimal response of price to firm-specific shock plus noise". The content of the latter is equal to z_i in (2.17). The content of the former is more complicated and depends on the price level, which is an endogenous variable. In this respect, let us denote by $q = m - c^*$ the equilibrium response of the price level to aggregate shocks under perfect information — observing a signal of the form "q plus noise" is equivalent (in any symmetric equilibrium) to observing a signal of the form "optimal response of price to aggregate shocks plus noise".[34]

Proposition 2.1. *Without loss of generality, we can restrict attention to equilibria in which each firm observes two private signals: one about the composite aggregate shock $q = m - c^*$ ($s_{i,1} = q + \varepsilon_{i,1} = m - c^* + \varepsilon_{i,1}$) and one about its idiosyncratic shock ($s_{i,2} = z_i + \varepsilon_{i,2}$), where $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$ are drawn independently for each i from the distributions $N \sim (0, \sigma_1^2)$ and $N \sim (0, \sigma_2^2)$ respectively.*[35]

Proof: See Appendix 2.8.2. □

Proposition 2.1 is useful because it simplifies the analysis regarding the equilibrium allocation of attention — now we only have to look for the variance of noise in private signals (σ_1^2 and σ_2^2) which is consistent with equilibrium. In turn, this is equivalent to determining how much attention firms pay to aggregate and idiosyncratic shocks in equilibrium — given Proposition 2.1, the information processing constraint rewrites as:

$$\underbrace{I[(m, c^*)', S_i^A]}_{\kappa_A} + \underbrace{I[z_i, S_i^I]}_{\kappa_I} \leq \kappa \iff \underbrace{\frac{1}{2} log_2(1 + \frac{\sigma_q^2}{\sigma_1^2})}_{\kappa_A} + \underbrace{\frac{1}{2} log_2(1 + \frac{\sigma_z^2}{\sigma_2^2})}_{\kappa_I} \leq \kappa \tag{2.18}$$

Where κ_A and κ_I denote how much attention firm i pays to aggregate and idiosyn-

[34]Because in equilibrium the price level will be proportional to q — a similar observation is made in Maćkowiak and Wiederholt (2009), where it is noted that observing a signal of nominal demand, or a signal of a linear combination of nominal demand and the price level is equivalent. Note that the problem here is more complicated because there are two aggregate shocks (to both nominal demand and the efficient level of output).

[35]i.e. we only need to analyse equilibria in which: $K = 2, F = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & \phi \end{pmatrix}, \Omega_\varepsilon = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$.

cratic shocks respectively.[36] Clearly, there is a one-to-one mapping between σ_1^2 and k_A (as well as between σ_2^2 and k_I) — in what follows, I present results in terms of the amount of attention that firms pay to aggregate and idiosyncratic shocks in equilibrium (rather than the variance of the noise terms in private signals).[37]

Some parts of the proposition are straightforward — for instance, that we can restrict attention to signal vectors featuring only one signal concerning the idiosyncratic shock. If the firm were to observe multiple signals about z_i, these would be equivalent[38] to a single signal with a higher precision (the same goes for observing multiple signals about $m - c^*$).[39]

The equilibrium content of signals containing information about aggregate shocks is less trivial because firms track two correlated aggregate variables (the efficient level of output (c^*) and the money supply (m))[40] and because of the strategic complementarity in price-setting. I show that in any symmetric equilibrium firms only learn about the linear combination $m - c^*$ (i.e. in equilibrium, firms do not observe any signals concerning aggregate shocks of the form $s_i = d_1 m + d_2 c^* + \varepsilon_i$ with $d_1 \neq -d_2$).

Remark 2.3. *Following Proposition 2.1, to fully characterise the equilibrium allocation of attention, it suffices to find the pair (κ_A, κ_I) which is consistent with equilibrium. More generally, the approach to determining the equilibrium allocation of attention is the following: first, note that for any (κ_A, κ_I), it is without loss of generality to restrict attention to information structures as described in Proposition 2.1; then (later on, in Claim 2.2), find the (κ_A, κ_I) which is consistent with equilibrium.*

[36] I have dropped i subscripts because the allocation of attention is symmetric.

[37] To facilitate comparison of my results with Maćkowiak and Wiederholt (2009) — following Proposition 2.1, the firms' problem here turns out to be equivalent to the one in their paper.

[38] i.e. would use up the same amount of processing capacity and lead to the same posterior uncertainty about z_i.

[39] Thus, the number of signals received by firms in equilibrium is indeterminate. However, all equilibria are equivalent, in the sense that firms face the same posterior uncertainty about shocks in any equilibrium, and all equilibria are associated with the same distribution of prices across firms — hence, the "without loss of generality" part at the beginning of the proposition.

[40] For any $\rho \neq 0$. For $\rho = 0$, i.e. if the CB does not change the money supply (relative to the deterministic equilibrium), firms only track the efficient level of output.

2.4 Equilibrium when central bank information is exogenous

In this section, I derive the equilibrium behaviour of firms and optimal monetary policy when the CB's information is exogenous. Then, I analyse the determinants of price informativeness. Putting everything together, we can express the CB's problem as:

$$\min_{\rho} E[(c-c^*)^2 + \delta \int_0^1 (p_i - p + z_i)^2 di] \tag{2.19}$$

Subject to:

$$c = m - p \tag{2.20}$$

$$m = \rho x \tag{2.21}$$

$$p = \int_0^1 p_i di \tag{2.22}$$

$$p_i = E[\hat{p}_i | S_i] \tag{2.23}$$

$$\hat{p}_i = r(m - c^*) + (1-r)p - z_i \tag{2.24}$$

$$S_i = \begin{pmatrix} m - c^* + \varepsilon_{i,1} \\ z_i + \varepsilon_{i,2} \end{pmatrix} \tag{2.25}$$

$$\Omega_{\varepsilon i} = \begin{pmatrix} \sigma_1^* & 0 \\ 0 & \sigma_2^* \end{pmatrix} \tag{2.26}$$

$$(\sigma_1^*, \sigma_2^*) = \arg\max_{\sigma_1, \sigma_2} \frac{1}{2} E\left[(p_i - \hat{p}_i)^2\right] \text{ subject to (2.20)-(2.26) and (2.18)} \tag{2.27}$$

Where:

$$c^* = -\frac{1}{\gamma + \frac{1+\psi}{\alpha} - 1} v \tag{2.28}$$

$$z_i = \frac{\Lambda}{1 + \Lambda - \alpha} a_i \tag{2.29}$$

$$\delta = \frac{1}{\gamma + \frac{1+\psi}{\alpha} - 1} \left[\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)}\right] \left(\frac{1+\Lambda}{\Lambda}\right)^2 > 0 \tag{2.30}$$

Substituting the log-linearized demand function for good i into (2.15) and using (2.22) gives an objective function equivalent to (2.19). (2.20) follows from the definition of nominal demand. (2.21) is the CB's policy rule. (2.22) follows from log-linearizing the price index. (2.23) states that each firm sets its price equal to its conditional expectation of the profit-maximizing price, which is characterised in (2.24). (2.25) and (2.26) characterise firm i's signal vector following Proposition 2.1, and (2.27) states that each firm chooses its allocation of attention optimally to minimize expected profit losses.

I derive the equilibrium by backward induction. Hence, I first characterise the firms' price-setting behaviour conditional on the CB's policy (ρ) and a symmetric allocation of attention (Claim 2.1).[41] Then, in Claim 2.2, I determine the equilibrium allocation of attention given the CB's policy ρ.[42] Lastly, in Claim 2.3, I characterise optimal policy.[43]

Claim 2.1. (Equilibrium price-setting)

Fix any ρ and any symmetric allocation of attention ($\kappa_{i,A} = \kappa_A$ and $\kappa_{i,I} = \kappa_I \forall i$). In equilibrium, prices are set according to:

$$p_i^* = \tfrac{r\beta_1}{1-(1-r)\beta_1} s_{i,1} - \beta_2 s_{i,2} \forall i$$

Where $\beta_1 = 1 - 2^{-2\kappa_A}$ and $\beta_2 = 1 - 2^{-2\kappa_I}$ denote the Kalman-gains of signals $s_{i,1}$ and $s_{i,2}$ respectively (i.e. $E[q|S_i] = \beta_1 s_{i,1}$ and $E[z_i|S_i] = \beta_2 s_{i,2}$).

Proof. See Appendix 2.8.3.[44] □

Price-setting conditional on a symmetric allocation of attention is equivalent to a beauty contest game with an exogenous, Gaussian information structure, similar to Morris and Shin (2002). This guarantees the existence and uniqueness of the linear equilibrium in Claim 2.1. It also explains why the equilibrium weight given to the private signal $s_{i,1}$ is lower than the signal's Kalman-gain for any $r \in [0, 1)$:

$$\frac{r\beta_1}{1-(1-r)\beta_1} < \beta_1$$

For smaller values of r (i.e. a higher degree of strategic complementarity in price-setting), higher-order beliefs about q increasingly influence firm i's optimal price, as firms attach greater weight to the beliefs about other firms' beliefs about q, in-

[41] Step 5 in the timing of events.
[42] Step 2 in the timing of events.
[43] Step 1 in the timing of events.
[44] Given Proposition 2.1, the argument is standard in the literature — it follows Morris and Shin (2002), section D, and is virtually equivalent to the one in Adam (2007), section 5.

creasing the relative response to public information (i.e. the prior mean of q, which is equal to zero and common knowledge among firms) and decreasing the relative response to private information (signals $s_{i,1}$). In turn, this lowers the responsiveness of the price level to aggregate shocks and amplifies the real effects of nominal shocks due to imperfect information.[45]

Claim 2.2. (Equilibrium allocation of attention)

Fix any policy reaction function ρ.[46] The firms' equilibrium allocation of attention is given by:

$$\kappa_A^* = \begin{cases} \kappa & if\ \frac{r\sigma_q}{\sigma_z} \geq 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa}) \\ \frac{1}{2}log_2(1 - \frac{1}{r} + \frac{2^{\kappa}\sigma_q}{\sigma_z}) & if\ \frac{r\sigma_q}{\sigma_z} \in (2^{-\kappa}, 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa})) \\ 0 & if\ \frac{r\sigma_q}{\sigma_z} \leq 2^{-\kappa} \end{cases}$$

$$\kappa_I^* = \kappa - \kappa_A^*$$

Proof. See Appendix 2.8.3. □

The equilibrium above is equivalent to the one in Maćkowiak and Wiederholt (2009). As such, the intuition behind it and its interpretation are familiar. In equilibrium, firms pay more attention to shocks which are more volatile. Increasing the variance of the idiosyncratic shock (σ_z^2) relative to the variance of the aggregate shock (σ_q^2) will lead firms to pay more attention to the idiosyncratic shock (and vice-versa). Consequently, prices will also respond more strongly to the idiosyncratic shock and less strongly to the aggregate shock. Note that the variance of the aggregate shock here depends on the setting of policy — in particular, if the CB stabilizes shocks and reduces the variance of aggregate shocks to optimal prices (σ_q^2), the firms' allocation of attention will shift towards idiosyncratic shocks.

The degree of strategic complementarity in price-setting is also relevant in determining the equilibrium allocation of attention, as it induces strategic complementarity in information acquisition — if all firms pay more attention to the aggregate shock q, the price level will respond more strongly to it, so each individual firm will find it

[45] This is standard — see Woodford (2001), Adam (2007), Maćkowiak and Wiederholt (2009) among others.

[46] And note that this pins down the variance of q: $\sigma_q^2 = (\rho - 1)^2\sigma_{c^*}^2 + \rho^2\sigma_x^2$.

optimal to pay more attention to q.[47] Thus, pinning down the equilibrium amounts to solving a fixed point problem — at the fixed point, the amount of attention allocated to aggregate shocks is (weakly) decreasing in the degree of strategic complementarity.[48]

Claim 2.3. (Optimal policy)

Optimal policy seeks to accommodate shocks to the efficient level of output — i.e. it minimizes the variance of the aggregate shock (σ_q^2) conditional on the CB's information:

$$\rho^* = \frac{\sigma_{c^*}^2}{\sigma_{c^*}^2 + \sigma_x^2}$$

In equilibrium, the variance of the aggregate shock is given by:

$$\sigma_q^{2*} = \frac{\sigma_{c^*}^2 \sigma_x^2}{\sigma_{c^*}^2 + \sigma_x^2}$$

Proof: See Appendix 2.8.3. □

Welfare losses occur solely due to the firms' limited capacity to process information. By stabilizing aggregate shocks, the central bank simplifies the firms' tracking problem — in turn, this leads prices to better align with shocks and consumption to be more efficient. Accommodating shocks to the efficient level of output is somewhat standard in the literature on optimal monetary policy, in settings of both exogenous[49] and endogenous[50] firms' information. Nonetheless, it is interesting to note here that the mechanism via which accommodation policy increases expected welfare — i.e. whether it improves the consumption mix or reduces the variance of the output gap — depends on whether firms pay attention to just aggregate or idiosyncratic shocks, or both.

If firms only pay attention to aggregate shocks ($\kappa_A^* = \kappa$), (small) changes in policy influence both the consumption mix and the variance of the real output gap. For instance, stabilization policy lowers the variance of firms' information processing errors — this both decreases inefficient price dispersion across firms (thereby improving the consumption mix) and lowers the variance of the output gap (the

[47] Maćkowiak and Wiederholt (2009) refer to these as "feedback effects".

[48] More precisely, strictly decreasing at an interior equilibrium allocation of attention. Equivalently, (weakly) increasing in r.

[49] Baeriswyl et al. (2020).

[50] Adam (2007), Paciello and Wiederholt (2013).

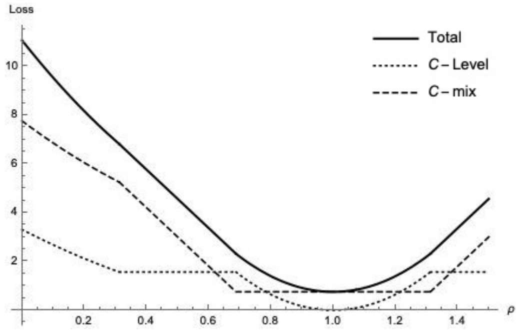


Figure 2.1: Welfare loss plotted as a function of the CB's reaction function ρ (in the presence of idiosyncratic shocks, $\sigma_z = 1$)

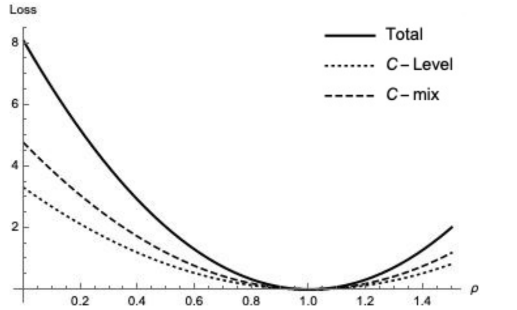


Figure 2.2: Welfare loss plotted as a function of the CB's reaction function ρ (in the absence of idiosyncratic shocks, $\sigma_z = 0$)

mechanism is similar to Adam (2007)). On the other hand, if firms only pay attention to idiosyncratic conditions ($\kappa_A^* = 0$), (small) changes in policy only influence the variance of the output gap.

If firms pay attention to both aggregate and idiosyncratic shocks ($\kappa_A^* \in (0, \kappa)$), (small) changes in policy affect welfare only by influencing inefficient price dispersion. Although accommodation policy lowers the variance of the aggregate shock (and, holding everything else unchanged, this should lead to a fall in the variance of the real output gap), it prompts firms to pay less attention to aggregate conditions — this fall in attention works in the opposite direction (to increase the variance of the output gap) and perfectly offsets the former effect (such that the variance of the output gap is constant for any $\kappa_A \in (0, \kappa)$).

Importantly, note from Claim 2.3 that optimal policy depends on the CB's information precision regarding the efficient level of output c^*. The CB intervenes more (and is able to achieve a lower welfare loss) if its information is more precise. In the limiting case when the CB is perfectly informed, the fundamental shock is perfectly accommodated, so the variance of the aggregate shock q (σ_q^2) is zero, and firms only pay attention to their idiosyncratic productivity shocks. This is illustrated in Figure 2.1, for a parametrisation where the CB's information is almost perfectly precise (so $\rho^* \approx 1$).[51]

For comparison, Figure 2.2 illustrates a situation where there are no idiosyncratic shocks (in which case the model corresponds almost exactly to the one in Adam

[51] More precisely, the parametrisation is chosen such that in the absence of policy intervention firms pay attention only to aggregate shocks (to illustrate the different ways in which policy affects welfare). The calibration is: $\kappa = 1, \sigma_{c^*}^2 = 4, \sigma_z^2 = 1, r = 0.4, \delta = 3, \sigma_x = 0.001$.

(2007)) — optimal policy is the same in both settings, but the mechanism via which policy intervention affects the welfare loss is different.

2.4.1 Policy intervention, monetary non-neutrality and the Phillips curve

Interestingly, for the economy without idiosyncratic shocks, the degree of monetary non-neutrality and the slope of the Phillips curve are constants (which depend on parameters only), while in the presence of idiosyncratic shocks, changes in the CB's policy rule may affect both the degree of monetary non-neutrality in the economy, as well as the slope of the Phillips curve. In the latter case, this is possible because changes in policy may prompt firms to pay more or less attention to aggregate shocks (see Claim 2.2) and, in turn, this affects how prices respond to aggregate shocks (see Claim 2.1).

To be more precise, from the firms' equilibrium behaviour,[52] it follows that the price level can be expressed as $p = \alpha(m - c^*)$, where:

$$\alpha = \begin{cases} \alpha_H = \frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})} & \text{if } \frac{r\sigma_q}{\sigma_z} \geq 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa}) \\ \alpha_M = 1 - \frac{\sigma_z}{2^{\kappa} r \sigma_q} & \text{if } \frac{r\sigma_q}{\sigma_z} \in (2^{-\kappa}, 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa})) \\ \alpha_L = 0 & \text{if } \frac{r\sigma_q}{\sigma_z} \leq 2^{-\kappa} \end{cases} \tag{2.31}$$

Thus, the responsiveness of the price level to a nominal demand shock m is captured by α, so the degree of monetary non-neutrality is $(1 - \alpha)$. The slope of the Phillips curve — relating the price level[53] to the real output gap $(c - c^*)$ — is given by $\alpha/(1 - \alpha)$.[54] Intuitively, α depends on how strongly firms respond to their private signals about aggregate conditions when setting their prices — in turn, this depends on the degree of strategic complementarity in price-setting and on how much attention firms pay to aggregate shocks (see Claim 2.1).

Hence, if there are no idiosyncratic shocks ($\sigma_z = 0$), firms pay attention solely to aggregate shocks ($\kappa_A^* = \kappa$) and α is always equal to α_H. Note that this is inde-

[52]Equation (2.31) can be obtained by substituting κ_A^* from Claim 2.2 into the price-setting equation from Claim 2.1 and aggregating across firms. Note that it corresponds to equation (38) in Maćkowiak and Wiederholt (2009).

[53]More precisely, the deviation of the price level from the deterministic equilibrium.

[54]Using $p = \alpha(m - c^*)$ and the definition of nominal demand $m = c + p$, one obtains $p = [\alpha/(1 - \alpha)](c - c^*)$.

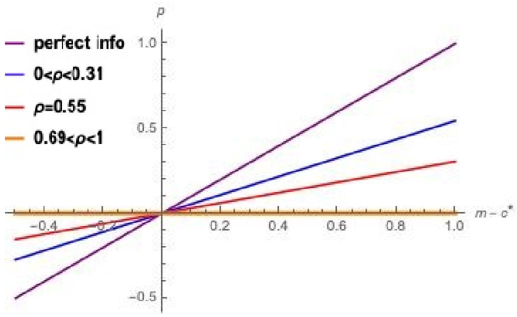


Figure 2.3: Response of price level to aggregate shocks (plotted for different policy reaction functions ρ)

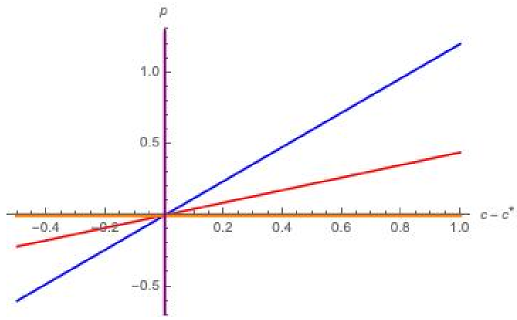


Figure 2.4: Phillips curve (plotted for different policy reaction functions ρ)

pendent of the CB's policy (ρ) or the variance of the aggregate fundamental ($\sigma^2_{c^*}$).[55] Consequently, in an economy without idiosyncratic shocks, the degree of monetary non-neutrality and the slope of the Phillips curve depend only on the degree of strategic complementarity in price-setting and on the firms' capacity to process information.

On the other hand, in the presence of idiosyncratic shocks, α also depends on the variance of aggregate and idiosyncratic shocks, as these affect the equilibrium allocation of attention. In particular, note that stabilization policy (by reducing the variance of aggregate shocks) may decrease the responsiveness of the price level to aggregate shocks and flatten the Phillips curve. Figures 2.3 and 2.4 plot the response of the price level to aggregate shocks and the Phillips curve (for different policy rules) in the example from before (where the CB is almost perfectly informed).[56] At the optimal policy, the variance of aggregate shocks is (almost) zero so firms only pay attention to idiosyncratic shocks. Thus, the price level never responds to aggregate shocks and the Phillips curve is perfectly flat. If there are no idiosyncratic shocks (σ_z=0), the two blue lines do not rotate in response to changes in policy.[57]

2.4.2 The signal value of prices

Having characterised optimal policy and the equilibrium behaviour of firms, we are in a position to assess what information is revealed by prices (and how policy inter-

[55] This happens because the signal-to-noise ratio associated with firms' private signals (as well as the Kalman-gain of signals $s_{i,1}$) depends only on how much attention firms pay to q (and not on the variance of q which is indeed influenced by policy). For instance, if the CB reduces the variance of the aggregate shock (σ^2_q), the variance of noise in private signals (σ^2_1) decreases proportionally, so firms respond the same way to their private signals when setting prices.

[56] See footnote 51 for the parametrisation.

[57] They represent the response of the price level to aggregate shocks and the Phillips curve when firms pay attention solely to aggregate shocks.

vention affects price informativeness). In this respect, it is useful to define exactly what is meant by price informativeness. In line with the literature on the signal value of prices (Morris and Shin (2005); Baeriswyl (2011); Baeriswyl et al. (2020)), in the endogenous central bank information setup (next section), I assume that the CB's only information source is a noisy signal of the aggregate price level $\tilde{p} = p + \tilde{\varepsilon}$, where $\tilde{\varepsilon} \sim N(0, \sigma_{\tilde{p}}^2)$ is distributed independently of all other random variables. Thus, by price informativeness, I refer to the incremental information precision that the CB gains from observing the signal $\tilde{p}$:

Definition. *Price informativeness (denoted τ_p) is defined as the precision of the unbiased signal of the fundamental (referring to the efficient level of output (c^*)) which the CB can construct from $\tilde{p}$. Equivalently, price informativeness captures the difference between the CB's posterior and prior information precision about c^* (where prior and posterior refer to before and after observing the signal $\tilde{p}$).*[58]

It might seem arbitrary that price informativeness is defined in terms of what the CB can infer from $\tilde{p}$. However, we will note that this is equivalent to the incremental information precision that an external observer (who knows m and is fully attentive) gains from observing $\tilde{p}$.[59]

The responsiveness of the price level to shocks concerning the efficient level of output is crucial in determining what the CB can infer from $\tilde{p}$. It follows from before that we can express $\tilde{p}$ as:

$$\tilde{p} = \alpha(m - c^*) + \tilde{\varepsilon}$$

Where α is characaterised in equation (2.31). If firms only pay attention to firm-specific productivity ($\kappa_A^* = 0$) the price level does not respond to aggregate shocks ($\alpha = 0$), so $\tilde{p}$ is completely uninformative (it reveals only the noise term $\tilde{\varepsilon}$). However, if firms pay some attention to aggregate conditions ($\kappa_A^* > 0$), $\tilde{p}$ reveals information about both policy and the efficient level of output. As the CB knows m, it can construct an unbiased signal of c^*:

$$\tilde{c} = m - \frac{\tilde{p}}{\alpha} = c^* + \frac{\tilde{\varepsilon}}{\alpha} \tag{2.32}$$

[58] i.e. $\tau_p = [var(c^*|I_{CB}, \tilde{p})]^{-1} - [var(c^*|I_{CB})]^{-1}$, where I_{CB} denotes the CB's information set before observing $\tilde{p}$.

[59] Note that the setting of policy is observable to firms. If we assume that the setting of policy is also observable to an external observer (who does not otherwise have any private information about c^* but) who is fully attentive, he perfectly infers the CB's private information about c^* by observing m (because there is a one-to-one mapping between x and m for any $\rho \neq 0$).

With precision $\tau_p = \alpha^2/\sigma_{\tilde{p}}^2$. Equivalently, price informativeness is given by:

$$\tau_p = \begin{cases} \alpha_H^2/\sigma_{\tilde{p}}^2 = \left[\frac{r(1-2^{-2k})}{1-(1-r)(1-2^{-2k})}\right]^2/\sigma_{\tilde{p}}^2 & \text{if } \frac{r\sigma_q}{\sigma_z} \geq 2^{-k} + r(2^k - 2^{-k}) \\ \alpha_M^2/\sigma_{\tilde{p}}^2 = \left[1 - \frac{\sigma_z}{2^k r\sigma_q}\right]^2/\sigma_{\tilde{p}}^2 & \text{if } \frac{r\sigma_q}{\sigma_z} \in (2^{-k}, 2^{-k} + r(2^k - 2^{-k})) \\ 0 & \text{if } \frac{r\sigma_q}{\sigma_z} \leq 2^{-k} \end{cases} \tag{2.33}$$

Note that besides the observation noise ($\sigma_{\tilde{p}}^2$), price informativeness depends only on α. Because the CB can perfectly disentangle the money supply from $\tilde{p}$, changes in policy (ρ) are not associated with changes in price informativeness so long as firms do not reallocate attention from aggregate to idiosyncratic shocks or vice-versa (such that α stays constant).

Remark 2.4: *It should be clear that price informativeness, the slope of the Phillips curve and the degree of monetary non-neutrality are all determined by α. A lower α leads to:*

- *Lower price informativeness*
- *A flatter Phillips curve*
- *A higher degree of monetary non-neutrality*

Determinants of price informativeness: The observation that price informativeness depends on the degree of strategic complementarity in price-setting and on the firms' information precision regarding fundamentals is somewhat standard in the literature on the signal value of prices: *"[. . .] the central bank's information precision is a function of the private sector agents' information precision. This is very natural, since the central bank learns by observing what the individual agents do. The reason why [private agents' information precision] enters in this relation is because the aggregate actions are revealing only to the extent that private agents put weight on their own private signals. The more informative are their private signals, the greater is the information value of the aggregate action. In this sense, the central bank's information value is dependent on (and is derivative of) the private sector agents' information precision"* (Morris and Shin (2005)). The novelty in the model presented here is that the private sector agents' information is endogenous[60] (because it depends on their chosen allocations of attention).

[60]Referring to both the content and precision of firms' signals.

In line with our analysis from the previous section (regarding α), if firms pay attention solely to aggregate shocks, price informativeness depends only on the degree of strategic complementarity $(1-r)$ and the firms' capacity to process information (κ) — a higher κ or a higher r both prompt firms to respond more strongly to their private information, thereby improving price informativeness. Note that in this case changes in the CB's policy rule (ρ) do not affect τ_p — when firms pay a fixed amount of attention to aggregate conditions, changes in policy do not prompt firms to react more or less strongly to their private information (and therefore do not affect price informativeness).

If firms pay attention to both aggregate and idiosyncratic shocks, price informativeness depends not only on κ and $(1-r)$, but also on the variance of aggregate and idiosyncratic shocks (σ_q^2 and σ_z^2) — thus, it also depends on the setting of policy (as this determines σ_q^2). A higher σ_q^2, a lower σ_z^2, or a higher κ, all lead firms to pay more attention to aggregate shocks (and prices to be more informative, as they respond more strongly to q). Increasing the degree of strategic complementarity (lowering r) has a dual detrimental effect on price informativeness — it both prompts firms to pay less attention to aggregate shocks (as per Claim 2.2), and to put a lower weight on their private signals concerning aggregate shocks when setting prices (for any given allocation of attention, as per Claim 2.1). Figure 2.6 illustrates the effect of changes in the degree of strategic complementarity on price informativeness — the dotted line corresponds to a setting where there are no idiosyncratic shocks ($\sigma_z = 0$), so only the latter effect is at play;[61] the solid line corresponds to a situation where there are both aggregate and idiosyncratic shocks, so changes in r also lead to changes in the equilibrium allocation of attention.[62]
Figure 2.5 illustrates the effect of policy intervention on price informativeness in the example from the previous section (in which the CB is almost perfectly informed). In line with the discussion above, policy influences price informativeness only to the extent that it leads firms to pay more or less attention to aggregate shocks. Note that optimal policy ($\rho^* \approx 1$) prompts firms to not pay any attention to aggregate conditions, implying that prices reveal no information regarding the efficient level of output (c^*).

[61] i.e. increasing the degree of strategic complementarity lowers price informativeness only by decreasing the firms' response to their private signals (while the precision of the signals remains constant).

[62] Figure 2.6 is plotted for a different calibration chosen such that when $r = 1$ firms pay attention solely to aggregate shocks (for the solid line $\sigma_z = 1$); for the dotted line $\sigma_z = 0$ so firms only pay attention to aggregate shocks (for any r). The other parameters are: $\kappa = 1, \sigma_q = 2, \sigma_{\tilde{p}} = 1$.

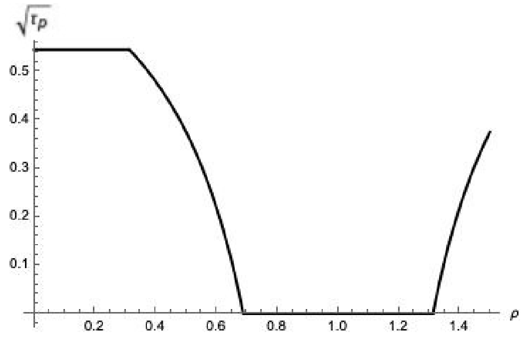


Figure 2.5: Square root of price informativeness plotted as a function of the CB's reaction function (ρ)

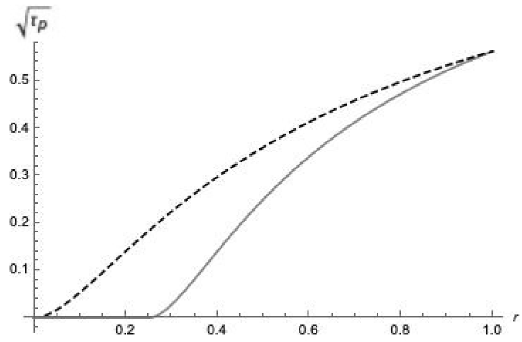


Figure 2.6: Strategic complementarities and price informativeness

2.4.3 Exogenous vs endogenous firms' information

Before proceeding to analyse the reflection problem, let us illustrate in more detail how the endogeneity of the firms' information plays an important role in our problem. This also allows us to compare the mechanism via which policy intervention deteriorates price informativeness here with the one in Baeriswyl et al. (2020), section 3.

To do so, we will consider our example from the previous sections, in which the CB is almost perfectly informed.[63] Start from a situation in which the CB does not intervene ($\rho = 0$), then analyse the equilibrium behaviour of firms (and price informativeness) as the CB (gradually) changes its policy towards $\rho^* \approx 1$ in three different cases:

1. Firms do not reoptimize their information choice (exogenous information)
2. Firms reoptimize what they pay attention to but not how much attention they pay to aggregate vs idiosyncratic shocks (fixed capacity per channel)
3. Firms fully reoptimize their allocations of attention (endogenous information)

Starting point ($\rho = 0$): Given our parametrisation, in the absence of any policy intervention, firms only pay attention to aggregate shocks, so each firm observes the signal $s^0_{i,1} = c^* + \varepsilon^0_{i,1}$, with $\varepsilon^0_{i,1} \sim N(0, \sigma^2_{c^*}/(2^{2\kappa} - 1))$. Conditional on this information structure, for $\rho = 0$, the price level is proportional to c^* in equilibrium (more precisely $p = \alpha_H c^*$). This is illustrated on the L.H.S. of Figure 2.7, which represents $q = m - c^*$ and p as vectors in (c^*, ε_x) space — at the starting point ρ is zero, so q is represented by the (black) vector OQ_0 and the price level is represented

[63] See footnote 51 for the parametrisation.

by the blue vector in the direction of OQ_0 (and of length α_H). For $\rho = 0$ observing a noisy signal of the price level ($\tilde{p} = p + \tilde{\varepsilon}$) is equivalent to observing a noisy signal of the fundamental (c^*) with precision equal to:

$$\tau_p^0 = \alpha_H^2 / \sigma_{\tilde{p}}^2$$

In what follows τ_p^0 denotes price informativeness at the inaction policy ($\rho = 0$).

Case 1: I first illustrate how policy intervention crowds out the information contained in the price level in a setting where firms do not reoptimize their choice of information in response to changes in policy. In such a setting, each firm still observes the signal $s_{i,1}^0 = c^* + \varepsilon_{i,1}^0$, with $\varepsilon_{i,1}^0 \sim N(0, \sigma_{c^*}^2/(2^{2\kappa} - 1))$ for any ρ (i.e. each firm still pays attention solely to the aggregate fundamental (c^*), regardless of the setting of policy). Note that firms know the CB's reaction function (ρ), but do not know anything else concerning the setting of the monetary instrument (m) because (although m is observable) they are not paying any attention to it. More precisely, each firm's expectation of m is given by:

$$E[m|s_{i,1}^0] = E[\rho(c^* + \varepsilon_x)|s_{i,1}^0] = \rho E[c^*|s_{i,1}^0]$$

This delivers a similar equilibrium pricing rule as in Baeriswyl et al. (2020) section 3.2, where the CB's policy instrument is unobservable to firms (because the CB finds it optimal to act under full opacity) — more specifically, given this information structure, one can show that accommodation policy ($\rho \in (0, 1]$) prompts firms to underreact to their private signals $s_{i,1}^0$:

$$p_i = (\rho - 1)\frac{r(1 - 2^{-2\kappa})}{1 - (1 - r)(1 - 2^{-2\kappa})} s_{i,1}^0 = (\rho - 1)\alpha_H s_{i,1}^0$$

In turn, this implies that the price level reacts less to the aggregate fundamental (c^*), relative to a situation where there is no policy intervention:

$$p = \int_0^1 p_i di = (\rho - 1)\alpha_H c^*$$

Which in turn prompts a deterioration of price informativeness:

$$\begin{aligned} \tau_p^1 &= (\rho - 1)^2 \alpha_H^2 / \sigma_{\tilde{p}}^2 \\ \tau_p^1 &< \tau_p^0 \forall \rho \in (0, 1] \end{aligned} \tag{2.34}$$

The grey line in Figure 2.7 plots price informativeness as a function of the CB's reaction function in this case. Vectors depicting the price level associated with different policy reaction functions are illustrated by the grey arrows in the graph on the L.H.S. of Figure 2.7 — as the CB changes its policy from 0 to 1, firms increasingly underreact to their private signals and the price level responds less and less to the aggregate fundamental (so the vector depicting the price level shrinks in length and in the limit tends towards zero as $\rho \to 1$). Note that the price level never responds to the CB's observation noise (ε_x), because firms do not pay any attention to the monetary instrument (m).

Case 2: If firms do not reoptimize how much attention they pay to aggregate vs idiosyncratic shocks in response to changes in the CB's reaction function (ρ), then $\kappa_A = \kappa$ so following Proposition 2.1, each firm observes the signal $s_{i,1} = q + \varepsilon_{i,1}$, with $\varepsilon_{i,1} \sim N(0, \sigma_q^2/(2^{2\kappa} - 1))$ for any ρ — i.e. each firm pays attention solely to $q = m - c^*$ for any policy reaction function. It follows from our equilibrium analysis that in this case the price level is proportional to q (for any ρ); more specifically:

$$p = \alpha_H q = \alpha_H (m - c^*)$$

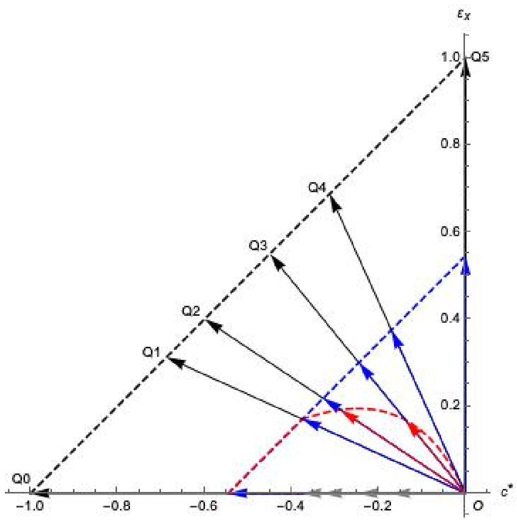
The blue vectors in the directions of $\{OQ_1; ...; OQ_5\}$ illustrate p for $\rho \in \{0.31; 0.4; 0.55; 0.69; 1\}$ respectively. Price informativeness is always equal to that at the inaction policy and is independent of the CB's reaction function ρ:

$$\tau_p^2 = \tau_p^0 = \alpha_H^2 / \sigma_{\tilde{p}}^2$$

Case 3: See the discussion in the previous section — the price level is given by (2.31) and price informativeness (denoted by τ_p^3 in Figure 2.7) is characterised in (2.33). Given our calibration, for $\rho \in (0.31, 0.69)$, the equilibrium allocation of attention is interior — the red vectors in the directions of $\{OQ_2; OQ_3\}$ depict p for $\rho \in \{0.4; 0.55\}$. For any $\rho \in [0.69; 1]$ firms pay attention only to idiosyncratic shocks, so $p = 0$. For any $\rho \in [0; 0.31]$ firms pay attention only to aggregate shocks, so $p = \alpha_H q$.

Discussion: Although in our example optimal policy intervention ($\rho^* \approx 1$) completely crowds out the information contained in the price level under both exogenous and endogenous firms' information,[64] the mechanisms at play are clearly distinct — under exogenous information, it is the firms' *under-reaction* to their private signals

[64] Referring to cases 1 and 3 above.



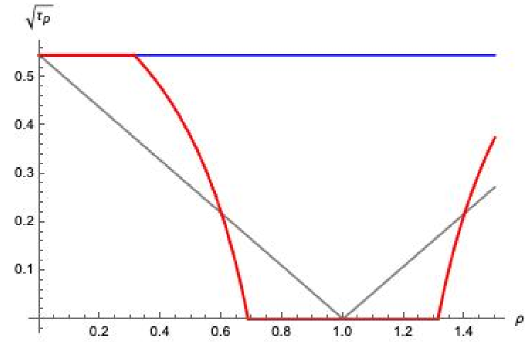


Figure 2.7: Exogenous vs endogenous firms' information and price informativeness
Grey: Exogenous information ($\sqrt{\tau_p^1}$)
Blue: Fixed Capacity ($\sqrt{\tau_p^2}$)
Red: Endogenous information ($\sqrt{\tau_p^3}$)

(about c^*) which lowers price informativeness (see (2.34)), while under endogenous information it is the firms' *under-acquisition* of information (about q), which deteriorates price informativeness.[65] To see this more clearly, note that under exogenous firms' information, price informativeness is always zero if $\rho = 1$ (regardless of parameter values), while this is not the case under endogenous firms' information — one can easily construct an example (with a different calibration) in which the CB is poorly informed and a reaction function specifying $\rho = 1$ leads prices to be more informative in equilibrium than at the inaction policy.

Note that the tension between implementing the optimal policy and learning from prices will be more pervasive under endogenous than under exogenous firms' information. Under exogenous firms' information, optimal policy intervention always *lowers* price informativeness (see Baeriswyl et al. (2020), section 3). On the other hand, under endogenous firms' information, if optimal policy leads firms to pay less attention to aggregate shocks, then optimal policy will necessarily *minimize* price informativeness (this is also stated more formally as a result in what follows).

[65] Referring to the fact that firms pay less attention to the aggregate shock q when policy is set optimally — see (2.33) and its derivation.

2.5 Equilibrium when central bank information is endogenous

In this section, I present results regarding optimal policy in a setting where the CB learns from prices, while simultaneously influencing them via the conduct of its policy. Thus, I assume here that the CB's only information source is the signal $\tilde{p} = p + \tilde{\varepsilon}$ (introduced in the previous section). As such, the CB's policy rule now reads: $m = \rho\tilde{p}$. Intuitively, given any m, a higher observation of the price level ($\tilde{p}$) signals to the CB a lower efficient level of output c^* (see (2.32)). By exactly how much the money supply should optimally respond to the noisy observation of the price level is the question I investigate here. I restrict attention to $\rho \leq 0$ and argue in Appendix 2.8.4 that the policy rule $m = \rho\tilde{p}$ is equivalent to a policy rule which maps the unbiased signal of the fundamental which the CB constructs from $\tilde{p}$ to the money supply — i.e. to a policy rule of the form: $m = \tilde{\rho}\tilde{c}$, where $\tilde{c}$ is characterised in (2.32) and where $\tilde{\rho} \in [0, 1)$.

I look for "simultaneous equilibria" where firms choose prices and the CB sets the policy instrument (m) at the same time — this is the same equilibrium concept employed in Baeriswyl (2011), Baeriswyl et al. (2020). The equilibrium definition is the same as before and the only difference with regards to the timing of events is that the CB and firms now act simultaneously.[66]

Proposition 2.2. *Suppose that in the absence of policy intervention ($\rho = 0$) the equilibrium allocation of attention is interior (i.e. $\kappa^*_{A|\rho=0} \in (0, \kappa)$).[67] Then, there is a unique equilibrium where optimal policy solves:*

$$\frac{\sigma^2_{c^*} + \rho^*\sigma^2_{\tilde{p}}}{\sqrt{\sigma^2_{c^*} + \rho^{*2}\sigma^2_{\tilde{p}}}} = \frac{\sigma_z}{2^\kappa r},$$

prices are set according to:

$$p^*_i = \frac{r(1 - 2^{-2k^*_A})}{1 - (1 - r)(1 - 2^{-2k^*_A})} s_{i,1} - (1 - 2^{-2k^*_I}) s_{i,2} \forall i,$$

[66]Namely, steps 4 and 5 in the previous timing of events take place simultaneously. More precisely, the CB still chooses ρ first; secondly, firms choose their allocations of attention; then, shocks are realized; lastly, all agents (i.e. both the CB and firms) receive their information and take actions simultaneously (firms set prices and the CB sets m).

[67]Expressed as a condition on parameters, this writes as: $\frac{r\sigma_{c^*}}{\sigma_z} \in (2^{-\kappa}, 2^{-\kappa} + r(2^\kappa - 2^{-\kappa}))$.

and the equilibrium allocation of attention is given by:

$$\kappa_A^* = \frac{1}{2} log_2(1 - \frac{1}{r} + \frac{2^k \sigma_q^*}{\sigma_z}),$$
$$\kappa_I^* = \kappa - \kappa_A^*,$$

where the standard deviation of the aggregate shock is:

$$\sigma_q^* = \frac{\sigma_{c^*}^2}{\sqrt{\sigma_{c^*}^2 + \rho^{*2}\sigma_{\tilde{p}}^2}}.$$

Corollary 2.1. *Optimal policy minimizes price informativeness.*[68]

Corollary 2.2. *Optimal policy is more activist*[69] *whenever prices are more informative in equilibrium.*

Comparative statics: $\frac{\partial|\rho^*|}{\partial\kappa} > 0$; $\frac{\partial|\rho^*|}{\partial\sigma_{c^*}^2} > 0$; $\frac{\partial|\rho^*|}{\partial\sigma_z^2} < 0$; $\frac{\partial|\rho^*|}{\partial r} > 0$; $\frac{\partial|\rho^*|}{\partial\sigma_{\tilde{p}}^2} < 0$.

Proof: See Appendix 2.8.4. □

It is easiest to interpret the equilibrium by referring to our analysis from the exogenous CB information section — the only difference here concerns the determination of the money supply, which is now set as a linear function of the CB's noisy observation of the price level.[70] Firstly, remark that for each individual firm the problem is unchanged (so the firms' equilibrium behaviour will be the same as before). A relevant observation is that although the money supply is determined differently, it is still Gaussian and observable to firms. To be more specific, each firm still takes as given the CB's reaction function (ρ) and the other firms' allocations of attention (and it anticipates how prices are set in equilibrium given the other firms' allocations of attention). Hence, each firm anticipates how the price level is determined in equilibrium, and thus how the policy instrument is set in equilibrium as a function of the CB's noisy observation of the price level — essentially, this implies that each firm can compute the optimal response of its price to aggregate shocks when

[68]Optimal policy also minimizes the slope of the Phillips curve and maximizes the degree of monetary non-neutrality — see Remark 2.4.

[69]By "policy activism" I refer to the absolute value of ρ — when $|\rho|$ is higher, policy is more activist in the sense that the CB responds more strongly to a perceived change in the price level.

[70]Which is itself a function of the money supply (whereas before the money supply was set as a linear function of the CB's noisy observation of the fundamental itself).

choosing its allocation of attention.[71] Because the firm's problem of choosing its allocation of attention is the same as before, we can use the same reasoning to prove Proposition 2.1 in this (endogenous CB information) setting. Then, Claim 2.1 and Claim 2.2 follow as before.[72] Claim 2.3 does not apply anymore because the policy instrument is set differently, but it is still relevant (in the sense that the expected welfare loss remains strictly increasing in the variance of the aggregate shock q (σ_q^2) and thus the CB still seeks to minimize it).[73,74]

Importantly, note that the equilibrium allocation of attention in Proposition 2.2 is interior (i.e. in equilibrium, firms pay attention to both aggregate and idiosyncratic shocks). We will note that the corollaries follow immediately from this observation (and our previous analysis). To see why in equilibrium (when policy is set optimally $\rho = \rho^*$) firms pay attention to both aggregate and idiosyncratic shocks, recall that in the absence of policy intervention ($\rho = 0$) the equilibrium allocation of attention is interior:

i) This implies that there are no equilibria where firms pay attention only to idiosyncratic shocks. It is easy to argue this by contradiction. Suppose there is an equilibrium where the allocation of attention is given by $(\kappa_A, \kappa_I) = (0, \kappa)$ — then, the price level is zero and it follows that $\sigma_q^2 \geq \sigma_{c^*}^2$ (for any ρ); but then each individual firm finds it optimal to pay some attention to aggregate shocks (hence, a contradiction).

ii) This also implies that, if policy is set optimally $\rho = \rho^*$, then firms cannot pay attention only to aggregate shocks. Firms would pay attention solely to aggregate conditions only if their variance (σ_q^2) was higher than the variance of the fundamen-

[71]More precisely, in equilibrium, the optimal response of firm i's price to aggregate shocks (given by $\hat{p_i}^A = r(m - c^*) + (1-r)p)$) will again be a linear combination of m and c^*. Because each firm takes as given the other firms' allocation of attention, it can compute p as a linear combination of m and c^* as before, and in any symmetric equilibrium it must still be the case that the price level is proportional to $m - c^*$, so Proposition 2.1 still holds.

[72]However, note that q (and thus σ_q^2) are determined differently to before (because the money supply is determined differently). For instance, from the perspective of firm i, the money supply now depends on the other firms' allocation of attention for any $\rho \neq 0$ (whereas in the exogenous CB information setting this was not the case) — for details, see Appendix 2.8.4

[73]Of course, this is just a consequence of the fact that the firms' equilibrium behaviour is the same as before (as is the welfare function).

[74]Also, our optimal (alternative) reaction function ($m = \tilde{\rho}^* \tilde{c}$) will be similar to the one from Claim 2.3, except for the fact that the CB's observation noise is endogenous — more precisely, the variance of the CB's observation noise is now the inverse of price informativeness (i.e. is equal to $1/\tau_p = \sigma_{\tilde{p}}^2/\alpha^2$) and is endogenous because α is endogenous (as it depends on the equilibrium allocation of attention), whereas before the CB's observation noise was just σ_x^2 (and was exogenous).

tal ($\sigma_{c^*}^2$) — but this would imply that the welfare loss in equilibrium (given optimal policy) is higher than the welfare loss in the absence of policy intervention (i.e. for $\rho = 0$), which cannot be the case (because the central bank chooses ρ to minimize the welfare loss, and it can choose $\rho = 0$).

Consequently, in equilibrium the firms' allocation of attention must necessarily be interior — for a more precise formulation of the argument, see Appendix 2.8.4. From section 2.4.1, we know that if the equilibrium allocation of attention is interior, then the responsiveness of the price level to aggregate shocks (α_M) is strictly increasing in the variance of the aggregate shock (σ_q^2), and does not otherwise depend on the CB's policy rule (see (2.31)). On the other hand, expected welfare is strictly decreasing in σ_q^2, so optimal policy must minimize σ_q^2. It follows that optimal policy minimizes α. Hence, optimal policy also minimizes price informativeness and the slope of the Phillips curve, and maximizes the degree of monetary non-neutrality (see Remark 2.4). Because the CB's sole information source is the noisy signal of the price level, it follows from Corollary 2.1 that by setting policy optimally, the CB minimizes its own information precision about the efficient level of output.

Taking into account that firms reallocate attention from aggregate to idiosyncratic shocks (or vice-versa) in response to changes in policy weakens optimal accommodation policy to shocks because it worsens the CB's information. Recall from Claim 2.3 (and the discussion following it) that the CB finds it optimal to intervene more whenever its information about the fundamental is more precise — when the allocation of attention is endogenous, the CB's information is always less precise at the optimal policy ($\rho = \rho^*$) than at the inaction policy ($\rho = 0$) (because optimal policy minimizes the amount of attention paid by firms to aggregate shocks), so the policymaker finds it optimal to accommodate less strongly shocks to the fundamental (relative to a setting where firms do not reallocate attention in response to changes in policy).

I provide a numerical example for illustrative purposes. I first provide a parsimonious calibration of the model. In line with the benchmark parametrisation of Maćkowiak and Wiederholt (2009), I let $\kappa = 3$ and $r = 0.15$. I normalize $\sigma_z = 1$ and calibrate σ_{c^*} such that in the absence of policy intervention ($\rho = 0$), firms allocate 20 percent of their information processing capacity towards aggregate shocks (and 80 percent towards idiosyncratic shocks) — this implies a value of σ_{c^*} roughly equal to one.[75,76] I calibrate $\sigma_{\tilde{p}}$ such that when policy is set optimally ($\rho = \rho^*$), firms allocate 10 percent of their information processing capacity towards aggregate shocks (and 90 percent towards idiosyncratic shocks) — this implies a value of $\sigma_{\tilde{p}} = 0.15$. Under this calibration, a shift from inaction ($\rho = 0$) to optimal policy ($\rho^* = -3.19$, or equivalently $\tilde{\rho}^* = 0.18$) lowers α from 0.16 to 0.07 (implying a flattening of the Phillips curve from 0.19 to 0.08 and a decline of price informativeness from 1.19 to 0.23 — see Figures 2.8 and 2.9). Also, note from Figure 2.9 that optimal policy ($\tilde{\rho}^* = 0.18$) minimizes price informativeness.

To make the discussion concerning optimal policy intervention more transparent, I present results in terms of the alternative (equivalent) policy rule $m = \tilde{\rho}\tilde{c}$, whereby the monetary instrument responds to the unbiased estimate of c^* constructed by the CB from $\tilde{p}$. This is easier to interpret, since the optimal reaction coefficient lies between 0 and 1 as in the exogenous CB information section (see Claim 2.3). Expressed in these terms, taking into account that firms reallocate attention in response to changes in the CB's reaction function weakens optimal accommodation policy to shocks from $\tilde{\rho}^* = 0.54$ to $\tilde{\rho}^* = 0.18$ (see Figure 2.10).

Figure 2.10 plots the welfare loss as a function of the policy reaction function ($\tilde{\rho}$). The red line depicts the welfare loss as a function of policy ($\tilde{\rho}$) if firms do not reallocate attention from aggregate to idiosyncratic shocks (or vice-versa) in response to changes in the CB's reaction function, such that they always allocate 20 percent of their attention to aggregate shocks (as in the absence of policy intervention $\tilde{\rho} = 0$). In other words, this is equivalent to a fixed capacity per channel setting where firms pay a fixed amount of attention to aggregate shocks ($\kappa_A = 0.6$) — in such a setting, price informativeness is constant at $\tau_p = 1.19$ (see the discussion in section 2.4.3 and Figure 2.9).

Note from Figure 2.10 that when the allocation of attention is endogenous, the policymaker's ability to reduce the welfare loss is significantly impaired — in fact, the expected welfare loss would have been lower in equilibrium if firms had not reallocated attention from aggregate to idiosyncratic shocks in response to changes in policy. An externality arises when the CB learns from prices because firms do not

[75] Although the variance of aggregate and idiosyncratic shocks is equal, firms allocate more attention to idiosyncratic shocks because of the strategic complementarity in price-setting and the associated "feedback effects".

[76] I focus the discussion here on price informativeness, the slope of the Phillips curve and the optimal degree of policy activism. Hence, I only provide a parsimonious calibration of the model. To calculate the welfare loss in Figure 2.10, I set $\delta = 3$.

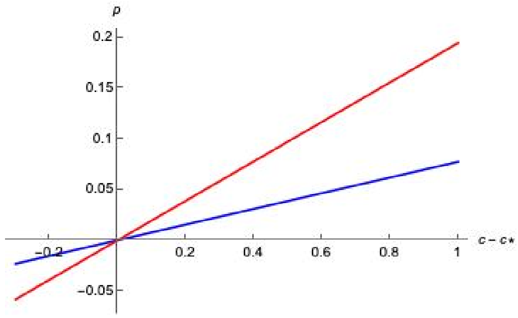


Figure 2.8: Phillips curve for different policy rules: red (inaction policy), blue (optimal policy)
.

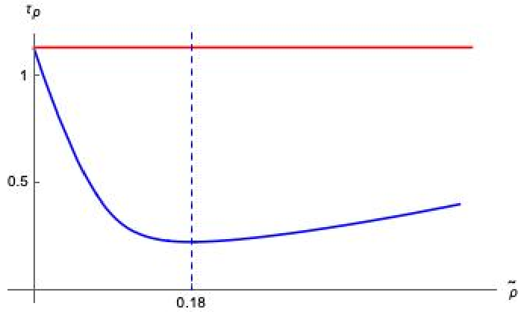


Figure 2.9: Price informativeness plotted as a function of policy ($\tilde{\rho}$) if firms always allocate 20 percent of their attention to aggregate shocks (red) vs if firms' allocation of attention is endogenous (blue)

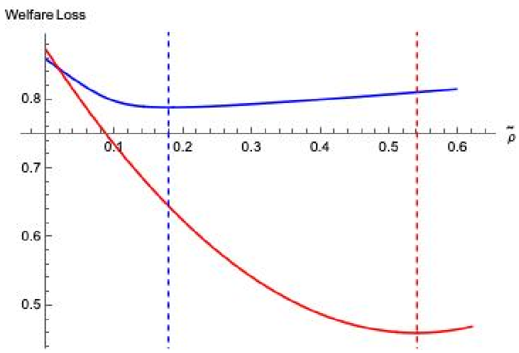


Figure 2.10: Welfare loss plotted as a function of policy ($\tilde{\rho}$) if firms always allocate 20 percent of their attention to aggregate shocks (red) vs if firms' allocation of attention is endogenous (blue)

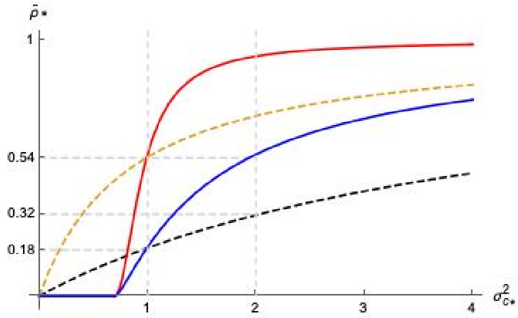


Figure 2.11: Optimal policy plotted as a function of the variance of the aggregate fundamental (σ^2_{c*}): black dotted line ($\kappa_A = 0.3$), yellow dotted line ($\kappa_A = 0.6$), blue line ($\kappa_A = \kappa^*_A$), red line ($\kappa_A = \kappa^*_{A|\rho=0}$)

internalize how their allocation of attention affects price informativeness.[77]

Because the CB learns from prices, its information (and thus, optimal policy) depends on the parameters of the environment, which shape the equilibrium allocation of attention and price informativeness. In particular, optimal policy is more activist whenever a change in parameters leads prices to be more informative in equilibrium (see Corollary 2.2) — a higher κ, a lower σ_z, or a higher r all prompt firms to pay more attention to aggregate shocks in equilibrium (and thus prices to be more informative, which in turn justifies a higher degree of policy activism).[78]

A higher σ_{c^*} increases the optimal reaction coefficient $\tilde{\rho}^*$ both because it increases price informativeness and because it worsens the CB's prior information precision concerning the fundamental.[79] To see this more clearly, consider what happens in our example if we double the variance of the aggregate fundamental ($\sigma^2_{c^*}$) — the amount of attention paid by firms to aggregate shocks in equilibrium increases to 15.3 percent and the optimal policy reaction coefficient rises back to $\tilde{\rho}^* = 0.54$. If firms still allocated only 10 percent of their attention to aggregate shocks (such that price informativeness stayed constant at $\tau_p = 0.23$), the optimal policy reaction function would specify $\tilde{\rho}^* = 0.32$.

This is depicted in Figure 2.11 which plots the optimal policy reaction coefficient ($\tilde{\rho}^*$) as a function of the variance of the aggregate fundamental ($\sigma^2_{c^*}$). In line with our analysis, the optimal degree of policy activism when the allocation of attention is endogenous (depicted in blue) is always lower relative to a situation where the amount of attention paid by firms to aggregate conditions is fixed at the inaction policy (depicted in red)[80] — this reflects the fact that the CB takes into account that policy intervention makes firms pay less attention to aggregate shocks (thus crowding out some of the information contained in the noisy signal of the price level).

[77] There are also other externalities at play (even when the CB does not learn from prices) as firms do not make socially optimal use of their private information (and do not choose the socially optimal allocation of attention).

[78] A higher r also prompts firms to respond more strongly to their private information conditional on any given allocation of attention — see the discussion in section 2.4.2.

[79] More specifically, the CB sets the policy instrument equal to its conditional expectation of the fundamental, which is a weighted sum of $\tilde{c}$ (i.e. the unbiased estimate of c^* constructed from $\tilde{p}$) and the prior mean of c^* (which is equal to zero) and where the weights correspond to the relative precision of each signal. A higher σ_{c^*} both lowers the precision of the prior and increases the precision of the signal in equilibrium.

[80] Except for any $\sigma^2_{c^*} < 0.69$ (when in the absence of policy intervention ($\rho = 0$) firms pay attention solely to idiosyncratic shocks) and the CB finds it optimal to not intervene ($\tilde{\rho}^* = 0$).

2.6 Extensions/Discussion

2.6.1 Information processing

The assumptions concerning firms' information processing play a crucial role in the results. In this section, I relax these assumptions and discuss their implications.

Firstly, consider the independence assumption (constraining firms to learn about aggregate and idiosyncratic shocks separately) — this can be either dropped or strengthened.

Dropping the independence assumption: If we drop the independence assumption, then it is straightforward to show that:

- Optimal policy should still seek to accommodate shocks to the efficient level of output
- Price informativeness is independent of the central bank's policy rule

Hence, it is easy to characterise the optimal reaction function — this is similar to the case featuring the independence assumption, but with a fixed capacity per channel.[81]

Strengthening the independence assumption: One may also consider strengthening the independence assumption (and constraining firms to learn about nominal demand and the efficient level of output separately) — more specifically, firm i's signal vector would be given by:

$$S_i = \begin{pmatrix} s_{i,1} \\ s_{i,2} \\ s_{i,3} \end{pmatrix} = \begin{pmatrix} c^* + \varepsilon_{i,1} \\ m + \varepsilon_{i,2} \\ z_i + \varepsilon_{i,3} \end{pmatrix}$$

And the firm would choose $\sigma_{i,1}^2, \sigma_{i,2}^2, \sigma_{i,3}^2$ subject to the information processing constraint. With such an information structure, the problem would be notably different.[82] For instance, even if firms paid a fixed amount of attention to aggregate shocks, a shift from inaction ($\rho = 0$) to accommodation policy ($\rho \in (0, 1)$) would deteriorate price informativeness both by: i) prompting firms to under-react to signals $s_{i,1}$ (for any fixed information structure)[83]; ii) prompting firms to shift some

[81] Or in the absence of idiosyncratic shocks ($\sigma_z = 0$).

[82] And it is not as straightforward to characterise the equilibrium as in the case where we drop the independence assumption.

[83] As in the "exogenous firms' information" example in section 2.4.3.

attention from the fundamental to the policy instrument (i.e. higher σ_1^2 than at the inaction policy).

Deriving the equilibrium with such an information structure is a notably different problem which I do not address here. If the reader is unconvinced by the model (because he believes that the stronger independence assumption should be employed), note that we can rewrite an exactly equivalent model where the aggregate fundamental is an (exogoenous) shock to nominal demand (as in Baeriswyl and Cornand (2011)) — in this case, we just need to assume that nominal demand (i.e. the sum of the fundamental and policy) is observable to firms (instead of the stronger assumption that any linear combination of the fundamental and policy is observable). The interpretation of the model would be slightly different, but the equilibrium and the main message would be the same.

Fixed capacity vs fixed marginal cost: In the model, we have assumed that firms have a fixed capacity to process information. Suppose instead that firms can choose how much attention they pay to their environment, subject to a fixed marginal cost of increasing the channel capacity (κ). In this case (assuming the independence assumption still holds), choosing how much attention to pay to aggregate vs idiosyncratic shocks are two independent decisions. One can show that in such a setting multiple equilibria can arise if the degree of strategic complementarity is sufficiently high ($r < 0.5$). Furthermore, the two equilibria feature different comparative statics (for instance, at the low-attention equilibrium, lowering the cost of information may actually prompt firms to pay less attention to aggregate shocks in equilibrium, while this is not the case at the high-attention equilibrium). Note that this occurs in the standard static version of Maćkowiak and Wiederholt (2009), if firms are allowed to choose how much attention they pay to their environment (subject to a fixed marginal cost of increasing information flow) — a similar point is made in Fulton (2017).

Thus, in the model I abstract from allowing firms to choose how much attention they pay to their environment for various reasons:

1. There are multiple equilibria featuring different comparative statics (and we would need a criterion to select among these equilibria)

2. The model does not allow for an analytical solution (of the reflection problem)

3. There is an additional complication (relative to the situation where the firms'

information processing capacity is fixed): a natural question arises whether firms' information processing costs should enter the welfare function or not[84]

2.6.2 Interpreting recessions

Throughout the paper I interpret recessions as times of high aggregate fundamental volatility (i.e. high $\sigma^2_{c^*}$) — one might object to this on various grounds. For instance, one might think of a recession as a low realization of c^*, rather than an increase in the variance of the distribution from which it is drawn. Nonetheless, remark that Song and Stern (2022) think of recessions in a similar way:[85]

"Countercyclical attention exhibited in Figure [2.12] is consistent with predictions in Mackowiak and Wiederholt (2009), which models firms that allocate attention between aggregate and idiosyncratic conditions. Their model predicts that firms will pay more attention to aggregate conditions in downturns if those conditions become more uncertain. This result is also consistent with Chiang (2021), which develops a generalized information structure where agents pay greater attention to uncertain aggregate conditions when expecting a bad economic state, which subsequently generates countercyclical attention and uncertainty."

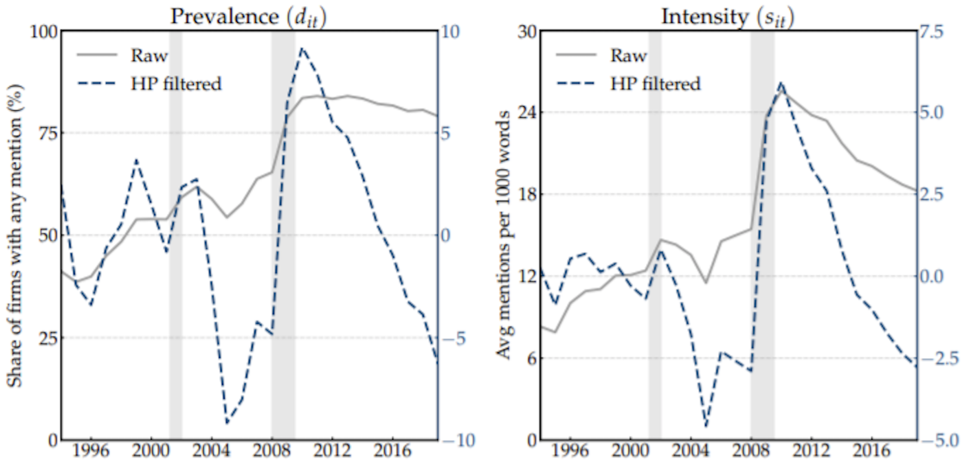
Furthermore, my interpretation is consistent with their empirical evidence that firm attention to aggregate shocks is counter-cyclical.[86] Importantly, their textual measure of firm attention seeks to distinguish between the amount of attention paid by firms to aggregate vs idiosyncratic shocks:

"We identify instances in which firms discuss the following nine macroeconomic topics: general economic conditions, output, labor market, consumption, investment, monetary policy, housing, inflation, and oil. [...] Any words or phrases that might apply to both aggregate and firm-specific conditions are removed to avoid misidentification. For example, the phrase 'interest rates' is excluded from the monetary policy dictionary because firms may mention interest rates in the context of their own liabilities."

[84] And if so, how exactly they enter the welfare function (for instance, one could assume that labour is allocated towards processing information).

[85] For more details, see section 2 in their paper.

[86] They also document that firm attention to aggregate shocks is polarized, in the sense that some firms pay more attention than others to aggregate shocks. This could be easily accommodated in the model I present here — for instance, by making firms heterogeneous in the amount of attention they pay to the environment (κ) — and the main message of the paper would not change.



Notes: Time series of firm attention to the keyword "economic conditions." The left panel plots the prevalence measure and reports the share of firms that mention the keyword. The right panel plots the intensity measure and reports the average mentions of the keyword per 1,000 words. "Raw" refers to the unfiltered series and "HP filtered" refers to the cyclical components of the HP-filtered series with smoothing factor 400. Shares are reported in percent.

Figure 2.12: "Time series of attention to 'economic conditions' " (source: Song and Stern (2022))

For a rational inattention model (of asset pricing) where recessions are conceptualised in a similar manner, see Kacperczyk et al. (2016).

2.6.3 Learning from prices

Throughout the paper, "learning from prices" refers to extracting information about the fundamental from a noisy signal of the price level. As previously mentioned, this modelling assumption follows the literature on the signal value of prices (Morris and Shin (2005); Baeriswyl (2011); Baeriswyl et al. (2020)).

Note that allowing the CB to also observe a noisy signal of real output does not affect the results in any way (because the noisy signal of real output is informationally equivalent to the noisy signal of the price level).[87] Hence, the model relates more broadly to the literature on endogenous CB information (see, for instance, Aoki (2003)).

[87]So allowing the CB to also observe an independent noisy signal of real output is equivalent to changing the variance of the CB's observation noise (σ_p^2).

2.7 Conclusion

In this paper, I analyse optimal monetary policy in a setting where the central bank learns from prices, and firms are rationally inattentive. I argue that optimal policy necessarily minimizes the attention paid by firms to aggregate shocks and that, in turn, this implies that optimal policy minimizes price informativeness. Because the CB learns from prices, it follows that by setting policy optimally, the CB must minimize its own information precision about the fundamental, so its ability to accommodate shocks is partly self-defeating. I also argue that policy should respond more strongly to perceived changes in the price level whenever firms pay more attention to aggregate shocks in equilibrium.

2.8 Proofs and complementary results

2.8.1 Derivation of deterministic equilibrium and objective functions

Deterministic equilibrium

In this section only, suppose that $M = 1, V = 1,\ A_i = 1 \forall i$ and this is common knowledge among all firms. We will present the derivation generally, then substitute these particular values at the end (because we will log-linearize the equilibrium conditions afterwards).

From the first-order conditions of the household's problem, it follows that the optimal consumption-labour choice must satisfy:

$$V L^{*\psi} C^{*\gamma} = WP \tag{2.35}$$

And the household's demand for each good is:

$$C_i^* = \left(\frac{P_i}{P}\right)^{-\frac{1+\Lambda}{\Lambda}} C^* \tag{2.36}$$

The market-clearing conditions are:

$$C_i = Y_i \forall i \tag{2.37}$$

$$L = \int_0^1 L_i di \tag{2.38}$$

Using (2.36), (2.37) and the production function, we can express firm i's profits as:

$$\Pi_i = (1 + t_s) P_i^{1-\frac{1+\Lambda}{\Lambda}} P^{\frac{1+\Lambda}{\Lambda}} C - W \left(\frac{C}{A_i}\right)^{1/\alpha} \left(\frac{P_i}{P}\right)^{\frac{1+\Lambda}{\Lambda\alpha}} \tag{2.39}$$

As mentioned in the main text, we think of price-setting in the deterministic case as a static game of complete information. Each firm chooses its price P_i to maximize Π_i, taking as given other firms' prices $P_{j\neq i}$. Note that because we are in the deterministic case, taking as given $P_{j\neq i}$ implies taking as given C and W as well (thus, it is exactly as in the standard case in which the firm takes as given the price index, the wage rate and composite consumption).[88]

[88] Given $P_{j\neq i}$, one can compute the price index, as firm i is atomistic. Given $M = 1$ and P, one can compute C. Then, W can be obtained using the equilibrium conditions.

After taking into account that the production subsidy is set optimally ($t_s = \Lambda$), firm i's first-order condition writes as:

$$\left(\frac{P_i^*}{P}\right)^{\frac{1+\Lambda-\alpha}{\alpha\Lambda}} = \frac{1}{\alpha} W \frac{1}{PC} \left(\frac{C}{A_i}\right)^{1/\alpha} \tag{2.40}$$

Because we are in the deterministic case, $A_i = 1 \forall i$, so (2.40) implies that in equilibrium all firms set the same price. Thus, the L.H.S. of (2.40) is one. Using the equilibrium conditions to find an expression for the wage and substituting in (2.40) yields:

$$1 = \frac{1}{\alpha} C^{\gamma + \frac{\psi+1}{\alpha} - 1}. \tag{2.41}$$

Note from above that if the money supply was not fixed, there would be nominal indeterminacy. However, recalling that $M = 1$, the only price-setting equilibrium is the one in which each firm sets $\overline{P_i} = \frac{1}{\overline{C}}$, where $\overline{C}$ solves (2.41). Also note that changes in M lead to proportional changes in P (implying that changes in the money supply have no real effects when they are common knowledge).

Welfare (derivation of central bank's objective)

For the derivation of welfare, we will work with a large but finite number of firms, instead of a unit mass — this does not alter the results in any meaningful way, but it makes the derivation more straightforward. This also facilitates comparison with Paciello and Wiederholt (2013).[89]

The consumption aggregator with a finite number of firms is given by:

$$C = \left(\frac{1}{I} \sum_{i=1}^{I} C_i^{\frac{1}{1+\Lambda}}\right)^{1+\Lambda} \tag{2.42}$$

Which can be rearranged as:

$$\hat{C}_I := \frac{C_I}{C} = \left(I - \sum_{i=1}^{I-1} \hat{C}_i^{\frac{1}{1+\Lambda}}\right)^{1+\Lambda} \tag{2.43}$$

[89] I follow their derivation here. They work with a finite number of firms (instead of a unit mass) because they "find that it makes the derivation of the central bank's objective [...] more transparent".

Noting again that the household needs to supply the labour required for production yields the counterpart of equation (2.9) for a finite number of firms:

$$L = \sum_{i=1}^{I} \left(\frac{C_i}{A_i} \right)^{1/\alpha} \tag{2.44}$$

Substituting (2.43) and (2.44) into the utility function gives an expression for utility at any feasible allocation:

$$U(C, \hat{C}_1, \hat{C}_2, ...\hat{C}_{I-1}, V, A_1, A_2, ..., A_I) =$$

$$= \frac{C^{1-\gamma} - 1}{1-\gamma} - \frac{V}{1+\psi} C^{\frac{1+\psi}{\alpha}} \left[\sum_{i=1}^{I-1} \left(\frac{\hat{C}_i}{A_i} \right)^{1/\alpha} + \left(\frac{1}{A_I} \right)^{1/\alpha} \left(I - \sum_{i=1}^{I-1} \hat{C}_i^{\frac{1}{1+\Lambda}} \right)^{\frac{1+\Lambda}{\alpha}} \right]^{1+\psi} \tag{2.45}$$

Maximizing (2.45) gives:

$$C^* = \left\{ \frac{\alpha}{V I^{1+\psi}} \left[\left(\frac{1}{I} \sum_{i=1}^{I} A_i^{\frac{1}{1+\Lambda-\alpha}} \right)^{-(1-\frac{1+\Lambda}{\alpha})(1+\psi)} \right] \right\}^{\frac{1}{\gamma-1+\frac{1+\psi}{\alpha}}} \tag{2.46}$$

$$\hat{C}_i^* = \left(\frac{A_i^{\frac{1}{1+\Lambda-\alpha}}}{\frac{1}{I} \sum_{i=1}^{I} A_i^{\frac{1}{1+\Lambda-\alpha}}} \right)^{1+\Lambda} \quad \forall i \tag{2.47}$$

The equations above correspond to (2.11) and (2.12) in the main text. Using (2.46) and (2.47), we can easily see that in the absence of shocks ($V = 1$ and $A_i = 1 \forall i$), optimal consumption is characterised by:

$$C^* = \left(\frac{\alpha}{I^{1+\psi}} \right)^{\frac{1}{\gamma-1+\frac{1+\psi}{\alpha}}} \tag{2.48}$$

$$\hat{C}_i^* = 1 \tag{2.49}$$

Note that the steady-state is the same as the one in Paciello and Wiederholt (2013).

Expressing equation (2.45) in terms of log-deviations from steady-state:

$$u(c,\hat{c}_1,...,\hat{c}_{I-1},v,a_1,...,a_I) = \frac{\overline{C}^{1-\gamma}e^{(1-\gamma)c}-1}{1-\gamma} -$$
$$-\frac{\alpha e^v}{1+\psi}\overline{C}^{1-\gamma}e^{\frac{1+\psi}{\alpha}c}\left[\frac{1}{I}\sum_{i=1}^{I-1}e^{\frac{1}{\alpha}(\hat{c}_i-a_i)}+\frac{1}{I}\frac{1}{e^{\frac{1}{\alpha}a_I}}\left(I-\sum_{i=1}^{I-1}e^{\frac{1}{1+\Lambda}\hat{c}_i}\right)^{\frac{1+\Lambda}{\alpha}}\right]^{1+\psi} \tag{2.50}$$

Our measure of welfare is a second-order approximation of (2.50) around the origin. Denote this by $\widetilde{u}(c,\hat{c}_1,...,\hat{c}_{I-1},v,a_1,...,a_I)$. Let $\mathbf{x}$ be the input vector of the function $u(.)$:

$$\mathbf{x} = \begin{pmatrix} c & \hat{c}_1 & ... & \hat{c}_{I-1} & v & a_1 & ... & a_I \end{pmatrix}^T$$

We can express our second-order Taylor approximation as:

$$\widetilde{u}(\mathbf{x}) = u(0) + \nabla u(0)\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{H}_u(0)\mathbf{x} \tag{2.51}$$

Where $\nabla u(0)$ denotes the gradient of $u(.)$ evaluated at the origin:

$$\nabla u(0) = \begin{pmatrix} \frac{\partial u}{\partial c}(0) & \frac{\partial u}{\partial \hat{c}_1}(0) & ... & \frac{\partial u}{\partial \hat{c}_{I-1}}(0) & \frac{\partial u}{\partial v}(0) & \frac{\partial u}{\partial a_1}(0) & ... & \frac{\partial u}{\partial a_I}(0) \end{pmatrix}$$

And $\mathbf{H}_u(0)$ denotes the Hessian of $u(.)$ evaluated at the origin:

$$\mathbf{H}_u(0) = \begin{pmatrix}
\frac{\partial^2 u}{\partial c^2}(0) & \frac{\partial^2 u}{\partial c\partial \hat{c}_1}(0) & ... & \frac{\partial^2 u}{\partial c\partial \hat{c}_{I-1}}(0) & \frac{\partial^2 u}{\partial c\partial v}(0) & \frac{\partial^2 u}{\partial c\partial a_1}(0) & ... & \frac{\partial^2 u}{\partial c\partial a_I}(0) \\
\frac{\partial^2 u}{\partial \hat{c}_1\partial c}(0) & \frac{\partial^2 u}{\partial \hat{c}_1{}^2}(0) & ... & \frac{\partial^2 u}{\partial \hat{c}_1\partial \hat{c}_{I-1}}(0) & \frac{\partial^2 u}{\partial \hat{c}_1\partial v}(0) & \frac{\partial^2 u}{\partial \hat{c}_1\partial a_1}(0) & ... & \frac{\partial^2 u}{\partial \hat{c}_1\partial a_I}(0) \\
\vdots & & & & & & & \\
\frac{\partial^2 u}{\partial \hat{c}_{I-1}\partial c}(0) & \frac{\partial^2 u}{\partial \hat{c}_{I-1}\partial \hat{c}_1}(0) & ... & \frac{\partial^2 u}{\partial \hat{c}_{I-1}^2}(0) & \frac{\partial^2 u}{\partial \hat{c}_{I-1}\partial v}(0) & \frac{\partial^2 u}{\hat{c}_{I-1}\partial a_1}(0) & ... & \frac{\partial^2 u}{\partial \hat{c}_{I-1}\partial a_I}(0) \\
\frac{\partial^2 u}{\partial v\partial c}(0) & \frac{\partial^2 u}{\partial v\partial \hat{c}_1}(0) & ... & \frac{\partial^2 u}{\partial v\partial \hat{c}_{I-1}}(0) & \frac{\partial^2 u}{\partial v^2}(0) & \frac{\partial^2 u}{\partial v\partial a_1}(0) & ... & \frac{\partial^2 u}{\partial v\partial a_I}(0) \\
\frac{\partial^2 u}{\partial a_1\partial c}(0) & \frac{\partial^2 u}{\partial a_1\partial \hat{c}_1}(0) & ... & \frac{\partial^2 u}{\partial a_1\partial \hat{c}_{I-1}}(0) & \frac{\partial^2 u}{\partial a_1\partial v}(0) & \frac{\partial^2 u}{\partial a_1^2}(0) & ... & \frac{\partial^2 u}{\partial a_1\partial a_I}(0) \\
\vdots & & & & & & & \\
\frac{\partial^2 u}{\partial a_I\partial c}(0) & \frac{\partial^2 u}{\partial a_I\partial \hat{c}_1}(0) & ... & \frac{\partial^2 u}{\partial a_I\partial \hat{c}_{I-1}}(0) & \frac{\partial^2 u}{\partial a_I\partial v}(0) & \frac{\partial^2 u}{\partial a_I\partial a_1}(0) & ... & \frac{\partial^2 u}{\partial a_I^2}(0)
\end{pmatrix}$$

Computing the gradient and the Hessian yields:

$$\nabla u(0) = \overline{C}^{1-\gamma}\begin{pmatrix} 0 & 0 & ... & 0 & -\frac{\alpha}{1+\psi} & \frac{1}{I} & ... & \frac{1}{I} \end{pmatrix}$$

$$
\mathbf{H}_u(0) =
$$

$$
\overline{C}^{1-\gamma}\begin{pmatrix}
-(\frac{1+\psi}{\alpha}-1+\gamma) & 0 & ... & 0 & -1 & \frac{1+\psi}{\alpha I} & ... & \frac{1+\psi}{\alpha I} \\
0 & -\frac{2}{I}\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)} & ... & -\frac{1}{I}\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)} & 0 & \frac{1}{\alpha I} & ... & 0 \\
\vdots & & & & & & & \\
0 & -\frac{1}{I}\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)} & ... & -\frac{2}{I}\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)} & 0 & 0 & ... & 0 \\
-1 & 0 & ... & 0 & -\frac{\alpha}{1+\psi} & \frac{1}{I} & ... & \frac{1}{I} \\
\frac{1+\psi}{\alpha I} & \frac{1}{\alpha I} & ... & 0 & \frac{1}{I} & \frac{1}{I}\left[\psi\left(\frac{1}{\alpha I}\right)^2-\frac{1}{\alpha}\right] & ... & -\frac{\psi}{\alpha I^2} \\
\vdots & & & & & & & \\
\frac{1+\psi}{\alpha I} & 0 & ... & 0 & \frac{1}{I} & -\frac{\psi}{\alpha I^2} & ... & \frac{1}{I}\left[\psi\left(\frac{1}{\alpha I}\right)^2-\frac{1}{\alpha}\right]
\end{pmatrix}
$$

Using the law of large numbers to argue that $\frac{1}{I}\sum_{i=1}^{I} a_i = 0$, (2.51) simplifies to:

$$
\begin{aligned}
\widetilde{u}(\mathbf{x}) =& u(0) - \overline{C}^{1-\gamma}\frac{1}{1+\psi}v + \frac{1}{2}\overline{C}^{1-\alpha}\left[-\left(\frac{1+\psi}{\alpha}-1+\gamma\right)c^2 - \frac{\alpha}{1+\psi}v^2\right] - \\
&- \frac{1}{2}\overline{C}^{1-\gamma}\left\{\frac{2}{I}\left(\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)}\right)\sum_{i=1}^{I-1}\hat{c_i}^2 - \frac{1}{I}\left[\psi\left(\frac{1}{\alpha I}\right)^2 - \frac{1}{\alpha}\right]\sum_{i=1}^{I}{a_i}^2\right\} - \\
&- \overline{C}^{1-\gamma}cv + \overline{C}^{1-\gamma}\frac{1}{\alpha I}\sum_{i=1}^{I-1}\hat{c}_i a_i - \\
&- \frac{1}{2}\overline{C}^{1-\gamma}\frac{1}{I}\left(\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)}\right)\left[\sum_{i=1}^{I-1}\left(\hat{c}_i\sum_{j\neq i}\hat{c}_j\right)\right] - \\
&- \frac{1}{2}\overline{C}^{1-\gamma}\frac{\psi}{\alpha I^2}\left[\sum_{i=1}^{I}\left(z_i\sum_{j\neq i}z_j\right)\right]
\end{aligned}
\tag{2.52}
$$

Maximizing (2.52) w.r.t. $(c, \hat{c}_1, ..., \hat{c}_{I-1})$ yields:

$$
c^* = -\frac{1}{\gamma + \frac{1+\psi}{\alpha} - 1}v \tag{2.53}
$$

$$
\hat{c_i}^* = \frac{1+\Lambda}{1+\Lambda-\alpha}a_i \tag{2.54}
$$

Which correspond to equations (2.13) and (2.14) in the main text. Computing the difference between utility at any feasible allocation and utility at the optimal allocation gives:

$$
\begin{aligned}
&\widetilde{u}(c, \hat{c}_1, ..., \hat{c}_{I-1}, v, a_1, ..., a_I) - \widetilde{u}(c^*, \hat{c}_1{}^*, ..., \hat{c}^*_{I-1}, v, a_1, ..., a_I) = \\
&= \underbrace{-\frac{1}{2}\overline{C}^{1-\gamma}\left(\frac{1+\psi}{\alpha} - 1 + \gamma\right)}_{\eta}(c - c^*)^2 \\
&\underbrace{-\frac{1}{2}\overline{C}^{1-\gamma}\frac{1}{I}\left[\frac{1+\Lambda-\alpha}{\alpha(1+\Lambda)}\right]}_{\zeta}\sum_{i=1}^{I-1}(\hat{c}_i - \hat{c}^*_i)^2
\end{aligned}
\tag{2.55}
$$

Equation (2.55) corresponds to the welfare function used in the main text. Note that all results regarding policy hold for any $\eta > 0$ and $\zeta > 0$, so changing parameters does not qualitatively alter our main results (as the central bank will not face a trade-off between stabilizing the consumption level and improving the consumption-mix).

Derivation of firm's objective

Firm i's optimal price readily follows after log-linearizing (2.35) and (2.40) around the deterministic equilibrium. Alternatively, the firm's objective can be derived by taking a second-order approximation of the firm's profit function around the deterministic equilibrium (see Maćkowiak and Wiederholt (2009)).

2.8.2 Details and proofs regarding the information structure

Choice regarding characteristics of signal vector

We can represent the firm's choice regarding its allocation of attention in an intuitive way reminiscent of Maćkowiak et al. (2018). As the firm's signals can be about any linear combination of shocks, we can write:

$$
S_i = F_i X_i + \varepsilon_i \tag{2.56}
$$

Where F_i is a K_i x 3 matrix of coefficients and $\varepsilon_i = (\varepsilon_{1,i}, \varepsilon_{2,i}, ..., \varepsilon_{K,i})'$ is a Gaussian white noise random vector (independent of X_i) with covariance matrix $\Omega_{\varepsilon i}$.

$$
\begin{pmatrix} s_{i,1} \\ s_{i,2} \\ \vdots \\ s_{i,K_i} \end{pmatrix} = \begin{pmatrix} f^i_{1,1} & f^i_{1,2} & f^i_{1,3} \\ f^i_{2,1} & f^i_{2,2} & f^i_{2,3} \\ \vdots & \vdots & \vdots \\ f^i_{K_i,1} & f^i_{K_i,2} & f^i_{K_i,3} \end{pmatrix} \begin{pmatrix} c^* \\ m \\ a_i \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \vdots \\ \varepsilon_{i,K_i} \end{pmatrix} \tag{2.57}
$$

Each firm chooses the number of signals it receives (K_i), the content of these signals (by choosing the components of the matrix F_i), as well as the variances and covariances of the noise terms in the signals (by choosing the covariance matrix $\Omega_{\varepsilon i}$). As all noise in signals is idiosyncratic (condition 2 in the main text), ε_i is independent of ε_j for any $j \neq i$.

The independence assumption (condition 3 in the main text) constrains the firm's choice regarding the components of matrix F_i and the covariance matrix $\Omega_{\varepsilon i}$ — for any $n \in \{1, 2, ..., K_i\}$, if $f_{n,1} \neq 0$ or $f_{n,2} \neq 0$, then $f_{n,3} = 0$; if $f_{n,3} \neq 0$, then $f_{n,1} = f_{n,2} = 0$ (such that firms cannot receive any signal which is informative about both aggregate and idiosyncratic shocks). Also, if signal s_{in} concerns aggregate shocks (i.e. $f_{n,1} \neq 0$ or $f_{n,2} \neq 0$) and signal $s_{i,p}$ concerns the idiosyncratic shock (i.e. $f_{p,3} \neq 0$), then $cov(\varepsilon_{i,n}, \varepsilon_{i,p}) = 0$.

To fix ideas and clarify the implications of the independence assumption, suppose that firm i decides to learn about the variables of interest by receiving n signals about aggregate shocks and $(K_i - n)$ signals about the idiosyncratic shock. Without any loss of generality, we can order the signals such that the subvector containing information about aggregate shocks comprises the first n signals (i.e. $S_i^A = (s_{i,1}, s_{i,2}, ..., s_{i,n})$) and the subvector containing information about the idiosyncratic shock comprises the last $(K_i - n)$ signals (i.e. $S_i^I = (s_{i,n+1}, s_{i,n+2}, ..., s_{i,K_i})$). The independence assumption (condition 3) then requires that:

$$S_i = \begin{pmatrix} s_{i,1} \\ \vdots \\ s_{i,n} \\ s_{i,n+1} \\ \vdots \\ s_{i,K_i} \end{pmatrix} = \begin{pmatrix} f^i_{1,1} & f^i_{1,2} & 0 \\ \vdots & \vdots & \vdots \\ f^i_{n,1} & f^i_{n,2} & 0 \\ 0 & 0 & f^i_{n+1,3} \\ \vdots & \vdots & \vdots \\ 0 & 0 & f^i_{K_i,3} \end{pmatrix} \begin{pmatrix} c^* \\ m \\ a_i \end{pmatrix} + \begin{pmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \\ \vdots \\ \varepsilon_{i,K_i} \end{pmatrix}; \qquad (2.58)$$

$$\Omega_{\varepsilon i} = \begin{pmatrix} \Omega^A_{\varepsilon i} & \mathbf{0} \\ \mathbf{0} & \Omega^I_{\varepsilon i} \end{pmatrix} \text{ (with sizes } \begin{pmatrix} n\mathbf{x}n & n\mathbf{x}(K_i - n) \\ (K_i - n)\mathbf{x}n & (K_i - n)\mathbf{x}(K_i - n) \end{pmatrix}) \qquad (2.59)$$

Where $\Omega^A_{\varepsilon i}$ denotes the covariance matrix of noise in signals concerning aggregate shocks and $\Omega^I_{\varepsilon i}$ denotes the covariance matrix of noise in signals containing information about the idiosyncratic shock:

$$\Omega^A_{\varepsilon i} = \begin{pmatrix} \sigma^2_{i,1} & \dots & cov(\varepsilon_{i,1}, \varepsilon_{i,n}) \\ \vdots & \vdots & \vdots \\ cov(\varepsilon_{i,n}, \varepsilon_{i,1}) & \dots & \sigma^2_{i,n} \end{pmatrix} \tag{2.60}$$

$$\Omega^I_{\varepsilon i} = \begin{pmatrix} \sigma^2_{i,n+1} & \dots & cov(\varepsilon_{i,n+1}, \varepsilon_{i,K_i}) \\ \vdots & \vdots & \vdots \\ cov(\varepsilon_{i,K_i}, \varepsilon_{i,n+1}) & \dots & \sigma^2_{i,K_i} \end{pmatrix} \tag{2.61}$$

We can then simply express the problem solved by firm i when choosing its allocation of attention as:

$$\begin{aligned} &\min_{\{K_i, F_i, \Omega_{\varepsilon i}\}} \frac{1}{2} E[p_i - r(m - c^*) - (1-r)p + z_i]^2 \\ &\text{Subject to: } p_i = E[r(m - c^*) + (1-r)p - z_i | S_i], \\ &\qquad (2.58),\ (2.59), I(S_i; X_i) \leq \kappa. \end{aligned} \tag{2.62}$$

The firm chooses the number of signals it receives (K_i), how many of these concern aggregate and idiosyncratic shocks, as well as their content (F_i) and the variances and covariances of the noise terms in the signals ($\Omega_{\varepsilon i}$), subject to the information processing constraint. It anticipates how it will optimally set its price conditional on the information it receives and takes into account how its decision regarding the allocation of attention will affect its profit losses. Equations (2.58) and (2.59) place additional constraints on the firm's available choices regarding F_i and $\Omega_{\varepsilon i}$, ensuring that the independence assumption (condition 3 in the main text) is satisfied.

Proof of Proposition 2.1

Using the independence assumption and taking into account that each firm sets its price optimally conditional on the information it receives, we can express firm i's problem when choosing its allocation of attention (2.62) as:

$$\begin{aligned} &\min_{\{K_i, F_i, \Omega_{\varepsilon i}\}} \frac{1}{2} var[r(m-c^*)+(1-r)p|S_i^A]+\frac{1}{2}var[z_i|S_i^I] \\ &\text{Subject to: } (2.58), (2.59), I[(c^*,m),S_i^A]+I(z_i,S_i^I)\leq \kappa. \end{aligned} \tag{2.63}$$

Where it was also noted that $I(a_i, S_i^I) = I(z_i, S_i^I)$ (for any S_i^I).

To determine the equilibrium allocation of attention, I approach the problem in two steps:

1) Fix the amount of attention paid by firms to aggregate and idiosyncratic conditions (i.e. suppose that $I[(c^*,m),S_i^A] \leq \kappa_A$ and $I(z_i,S_i^I) \leq \kappa_I$ for all i, with $\kappa_A+\kappa_I \leq \kappa$, $\kappa_A \geq 0$, $\kappa_I \geq 0$) and find the K, F, Ω_ε which are consistent with equilibrium.

2) Find the κ_A and κ_I which are consistent with equilibrium.

Proposition 2.1'. *In any symmetric equilibrium in which the amount of attention allocated by firms towards aggregate (idiosyncratic) shocks is given by κ_A (κ_I) — i.e. $I[(c^*,m),S_i^A] \leq \kappa_A$ and $I(z_i,S_i^I) \leq \kappa_I \forall i$ — it is WLOG to restrict our attention to an information structure in which each firm observes two private signals: one about the composite aggregate shock $q = m - c^*$ ($s_{i,1} = q + \varepsilon_{i,1} = m - c^* + \varepsilon_{i,1}$) and one about its idiosyncratic shock z_i ($s_{i,2} = z_i + \varepsilon_{i,2}$), where $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$ are drawn independently for each i from the distributions $N \sim (0, \sigma_1^2)$ and $N \sim (0, \sigma_2^2)$ respectively. Further, the variance of the noise terms in private signals is equal to $\sigma_1^2 = \sigma_q^2/(2^{2\kappa_A}-1)$ and $\sigma_2^2 = \sigma_z^2/(2^{2\kappa_I}-1)$, where $\sigma_q^2 = var(m-c^*)$.*[90]

Remark 2.3. Following Proposition 2.1, to fully characterise the equilibrium allocation of attention, it suffices to find the pair (κ_A, κ_I) which is consistent with equilibrium. More generally, the approach to determining the equilibrium alloca-

[90]i.e. we only need to analyse equilibria in which: $K = 2, F = \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & \phi \end{pmatrix}, \Omega_\varepsilon = \begin{pmatrix} \sigma_q^2/(2^{2\kappa_A}-1) & 0 \\ 0 & \sigma_z^2/(2^{2\kappa_I}-1) \end{pmatrix}$.

tion of attention is the following: first, note that for any (κ_A, κ_I), it is WLOG to restrict attention to information structures as described in Proposition 2.1; then (later on, in Claim 2.2), find the (κ_A, κ_I) which is consistent with equilibrium.

Proof: Firm i solves:

$$\begin{aligned} &\min_{\{K_i, F_i, \Omega_{\varepsilon i}\}} \frac{1}{2} var[r(m - c^*) + (1-r)p|S_i^A] + \frac{1}{2} var[z_i|S_i^I] \\ &\text{Subject to: } (2.58);\ (2.59)\ ; I[(c^*, m), S_i^A] \leq \kappa_A; I(z_i, S_i^I) \leq \kappa_I. \end{aligned} \tag{2.64}$$

Because of the way we split the problem, the firm's choice regarding its signals containing information about aggregate shocks and its choice regarding signals containing information about the idiosyncratic shock are now two separate (independent) smaller problems. Before, if the firm chose a particular signal vector concerning idiosyncratic shocks, it had to take into account how much processing capacity this used up and how much was then available to process information about the aggregate shocks. Now, this is no longer the case and the firm's decision regarding how to learn about aggregate shocks $(K_i^A, F_i^A, \Omega_{\varepsilon i}^A)$ is independent of its decision regarding how to learn about the idiosyncratic shock $(K_i^I, F_i^I, \Omega_{\varepsilon i}^I)$, so we can analyze these two choices in turn.

I start with the choice regarding learning about the idiosyncratic shock. The subproblem reads:

$$\begin{aligned} &\min_{K_i^I, F_i^I, \Omega_{\varepsilon i}^I} \frac{1}{2} var[z_i|S_i^I] \\ \text{Subject to: } \ & S_i^I = F_i^I a_i + \varepsilon_i^I; \\ & \varepsilon_i^I \sim N(0, \Omega_{\varepsilon i}^I); \\ & I(z_i, S_i^I) \leq \kappa_I. \end{aligned}$$

The constraint on information flow regarding the idiosyncratic shock implies that:

$$H(z_i) - H(z_i|S_i^I) \leq \kappa_I$$

$$\frac{1}{2} log_2(2\pi e \sigma_z^2) - \frac{1}{2} log_2(2\pi e\ var(z_i|S_i^I)) \leq \kappa_I$$

$$log_2(\frac{\sigma_z^2}{var(z_i|S_i^I)}) \leq 2\kappa_I$$

$$var(z_i|S_i^I) \geq \frac{\sigma_z^2}{2^{2\kappa_I}}$$

The objective function is increasing in $var(z_i|S_i)$ and the firm wants to minimize it, so the constraint is binding and it must be the case that an optimal signal vector leads to:

$$var(z_i|S_i^I) = \frac{\sigma_z^2}{2^{2\kappa_I}}$$

Any signal vector with the property that $var(z_i|S_i^I) = \frac{\sigma_z^2}{2^{2k_I}}$ is optimal, so the optimal choice regarding the number and content of signals concerning the idiosyncratic shock is indeterminate - there is an infinite number of signal vectors leading to the same conditional variance (as mentioned in the main text, observing multiple signals about z_i is equivalent to observing a single signal with a higher precision). However, the posterior uncertainty following the observation of the optimal signal vector is unique. Thus, without loss of generality, we can restrict attention to signal vectors in which the firm observes one signal about the idiosyncratic shock of the form $s_{i,2} = z_i + \varepsilon_{i,2}$, with $\varepsilon_{i,2} \sim N(0, \sigma_2^2)$ and $\sigma_2^2 = \frac{\sigma_z^2}{2^{2\kappa_I - 1}}$.[91]

(Bayesian updating given the normal information structure implies that $var(z_i|s_{i,2}) = \frac{\sigma_z^2 \sigma_2^2}{\sigma_z^2 + \sigma_2^2}$. Because $\sigma_2^2 = \frac{\sigma_z^2}{2^{2\kappa_I - 1}}$, it follows that $var(z_i|S_i^I) = \frac{\sigma_z^2}{2^{2\kappa_I}}$, so the signal $s_{i,2}$ is optimal.)

Now, let's turn our attention to firm i's choice regarding its signals containing information about aggregate shocks. The subproblem reads:

$$\min_{K_i^A, F_i^A, \Omega_{\varepsilon i}^A} \frac{1}{2} var[r(m - c^*) + (1 - r)p|S_i^A]$$

$$\text{Subject to: } S_i^A = F_i^A \begin{pmatrix} m \\ c^* \end{pmatrix} + \varepsilon_i^A;$$

$$\varepsilon_i^A \sim N(0, \Omega_{\varepsilon i}^A);$$

$$I[(m, c^*), S_i^A] \le \kappa_A.$$

The problem here is more complicated, because firms track two (correlated) shocks and because of the strategic complementarity in price-setting. Recall that we want to prove that it is WLOG to restrict attention to an information structure in which each firm observes a private signal of the form $s_{i,1} = m - c^* + \varepsilon_{i,1}$, where $\varepsilon_{i,1} \sim N(0, \sigma_q^2/(2^{2\kappa_A} - 1))$. I argue this in four steps:

- Step 1: Prove Lemma 1 below
- Step 2: Note that for any given symmetric allocation of attention, p is a linear

[91] i.e. a solution is $K_i^I = 1; F_i^I = (\phi); \Omega_{\varepsilon i}^I = (\frac{\sigma_z^2}{2^{2\kappa_I - 1}})$.

combination of m and c^* (in price-setting equilibrium among firms)

- Step 3: Using Lemma 1 and step 2, note that in any symmetric equilibrium the price level must be proportional to q (i.e. $p =$ constant x $(m - c^*)$)
- Step 4: Rest of proof follows immediately from Lemma 1 and step 3.

Lemma 1. *If the firm solves:*

$$\begin{aligned} &\min_{K_i^A, F_i^A, \Omega_{\varepsilon i}^A} [var(d_1 m + d_2 c^* | S_i^A)] \\ &\textit{Subject to: } S_i^A = F_i^A \begin{pmatrix} m \\ c^* \end{pmatrix} + \varepsilon_i^A; \\ &\varepsilon_i^A \sim N(0, \Omega_{\varepsilon i}^A); \\ &I[(m, c^*), S_i^A] \leq \kappa_A. \end{aligned} \tag{2.65}$$

(where d_1 and d_2 are non-zero constants) it is without loss of generality to restrict attention to one-dimensional signal vectors of the form $s_{i,1} = d_1 m + d_2 c^* + \varepsilon_{i,1}$, where $\varepsilon_{i,1} \sim N(0, \sigma_1^*)$, i.e. a solution is $K_i^A = 1, F_i^A = \begin{pmatrix} d_1 & d_2 \end{pmatrix}$ and $\sigma_1^* = var(d_1 m + d_2 c^*)/(2^{2\kappa_A} - 1)$.

Proof: See Appendix 2.8.2. □

Step 2: *Note that for any given symmetric allocation of attention, p is a linear combination of m and c^* (in price-setting equilibrium among firms).*

Fix any symmetric allocation of attention:

$$(K_i, F_i, \Omega_{\varepsilon i}) = (K, F, \Omega_\varepsilon) \forall i.$$

Price-setting conditional on this symmetric allocation of attention is equivalent to a beauty contest game with an exogenous (Gaussian) information structure similar to Morris and Shin (2002), as each firm sets its price to match a weighted sum of $m - c^*$ and the average price (as well as its idiosyncratic shock):

$$p_i = E[r(m - c^*) + (1 - r)p - z_i | S_i]$$

As such, it should be without loss of generality to restrict our attention to linear equilibria — i.e. equilibria in which each firm sets its price as a linear function of its signals:

$$p_i = \beta S_i = \beta_A S_i^A + \beta_I S_i^I \tag{2.66}$$

Where β is a $1\mathbf{x}K$ matrix of coefficients to be determined in equilibrium (β_A is the left $1\mathbf{x}K^A$ submatrix of β denoting firms' response to signals containing information about aggregate shocks and β_I is the right $1\mathbf{x}K^I$ submatrix of β denoting firms' response to signals containing information about idiosyncratic shocks). (2.66) is equivalent to:

$$p_i = \beta_A[F^A \begin{pmatrix} m \\ c^* \end{pmatrix} + \varepsilon_i^A] + \beta_I[F^I a_i + \varepsilon_I^I] \tag{2.67}$$

As all noise in signals is idiosyncratic and the a_i's are i.i.d., aggregating (2.67) across firms gives:

$$p = \beta_A F^A \begin{pmatrix} m \\ c^* \end{pmatrix}$$

Which is indeed a linear combination of m and c^*.

Note that this is by no means a (rigorous) proof, as we restricted attention to linear equilibria. More rigorously, one can also show that it is WLOG to restrict attention to linear equilibria by expressing each firm's optimal price as an infinite sum of its first and higher-order beliefs of $(m - c^*)$, then arguing that beliefs of various order are all linear combinations of signals (and the infinite sum is bounded), but this is superfluous at this point given the literature (see Morris and Shin (2002), section D for the initial argument which could also be applied here although the information structure here is more complicated; I also go over the argument in Appendix 2.8.3 for the information structure described in Proposition 2.1 (the argument proves the uniqueness of the linear equilibrium given that information structure)).

Step 3: *Using Lemma 1 and step 2, note that in any symmetric equilibrium the price level must be proportional to q (i.e. that we can express $p =$ constant x $(m - c^*)$).*

The approach is the following. We start by positing that in equilibrium the price level is a particular linear combination of m and c^* (i.e. that $p = \gamma_1 m + \gamma_2 c^*$). Taking this as given, we characterise firms' optimal signals (using the Lemma) and their equilibrium prices. We then check whether the price level is indeed given by the linear combination (of m and c^*) which we posited at the beginning. The only linear combinations which verify our initial guess are of the form $\gamma_1 = -\gamma_2$ (implying that in any symmetric equilibrium the price level is proportional to $m - c^*$).

Suppose that the price level is given by $p = \gamma_1 m + \gamma_2 c^*$. Firm i's profit-maximizing price is given by:

$$\hat{p}_i = r(m - c^*) + (1-r)(\gamma_1 m + \gamma_2 c^*) - z_i \Leftrightarrow$$
$$\Leftrightarrow \hat{p}_i = \underbrace{[r + (1-r)\gamma_1]}_{\xi_1} m + \underbrace{[(1-r)\gamma_2 - r]}_{\xi_2} c^* - z_i$$

Thus, when choosing the properties of its signal vector regarding aggregate shocks, firm i solves:

$$\min_{K_i^A, F_i^A, \Omega_{\varepsilon i}^A} [var(\xi_1 m + \xi_2 c^* | S_i^A)]$$
$$\text{Subject to: } S_i^A = F_i^A \begin{pmatrix} m \\ c^* \end{pmatrix} + \varepsilon_i^A;$$
$$\varepsilon_i^A \sim N(0, \Omega_{\varepsilon i}^A);$$
$$I[(m, c^*), S_i^A] \leq \kappa_A.$$

Using the Lemma, it is without loss of generality to restrict our attention to equilibria in which the firms' signal vector concerning aggregate shocks comprises only the signal $s_{i,1} = \xi_1 m + \xi_2 c^* + \varepsilon_{i,1}$, where $\varepsilon_{i,1} \sim N(0, \sigma_1^{2*})$ and $\sigma_1^* = var(\xi_1 m + \xi_2 c^*)/(2^{2\kappa_A} - 1)$. Given this signal, each firm sets its price according to:

$$p_i^* = \left(1 - 2^{-2\kappa_A}\right) s_{1,i} - E[z_i | S_i^I]$$

Aggregating across firms, the price level is given by:

$$p = \left(1 - 2^{-2\kappa_A}\right)(\xi_1 m + \xi_2 c^*)$$

In order for our guess to be verified, it must be the case that:

$$\begin{cases} \gamma_1 = (1 - 2^{-2\kappa_A})\xi_1 \\ \gamma_2 = (1 - 2^{-2\kappa_A})\xi_2 \end{cases} \tag{2.68}$$

Solving this yields:

$$\begin{cases} \gamma_1 = \frac{r\beta_1}{1-(1-r)\beta_1} \\ \gamma_2 = -\frac{r\beta_1}{1-(1-r)\beta_1} \end{cases} \tag{2.69}$$

Where $\beta_1 = 1 - 2^{-2\kappa_A}$. Note that $\gamma_1 = -\gamma_2$ for any κ_A and r.

Step 3 says that regardless of how much attention firms pay to aggregate shocks in equilibrium (for any κ_A), if firms optimally choose the number, content and precision of their signals and set prices optimally conditional on the information they receive, then the price level must be proportional to $m - c^*$. Note that this can be proved in multiple ways (and the argument above is just one of them).

Step 4: Because in any symmetric equilibrium the price level is proportional to $m - c^*$, we can rewrite firm i's problem as

$$\begin{aligned} \min_{K_i^A, F_i^A, \Omega_{\varepsilon i}^A} \quad & \text{constant x } [var(m - c^* | S_i^A)] \\ \text{Subject to: } \quad & S_i^A = F_i^A \begin{pmatrix} m \\ c^* \end{pmatrix} + \varepsilon_i^A; \\ & \varepsilon_i^A \sim N(0, \Omega_{\varepsilon i}^A); \\ & I[(m, c^*), S_i^A] \leq \kappa_A. \end{aligned}$$

The objective is equivalent to the one from Lemma 1 if $d_1 = 1$ and $d_2 = -1$. It follows that a solution is to observe the signal $s_{i,1} = m - c^* + \varepsilon_{i,1}$, with $\varepsilon_{i,1} \sim N(0, \sigma_1)$ and $\sigma_1^2 = var(m - c^*)/(2^{2\kappa_A} - 1)$.

Collecting results yields Proposition 2.1'. $\square$

Proof of Lemma 1

Proof: Recall the CB's policy rule $m = \rho x = \rho(c^* + \varepsilon_x)$ (we are solving the problem for $\rho \neq 0$, as $d_1 \neq 0$). The idea is to note that we can rewrite the problem as:

$$\begin{aligned} \min_{K_i^A, \tilde{F_i^A}, \Omega_{\varepsilon i}^A} \quad & [var(d_1 \rho \varepsilon_x + (d_1 \rho + d_2) c^* | S_i^A)] \\ \text{Subject to: } \quad & S_i^A = \tilde{F_i^A} \begin{pmatrix} \varepsilon_x \\ c^* \end{pmatrix} + \varepsilon_i^A; \\ & \varepsilon_i^A \sim N(0, \Omega_{\varepsilon i}^A); \\ & I[(\varepsilon_x, c^*), S_i^A] \leq \kappa_A. \end{aligned} \tag{2.70}$$

That is, we can set up the firm's tracking problem in terms of the independent variables c^* and ε_x — this is useful, because the solution to this problem follows directly from Adam (2007). Two observations are in order, to argue that problems (2.65) and (2.70) are equivalent.

Firstly, note that in both problems firms can choose from the same menu of signals — the set of all linear combinations of m and c^* is the same as the set of all linear combinations of ε_x and c^*.

Secondly, note that the information processing constraint is the same in the two problems. The vector (m, c^*) can be computed from the vector (ε_x, c^*) and vice-versa — thus, they contain the same information and it follows[92] that $I[(m, c^*), S_i^A] = I[(\varepsilon_x, c^*), S_i^A]$ for any signal vector S_i^A.

Hence, in both problems, the firm chooses from the same menu of signals to maximize the same objective function, subject to the same constraint — clearly, solutions must coincide.

It follows from Adam (2007) that when the firm solves the latter problem (expressed in terms of independent variables), it is without loss of generality to restrict attention to one-dimensional signal vectors of the form $\tilde{s_{i,1}} = d_1\rho\varepsilon_x + (d_1\rho + d_2)c^* + \varepsilon_{i,1}$, where $\varepsilon_{i,1} \sim N(0, \sigma_1^{2*})$, i.e. a solution is $K_i = 1, \tilde{F_i} = \begin{pmatrix} d_1\rho & d_1\rho + d_2 \end{pmatrix}$ and $\sigma_1^{2*} = var[d_1\rho\varepsilon_x + (d_1\rho + d_2)c^*]/(2^{2\kappa_A} - 1)$. Lastly, note that this is the same as the signal $s_{i,1}$ from the Lemma. $\square$

Remark 2.5: *See Adam (2007) — section 4.1 for the reference (more specifically, problem (18) and Appendix A.3). Similar results are discussed more generally in Sims (2010) and Fulton (2017).*

[92] This is a property of mutual information — see for instance, Fulton (2017), property 3 and corollary, or Maćkowiak et al. (2018).

2.8.3 Details regarding equilibrium when central bank information is exogenous

Proof of Claim 2.1

Note that this was actually already proved in Appendix 2.8.2 (see (2.69)). The following proof shows the uniqueness of the linear equilibrium associated with the information structure from Proposition 2.1 (recall that we restricted our attention to linear equilibria). The following argument can also be extended to show that it is WLOG to restrict attention to linear equilibria (in step 2 of Appendix 2.8.2). Also, note again that this is virtually equivalent to the argument in Adam (2007), section 5.

Proof: Each firm's optimal pricing rule specifies:

$$p_i = E_i[r(m - c^*) + (1 - r)p - z_i] \tag{2.71}$$

Where $E_i[.]$ denotes the expectation conditional on firm i's information set, i.e. $E_i[.] = E[.|S_i]$. We have fixed a symmetric allocation of attention $\kappa_{i,A} = \kappa_A$ and $\kappa_{i,I} = \kappa_I \forall i$. Following Proposition 2.1', it is WLOG to restrict our attention to an information structure in which each firm observes two private signals: one about the composite aggregate shock $q = m - c^*$ ($s_{i,1} = q + \varepsilon_{i,1} = m - c^* + \varepsilon_{i,1}$) and one about its idiosyncratic shock z_i ($s_{i,2} = z_i + \varepsilon_{i,2}$), where $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$ are drawn independently for each i from the distributions $N \sim (0, \sigma_q^2/(2^{2\kappa_A} - 1))$ and $N \sim (0, \sigma_z^2/(2^{2\kappa_I} - 1))$ respectively.

Bayesian updating given this information structure implies that:

$$E_i[q] = \underbrace{[1 - 2^{-2\kappa_A}]}_{\beta_1} s_{i,1}$$
$$E_i[z_i] = \underbrace{[1 - 2^{-2\kappa_I}]}_{\beta_2} s_{i,2}$$

Before proceeding, we introduce some notation in the spirit of Morris and Shin (2002) to be able to refer to beliefs of various order — denote by $\overline{E}[q]$ the average expectation of q across firms (i.e. $\overline{E}[q] = \int_0^1 E_i[q]di$), by $\overline{E}^2[q]$ the average expectation of the average expectation of q across firms (i.e. $\overline{E}^2[q] = \overline{E}[\overline{E}[q]]$) and so on.

Note that all noise in signals is idiosyncratic and the a_i's are i.i.d. across firms, so $\int_0^1 E_i[z_i] = 0$. By repeatedly integrating across firms, taking conditional expectations and substituting back in (2.71), we can rewrite it as:

$$p_i = rE_i[q] + r\sum_{n=1}^{\infty}(1-r)^n E_i\left[\overline{E}^n(q)\right] - E_i[z_i] \tag{2.72}$$

We can then compute the firm's beliefs of various order about q.

We know that $E_i[q] = \beta_1 s_{i,1} = \beta_1[q + \varepsilon_{i,1}]$. Aggregating across firms gives: $\overline{E}[q] = \beta_1 q$.

Hence, $E_i[\overline{E}(q)] = E_i[\beta_1 q] = \beta_1 E_i[q] = \beta_1^2 s_{i,1}$.

Repeatedly applying the same operation yields: $E_i[\overline{E}^n(q)] = \beta_1^{n+1} s_{i,1}$.

Substituting the higher-order beliefs into (2.72) and computing the infinite sum gives:

$$p_i = \frac{r\beta_1}{1-(1-r)\beta_1} s_{i,1} - \beta_2 s_{i,2}.$$

□

Proof of Claim 2.2

Note again that this is equivalent to the equilibrium in Maćkowiak and Wiederholt (2009), section 5 (for the derivation of the equilibrium, see the technical appendix of their paper). The derivation here is slightly different because firms do not observe the price level.

Proof: Consider firm i's problem when choosing its allocation of attention. The firm takes as given the CB's policy rule (thus, it takes as given the variance of the aggregate shock σ_q^2) and the other firms' allocations of attention. As there is a unit mass of firms (and firm i is atomistic), firm i's decision regarding its allocation of attention and firm i's price do not influence the aggregate price level. Thus, firm i effectively takes the price level as given (because it takes as given other firms' attention allocations and it anticipates how prices are set in equilibrium (see Claim 2.1)).

Following Proposition 2.1', firm i's problem writes as:

$$\min_{\kappa_A,\kappa_I} var[rq + (1-r)p - z_i|S_i]$$

$$\text{Subject to: } S_i = \begin{pmatrix} q + \varepsilon_{i,1} \\ z_i + \varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma_q^2}{2^{\kappa_A}-1} & 0 \\ 0 & \frac{\sigma_z^2}{2^{\kappa_I}-1} \end{pmatrix};$$

$$\kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A;\kappa_I) \in [0,\kappa]^2.$$

In what follows I use the following Lemma to simplify the exposition:

Lemma 2. *Suppose w is a constant (which the firm takes as given). Consider the problem:*

$$\min_{\kappa_A,\kappa_I} var[wq - z_i|S_i]$$

$$\text{Subject to: } S_i = \begin{pmatrix} q + \varepsilon_{i,1} \\ z_i + \varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma_q^2}{2^{\kappa_A}-1} & 0 \\ 0 & \frac{\sigma_z^2}{2^{\kappa_I}-1} \end{pmatrix};$$

$$\kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A;\kappa_I) \in [0,\kappa]^2.$$

The solution is:

$$\kappa_A^* = \begin{cases} \kappa & \text{if } \frac{w\sigma_q}{\sigma_z} \geq 2^{\kappa} \\ \frac{1}{2}\kappa + \frac{1}{4}log_2\left(\frac{w^2\sigma_q^2}{\sigma_z^2}\right) & \text{if } \frac{w\sigma_q}{\sigma_z} \in (2^{-\kappa}, 2^{\kappa}) \\ 0 & \text{if } \frac{w\sigma_q}{\sigma_z} \leq 2^{-\kappa} \end{cases}$$

$$\kappa_I^* = \kappa - \kappa_A^*$$

Proof. Bayesian updating given the signal vector S_i implies that $var[q|S_i] = \sigma_q^2/2^{2\kappa_A}$ and $var[z_i|S_i] = \sigma_z^2/2^{2\kappa_I}$. Thus, we can rewrite the problem as:

$$\min_{\kappa_A,\kappa_I} w^2\sigma_q^2/2^{2\kappa_A} + \sigma_z^2/2^{2\kappa_I}$$

$$\text{Subject to: } \kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A;\kappa_I) \in [0,\kappa]^2.$$

This is a standard constrained minimization problem whose solution is given in the Lemma. Alternatively, see Maćkowiak and Wiederholt (2009), section 5 (equations 35 and 36). □

Now let us analyse the equilibrium allocation of attention — the approach is the following: fix the allocation of attention of all firms $j \neq i$:

$$(\kappa_{j,A}, \kappa_{j,I}) = (\overline{\kappa}_A, \overline{\kappa}_I) \forall j \neq i$$

And compute the price level using Claim 2.1; then, analyse firm i's optimal allocation of attention. As we are looking for a symmetric equilibrium, it must be the case that firm i finds it optimal to choose the same allocation of attention as the other firms $(\kappa^*_{i,A}, \kappa^*_{i,I}) = (\overline{\kappa}_A, \overline{\kappa}_I)$. Essentially, this is a fixed point problem in which the equilibrium allocation of attention maps to itself.

Firstly, let us consider equilibria in which firms pay attention only to idiosyncratic conditions. Suppose $(\overline{\kappa}_A, \overline{\kappa}_I) = (0, \kappa)$. By Claim 2.1, prices are set according to $p_j = \beta_2 s_{j,2}$, so the price level is $p = 0$. Thus, firm i's problem writes as:

$$\begin{gathered} \min_{\kappa_A, \kappa_I} var[rq - z_i | S_i] \\ \text{Subject to: } S_i = \begin{pmatrix} q + \varepsilon_{i,1} \\ z_i + \varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma_q^2}{2^{\kappa_A} - 1} & 0 \\ 0 & \frac{\sigma_z^2}{2^{\kappa_I} - 1} \end{pmatrix}; \\ \kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A; \kappa_I) \in [0, \kappa]^2. \end{gathered}$$

By Lemma 2, firm i finds it optimal to choose $(\kappa^*_{i,A}, \kappa^*_{i,I}) = (0, \kappa)$ if $\frac{r\sigma_q}{\sigma_z} \leq 2^{-\kappa}$. Thus, for such policy and parameter values, there is an equilibrium in which firms only pay attention to idiosyncratic shocks.

Secondly, consider equilibria in which firms pay attention only to aggregate conditions. Suppose $(\overline{\kappa}_A, \overline{\kappa}_I) = (\kappa, 0)$. By Claim 2.1, prices are set according to $p_j = \frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})} s_{j,1}$, so the price level is given by $p = \underbrace{\frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})}}_{\alpha_H} q$.

Thus, firm i's problem writes as:

$$\begin{gathered} \min_{\kappa_A, \kappa_I} var[(r + (1-r)\alpha_H) q - z_i | S_i] \\ \text{Subject to: } S_i = \begin{pmatrix} q + \varepsilon_{i,1} \\ z_i + \varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma_q^2}{2^{\kappa_A} - 1} & 0 \\ 0 & \frac{\sigma_z^2}{2^{\kappa_I} - 1} \end{pmatrix}; \\ \kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A; \kappa_I) \in [0, \kappa]^2. \end{gathered}$$

By Lemma 2, firm i finds it optimal to choose $(\kappa^*_{i,A}, \kappa^*_{i,I}) = (\kappa, 0)$ if $\frac{[r+(1-r)\alpha_H]\sigma_q}{\sigma_z} \geq 2^{\kappa}$. This simplifies to $\frac{r\sigma_q}{\sigma_z} \geq 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa})$; for such policy and parameter values there is an equilibrium in which firms pay attention only to aggregate shocks.

Lastly, consider equilibria in which firms pay attention to both aggregate and idiosyncratic shocks. Suppose $(\overline{\kappa}_A, \overline{\kappa}_I) \in (0, \kappa)^2$. Using the same logic as before, the price level is given by:

$$p = \underbrace{\frac{r(1-2^{-2\overline{\kappa}_A})}{1-(1-r)(1-2^{-2\overline{\kappa}_A})}}_{\alpha_M} q \tag{2.73}$$

Using Lemma 2, it follows that at an interior solution the amount of attention firm i should pay to aggregate conditions is:

$$\kappa_{i,A}^{*} = \frac{1}{2}\kappa + \frac{1}{4}log_2 \left\{ [r + (1-r)\alpha_M]^2 \frac{\sigma_q^2}{\sigma_z^2} \right\} \tag{2.74}$$

Note again that we are looking for a symmetric equilibrium, so it must be the case that $\kappa_{i,A}^{*} = \overline{\kappa}_A$; denote the fixed point (i.e. the solution to (2.74)) by $\overline{\kappa}_A^{*}$. Substituting α_M from (2.73) into (2.74) and rearranging gives:

$$(\frac{1}{2^{2\overline{\kappa}_A}})^2 \left[2^{2\kappa} r^2 \frac{\sigma_q^2}{\sigma_z^2} - (1-r)^2 \right] - 2(1-r)r(\frac{1}{2^{2\overline{\kappa}_A}}) - r^2 = 0$$

Which is a quadratic equation. Solving this yields:

$$2^{2\overline{\kappa}_A^{*}} \in \left\{ -\left(\frac{2^{\kappa}\sigma_q}{\sigma_z} + \frac{1}{r} - 1 \right); \frac{2^{\kappa}\sigma_q}{\sigma_z} - \frac{1}{r} + 1 \right\}$$

The first root is negative, as $r \in (0,1)$ and thus cannot be a solution (as $2^{2\overline{\kappa}_A^{*}} \geq 1$). Rearranging gives:

$$\overline{\kappa}_A^{*} = \frac{1}{2} log_2 \left(1 - \frac{1}{r} + \frac{2^{\kappa}\sigma_q}{\sigma_z} \right)$$

Indeed, this conforms with our conditions for an interior solution if $\frac{r\sigma_q}{\sigma_z} \in (2^{-\kappa}, 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa}))$.

Note that there is a unique equilibrium for any policy rule and parameter values (as parameter regions for which we obtain distinct equilibria are mutually exclusive). Also, the equilibrium coincides with the one from Maćkowiak and Wiederholt (2009). □

Proof of Claim 2.3

Proof: Recall that the setting of policy determines the variance of the aggregate shock q:

$$\sigma_q^2 = var(m - c^*) = var[\rho(c^* + \varepsilon_x) - c^*] = (\rho - 1)^2\sigma_{c^*}^2 + \rho^2\sigma_x^2 \tag{2.75}$$

We will find an expression for expected welfare as a function of σ_q^2 (and note that the CB's policy ρ does not affect expected welfare in any other way except for its effect on σ_q^2). The expression for welfare takes different forms if firms pay attention only to aggregate or idiosyncratic shocks (or both), so we consider these cases in turn.

Case 1: $\frac{r\sigma_q}{\sigma_z} \geq 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa})$

If firms pay attention solely to aggregate shocks ($\kappa_A^* = \kappa$), then (following Claim 2.1) prices are set according to:

$$p_i = \underbrace{\frac{r(1 - 2^{-2\kappa})}{1 - (1 - r)(1 - 2^{-2\kappa})}}_{\alpha_H} s_{i,1}$$

So the price level is given by:

$$p = \int_0^1 p_i di = \alpha_H \int_0^1 (q + \varepsilon_{i,1})di = \alpha_H q$$

Recall that the expected welfare loss is given by:

$$E[L] = E[c - c^*] + \delta E[\int_0^1 (p_i - p + z_i)^2 di]$$

The loss due to inefficient consumption variance in this case is given by:

$$E[(c - c^*)^2] = E[(m - p - c^*)^2] = E[(q - \alpha_H q)^2] = (1 - \alpha_H)^2\sigma_q^2$$

And the loss due to an inefficient consumption mix is given by:

$$\delta E[\int_0^1 (p_i - p + z_i)^2 di] = \delta E[\int_0^1 (\alpha_H(s_{i,1} - q) + z_i)^2 di] = \delta(\alpha_H^2\sigma_1^2 + \sigma_z^2) = \delta(\alpha_H^2\frac{\sigma_q^2}{2^{2\kappa} - 1} + \sigma_z^2)$$

Hence, in this case the expected welfare loss is given by:

$$L_H = (1 - \alpha_H)^2\sigma_q^2 + \delta(\alpha_H^2\frac{\sigma_q^2}{2^{2\kappa} - 1} + \sigma_z^2)$$

Case 2: $\frac{r\sigma_q}{\sigma_z} \in (2^{-\kappa}, 2^{-\kappa} + r(2^{\kappa} - 2^{-\kappa}))$

If firms pay attention to both aggregate and idiosyncratic shocks, then by Claim 2.1, prices are set according to:

$$p_i = \underbrace{\frac{r(1-2^{-2\kappa_A})}{1-(1-r)(1-2^{-2\kappa_A})}}_{\alpha_M} s_{i,1} - (1-2^{-2\kappa_I})s_{i,2}$$

Where, by Claim 2.2, $\kappa_A = \frac{1}{2}log_2(1 - \frac{1}{r} + \frac{2^k \sigma_q}{\sigma_z})$ and $\kappa_I = \kappa - \kappa_A$.

Aggregating across firms, the price level is given by:

$$p = \alpha_M q = (1 - \frac{\sigma_z}{2^k r \sigma_q})q$$

Thus, in this case the variance of the real output gap is:

$$E[(c-c^*)^2] = E[(q-p)^2] = E[(1-\alpha_M)^2 q^2] = \left[\frac{\sigma_z}{2^\kappa r \sigma_q}\right]^2 \sigma_q^2 = \left[\frac{\sigma_z}{2^\kappa r}\right]^2$$

Note that this is independent of σ_q^2. The loss due to an inefficient consumption mix simplifies to:

$$\delta E[\int_0^1 (p_i - p + z_i)^2 di] = \frac{\delta \sigma_z^2}{r 2^{2\kappa}}\left[2^{\kappa+1} r \frac{\sigma_q}{\sigma_z} + r - 2\right]$$

Hence, in this case the expected welfare loss is:

$$L_M = \left[\frac{\sigma_z}{2^\kappa r}\right]^2 + \frac{\delta \sigma_z^2}{r 2^{2\kappa}}\left[2^{\kappa+1} r \frac{\sigma_q}{\sigma_z} + r - 2\right]$$

Case 3: $\frac{r\sigma_q}{\sigma_z} \leq 2^{-\kappa}$

If firms pay attention solely to idiosyncratic shocks, then prices are set according to:

$$p_i = -(1-2^{-2\kappa})s_{i,2}$$

And the price level is given by $p = 0$. Consequently, the variance of the real output gap is:

$$E[(c-c^*)^2] = \sigma_q^2$$

And the associated loss due to an inefficient consumption mix is given by:

$$\delta E[\int_0^1 (p_i - p + z_i)^2 di] = \delta \frac{\sigma_z^2}{2^{2\kappa}}$$

Hence, in this case the expected welfare loss is:

$$L_L = \sigma_q^2 + \delta \frac{\sigma_z^2}{2^{2\kappa}}$$

Collecting results from the three cases yields that we can express the expected welfare loss as:

$$E[L] = \begin{cases} (1-\alpha_H)^2\sigma_q^2 + \delta(\alpha_H^2 \frac{\sigma_q^2}{2^{2\kappa}-1} + \sigma_z^2) & \text{if } \frac{r\sigma_q}{\sigma_z} \geq 2^{-k} + r(2^k - 2^{-k}) \\ \left[\frac{\sigma_z}{2^\kappa r}\right]^2 + \frac{\delta\sigma_z^2}{r2^{2\kappa}}\left[2^{\kappa+1} r \frac{\sigma_q}{\sigma_z} + r - 2\right] & \text{if } \frac{r\sigma_q}{\sigma_z} \in (2^{-k}, 2^{-k} + r(2^k - 2^{-k})) \\ \sigma_q^2 + \delta\frac{\sigma_z^2}{2^{2\kappa}} & \text{if } \frac{r\sigma_q}{\sigma_z} \leq 2^{-k} \end{cases}$$

Note that this is (continuous and) strictly increasing in σ_q^2, and does not otherwise depend on the policy coefficient ρ. It follows that the CB sets ρ to minimize σ_q^2. Minimizing (2.75) w.r.t. ρ readily yields:

$$\rho^* = \frac{\sigma_{c^*}^2}{\sigma_{c^*}^2 + \sigma_x^2}$$

And evaluating σ_q^2 at the optimal policy ρ^* gives:

$$\sigma_q^{2*} = \frac{\sigma_{c^*}^2 \sigma_x^2}{\sigma_{c^*}^2 + \sigma_x^2}$$

Note also that this is equal to the CB's posterior uncertainty about c^*. $\square$

2.8.4 Details regarding equilibrium when central bank information is endogenous

Policy rule

In the endogenous CB information section, I assume that the CB commits to the policy rule $m = \rho\tilde{p}$ — i.e. the CB's policy instrument responds to the noisy observation of the *price level* (whereas in the exogenous CB information setup I assume that the monetary instrument responds to the CB's noisy observation of the *fundamental* itself). In the spirit of the exogenous information section, we will consider an alternative policy rule of the form $m = \tilde{\rho}\tilde{c}$ (i.e. a policy rule whereby the monetary instrument responds to the unbiased signal of the fundamental which is constructed by the CB from $\tilde{p}$)[93] and argue that it is equivalent to the policy rule $m = \rho\tilde{p}$.

Firstly, note that Proposition 2.1 and Claim 2.1 still hold in the endogenous CB information setup regardless of the form of the policy rule which the CB commits to, so long as m remains normally distributed (m is determined differently, but is still Gaussian and observable, so firms still find it optimal to observe noisy signals of $q = m - c^*$ and z_i and to use the same pricing rule as before). Hence (using Claim 2.1 and aggregating across firms), for any symmetric allocation of attention $(\kappa_{i,A}, \kappa_{i,I}) = (\overline{\kappa}_A, \overline{\kappa}_I)\forall i$, the price level is given by:

$$p = \underbrace{\frac{r\beta_1}{1-(1-r)\beta_1}}_{\alpha}(m - c^*) \text{ where } \beta_1 = 1 - \frac{1}{2^{2\overline{\kappa}_A}} \tag{2.76}$$

Where $\alpha \in [0, 1)$ is to be determined in equilibrium (because $\overline{\kappa}_A$ is to be determined in equilibrium). Thus, in any equilibrium the CB's signal is:

$$\tilde{p} = \alpha(m - c^*) + \tilde{\varepsilon}$$

For $\alpha \neq 0$, the unbiased estimate of c^* recovered by the CB from $\tilde{p}$ is (as before):

$$\tilde{c} = m - \frac{\tilde{p}}{\alpha} = c^* - \frac{\tilde{\varepsilon}}{\alpha}$$

So our alternative policy rule writes as:

$$m = \tilde{\rho}\tilde{c} = \tilde{\rho}\left(c^* - \frac{\tilde{\varepsilon}}{\alpha}\right) = -\frac{\tilde{\rho}}{\alpha(1-\tilde{\rho})}\tilde{p}$$

[93] Which is denoted by $\tilde{c}$, see (2.32).

While our original policy rule writes as:

$$m = \rho\tilde{p} = \rho[\alpha(m - c^*) + \tilde{\varepsilon}] = -\frac{\rho\alpha}{1-\rho\alpha}(c^* - \frac{\tilde{\varepsilon}}{\alpha})$$

The two policy rules are equivalent if:

$$\rho = -\frac{\tilde{\rho}}{\alpha(1-\tilde{\rho})} \tag{2.77}$$

As mentioned in the main text I restrict attention to $\rho \leq 0$. Note from (2.76) that because $\overline{\kappa}_A \in [0, \kappa]$ it must be the case that $\alpha \in [0, 1)$ in any equilibrium. Hence it follows from (2.77) that if $\rho > 0$ (and there is an equilibrium among firms), then the policy rule $m = \rho\tilde{p}$ is equivalent to a policy rule of the form $m = \tilde{\rho}\tilde{c}$ where $\tilde{\rho} < 0$ or $\tilde{\rho} > 1$. Hence restricting attention to finite $\rho \leq 0$ is equivalent to restricting attention to $\tilde{\rho} \in [0, 1)$. I justify restricting attention to $\tilde{\rho} \in [0, 1)$ on the following grounds:

- It conforms with the optimal reaction coefficient from Claim 2.3 which always lies between 0 and 1. The CB's problem here is similar to the one from the exogenous CB information setup, the only difference being that its information precision concerning the fundamental is now endogenous (and depends on the setting of the policy reaction function) — see also Remark 2.6 below.

- Baeriswyl et al. (2020)[94] also restrict attention to policy rules specifying $\tilde{\rho} \in [0, 1)$, which "means that the central bank seeks to accommodate shocks to the fundamental rather than to amplify them (i.e. $\tilde{\rho} < 0$) or overaccommodate them (i.e. $\tilde{\rho} > 1$), which sounds realistic" (Baeriswyl et al. (2020))

Note 1: *The argument above is somewhat incomplete as it presupposes that for any $\rho \leq 0$ there is a unique equilibrium allocation of attention (and hence a unique α associated with each $\rho \leq 0$) — I argue in what follows that this is indeed the case.*

Note 2: *To compute $\tilde{\rho}$ (for the graphs in the main text), I fix different values for ρ and compute the equilibrium among firms (and α), then use equation (2.77) to compute $\tilde{\rho}$.*

Remark 2.6: It should also be WLOG to restrict attention to $\tilde{\rho} \in [0, 1)$.

Sketch of argument: Given a policy rule of the form $m = \tilde{\rho}\tilde{c}$, the variance of q

[94] See section 1.6.

in any equilibrium is:

$$\sigma_q^2 = var(m - c^*) = (\tilde{\rho} - 1)^2\sigma_{c^*}^2 + \tilde{\rho}^2\sigma_{\tilde{p}}^2/\alpha^2 \tag{2.78}$$

Consider any $\tilde{\rho} < 0$ — for any $\alpha \in [0, 1)$, the variance of the aggregate shock q is higher than the variance of the fundamental $\sigma_q^2 > \sigma_{c^*}^2$. From the welfare analysis in the exogenous CB information setup (Claim 2.3), we know that the expected welfare loss is strictly increasing in σ_q^2. It follows that a policy rule specifying $\tilde{\rho} = 0$ leads to lower welfare losses in equilibrium than one specifying $\tilde{\rho} < 0$ (i.e. the CB does not find it optimal to amplify shocks to the fundamental).

For any $\tilde{\rho} > 1$, one can show that overaccommodating less (i.e. decreasing $\tilde{\rho}$) leads to lower welfare losses in equilibrium (i.e. for $\tilde{\rho} > 1$ a policy rule specifying $m = (\tilde{\rho} - \varepsilon)\tilde{c}$ does strictly better than a policy rule specifying $m = \tilde{\rho}\tilde{c}$ (where ε is small), so no $\tilde{\rho} > 1$ can be optimal). To see this more clearly, consider the derivative of (2.78) w.r.t. $\tilde{\rho}$:

$$\frac{\partial \sigma_q^2}{\partial \tilde{\rho}} = 2(\tilde{\rho} - 1)\sigma_{c^*}^2 + 2\tilde{\rho}\sigma_{\tilde{p}}^2/\alpha^2 - 2\tilde{\rho}^2\sigma_{\tilde{p}}^2\frac{1}{\alpha^3}\frac{\partial \alpha}{\partial \tilde{\rho}} \tag{2.79}$$

Using the chain rule we can express $\frac{\partial \alpha}{\partial \tilde{\rho}} = \frac{\partial \alpha}{\partial \sigma_q^2}\frac{\partial \sigma_q^2}{\partial \tilde{\rho}}$, which implies that (2.79) can be rearranged as:

$$\frac{\partial \sigma_q^2}{\partial \tilde{\rho}} = \left[2(\tilde{\rho} - 1)\sigma_{c^*}^2 + \frac{2\tilde{\rho}\sigma_{\tilde{p}}^2}{\alpha^2}\right] / \left[1 + 2\tilde{\rho}\sigma_{\tilde{p}}^2\frac{1}{\alpha^3}\frac{\partial \alpha}{\partial \sigma_q^2}\right] \tag{2.80}$$

For $\tilde{\rho} > 1$, both the numerator and denominator of the fraction on the R.H.S. of (2.80) are strictly positive (because $\frac{\partial \alpha}{\partial \sigma_q^2} \geq 0$ (see 2.31) and because $\alpha \in [0, 1)$ in any equilibrium). It follows that $\frac{\partial \sigma_q^2}{\partial \tilde{\rho}} > 0$ for $\tilde{\rho} > 1$, so overaccommodating less (slightly lowering $\tilde{\rho}$) is welfare-improving.

Optimal policy

Recall (from (2.76)) that we can express the CB's signal as:

$$\tilde{p} = \alpha(m - c^*) + \tilde{\varepsilon}$$

Using the CB's policy rule $m = \rho\tilde{p}$, we can express q as:

$$q = m - c^* = \rho\left[\alpha(m - c^*) + \tilde{\varepsilon}\right] - c^* \Leftrightarrow$$
$$\Leftrightarrow q = -\frac{1}{1-\rho\alpha}c^* + \frac{\rho}{1-\rho\alpha}\tilde{\varepsilon}$$

(The denominator is strictly positive because we have restricted attention to $\rho \leq 0$ and because $\alpha \geq 0$ in any equilibrium.) Squaring both sides and taking unconditional expectations yields:

$$\sigma_q^2 = \left(\frac{1}{1-\rho\alpha}\right)^2 \sigma_{c^*}^2 + \left(\frac{\rho}{1-\rho\alpha}\right)^2 \sigma_{\tilde{p}}^2 \tag{2.81}$$

Note that the variance of aggregate shocks in equilibrium is a function of α. Also, α is itself a function of σ_q^2 (as it depends on how much attention firms pay to aggregate shocks).

As mentioned in the main text, in equilibrium firms pay attention to both aggregate and idiosyncratic shocks — I argue this more precisely (and alleviate potential concerns) afterwards in Appendix 2.8.4. For now, let us just guess and verify that the equilibrium allocation of attention is interior — in this case, it follows from Claim 2.2 that the amount of attention allocated by firms to aggregate shocks (in equilibrium) is given by:

$$\kappa_A = \frac{1}{2}log_2\left(1 - \frac{1}{r} + \frac{2^\kappa\sigma_q}{\sigma_z}\right)$$

Using this and (2.76) yields (as before):

$$\alpha = 1 - \frac{\sigma_z}{2^\kappa\sigma_q r} \tag{2.82}$$

Substituting (2.82) into (2.81) gives:

$$\sigma_q^2 = \frac{2^{2\kappa}\sigma_q^2 r^2(\sigma_{c^*}^2 + \rho^2\sigma_{\tilde{p}}^2)}{\left[2^\kappa\sigma_q r(1-\rho) + \rho\sigma_z\right]^2}$$

Rearranging yields:

$$\left[2^\kappa\sigma_q r + \rho(\sigma_z - 2^\kappa\sigma_q r)\right]^2 = 2^{2\kappa}r^2(\sigma_{c^*}^2 + \rho^2\sigma_{\tilde{p}}^2) \tag{2.83}$$

Note that because firms pay some attention to aggregate shocks ($\kappa_A > 0$), it must be the case that $\sigma_z < 2^\kappa\sigma_q r$. Recalling that we restrict attention to $\rho \leq 0$ implies that the squared term on the L.H.S. of (2.83) is strictly positive. Taking square roots of both sides of (2.83) and rearranging gives us an expression for σ_q as a function of policy and parameters:

$$\sigma_q = \frac{2^\kappa r\sqrt{\sigma_{c^*}^2 + \rho^2\sigma_{\tilde{p}}^2} - \rho\sigma_z}{2^\kappa r(1-\rho)} \tag{2.84}$$

It follows from before (Claim 2.3) that the welfare loss is strictly increasing in σ_q, so the CB seeks to minimize it.[95] Taking the derivative of (2.84) w.r.t. ρ (and simplifying) yields:

$$\frac{\partial\sigma_q}{\partial\rho} = \left[2^\kappa r(\sigma_{c^*}^2 + \rho\sigma_{\tilde{p}}^2) - \sigma_z\sqrt{\sigma_{c^*}^2 + \rho^2\sigma_{\tilde{p}}^2}\right] / \left[2^\kappa r(\rho-1)^2\sqrt{\sigma_{c^*}^2 + \rho^2\sigma_{\tilde{p}}^2}\right] \tag{2.85}$$

Setting the first-order condition equal to zero is equivalent to:

$$\frac{\sigma_{c^*}^2 + \rho^*\sigma_{\tilde{p}}^2}{\sqrt{\sigma_{c^*}^2 + \rho^{*2}\sigma_{\tilde{p}}^2}} = \frac{\sigma_z}{2^\kappa r} \tag{2.86}$$

Let $f(\rho) = \frac{\sigma_{c^*}^2 + \rho\sigma_{\tilde{p}}^2}{\sqrt{\sigma_{c^*}^2 + \rho^2\sigma_{\tilde{p}}^2}}$. Note that:

- $f(0) = \sigma_{c^*} > \frac{\sigma_z}{2^\kappa r}$ (because in the absence of policy intervention ($\rho = 0$), the equilibrium allocation of attention is interior)
- $f(-\frac{\sigma_{c^*}^2}{\sigma_{\tilde{p}}^2}) = 0 < \frac{\sigma_z}{2^\kappa r}$
- $f(\rho)$ is continuous and strictly increasing in ρ (for $\rho \leq 0$)

It follows that there is a unique solution to $f(\rho) = \frac{\sigma_z}{2^\kappa r}$ (i.e. (2.86)), which is the optimal policy.[96] Note that because $f(\rho)$ is increasing in ρ (for $\rho \leq 0$), it follows that $\frac{\partial\sigma_q}{\partial\rho} > 0$ for $\rho > \rho^*$ and $\frac{\partial\sigma_q}{\partial\rho} < 0$ for $\rho < \rho^*$ so we have indeed found a (global) minimum.

We should check that when policy is set optimally ($\rho = \rho^*$) firms indeed find it optimal to pay attention to both aggregate and idiosyncratic shocks in equilibrium (such that our guess is verified) — I argue this in Appendix 2.8.4 (see (2.90) and the discussion following it).

Corollary 2.1 follows immediately from our previous analysis. Because the equilibrium allocation of attention is interior, α is equal to α_M (see 2.82) — hence α

[95] More precisely, because the equilibrium allocation of attention is interior, the welfare loss is given by $E[L] = \left[\frac{\sigma_z}{2^\kappa r}\right]^2 + \frac{\delta\sigma_z^2}{r2^{2\kappa}}\left[2^{\kappa+1}r\frac{\sigma_q}{\sigma_z} + r - 2\right]$ — see case 2 in Appendix 2.8.3.

[96] Also, $\rho^* \in (-\frac{\sigma_{c^*}^2}{\sigma_{\tilde{p}}^2}, 0)$.

is strictly increasing in σ_q. Because optimal policy minimizes σ_q, it also minimizes α. It follows that optimal policy minimizes price informativeness ($\tau_p = \alpha^2/\sigma^2_{\tilde{p}}$), the slope of the Phillips curve and the degree of monetary neutrality.

To analyse comparative statics, note that for $\rho^* < 0$, the L.H.S. of (2.86) is strictly increasing in ρ^* and $\sigma^2_{c^*}$, and strictly decreasing in $\sigma^2_{\tilde{p}}$. Hence it is straightforward to prove Corollary 2.2 by taking implicit derivatives (or inspecting) (2.86).

Using (2.84) and (2.86) we can express:

$$\sigma^*_q = \frac{\sigma^2_{c^*}}{\sqrt{\sigma^2_{c^*} + \rho^{*2}\sigma^2_{\tilde{p}}}} \tag{2.87}$$

Proof that equilibrium allocation of attention is interior

Let us first consider equilibria where firms pay attention only to idiosyncratic shocks. Suppose $(\overline{\kappa}_A, \overline{\kappa}_I) = (0, \kappa)$. By Claim 2.1, prices are set according to $p_j = \beta_2 s_{j,2}$, so the price level is $p = 0$. Thus, firm i's problem writes as:

$$\min_{\kappa_A, \kappa_I} var[rq - z_i | S_i]$$

$$\text{Subject to: } S_i = \begin{pmatrix} q + \varepsilon_{i,1} \\ z_i + \varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma^2_q}{2^{\kappa_A}-1} & 0 \\ 0 & \frac{\sigma^2_z}{2^{\kappa_I}-1} \end{pmatrix};$$

$$\kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A; \kappa_I) \in [0, \kappa]^2.$$

By Lemma 2, firm i finds it optimal to choose $(\kappa^*_{i,A}, \kappa^*_{i,I}) = (0, \kappa)$ if $\frac{r\sigma_q}{\sigma_z} \leq 2^{-\kappa}$. Note that because $p = 0$, the CB's signal is pure noise $\tilde{p} = \tilde{\varepsilon}$. This implies that the variance of the aggregate shock $q = m - c^*$ is greater than (or equal to) the variance of c^* for any policy reaction function: $\sigma^2_q \geq \sigma^2_{c^*} \forall \rho$. As $\frac{r\sigma_{c^*}}{\sigma_z} > 2^{-\kappa}$ (because in the absence of policy intervention, the equilibrium allocation of attention is interior), it follows that the inequality $\frac{r\sigma_q}{\sigma_z} \leq 2^{-\kappa}$ cannot be satisfied in equilibrium for any ρ — hence there are no equilibria where firms pay attention only to idiosyncratic shocks.

Secondly, consider equilibria where firms pay attention only to aggregate conditions. Suppose $(\overline{\kappa}_A, \overline{\kappa}_I) = (\kappa, 0)$. By Claim 2.1, prices are set according to $p_j = \frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})} s_{j,1}$, so the price level is given by $p = \underbrace{\frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})}}_{\alpha_H} q$. Thus, firm i's problem writes as:

$$\min_{\kappa_A,\kappa_I} var[(r+(1-r)\alpha_H)q - z_i|S_i]$$
$$\text{Subject to: } S_i = \begin{pmatrix} q+\varepsilon_{i,1} \\ z_i+\varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma_q^2}{2^{\kappa_A}-1} & 0 \\ 0 & \frac{\sigma_z^2}{2^{\kappa_I}-1} \end{pmatrix};$$
$$\kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A;\kappa_I) \in [0,\kappa]^2.$$

By Lemma 2, firm i finds it optimal to choose $(\kappa_{i,A}^*, \kappa_{i,I}^*) = (\kappa, 0)$ if $\frac{[r+(1-r)\alpha_H]\sigma_q}{\sigma_z} \geq 2^\kappa$. Let us now determine σ_q in this case (and denote it by $\sigma_{q,H}$). Using (2.81), we can express:

$$\sigma_{q,H}^2 = \left(\frac{1}{1-\rho\alpha_H}\right)^2 \sigma_{c^*}^2 + \left(\frac{\rho}{1-\rho\alpha_H}\right)^2 \sigma_{\tilde{p}}^2.$$

Hence, there is an equilibrium where firms pay attention solely to aggregate shocks if:

$$[r+(1-r)\alpha_H]\,\sigma_{q,H} \geq 2^\kappa \sigma_z \tag{2.88}$$

Lastly, consider equilibria where firms pay attention to both aggregate and idiosyncratic shocks. Firm i's problem reads:

$$\min_{\kappa_A,\kappa_I} var[(r+(1-r)\alpha_M)q - z_i|S_i]$$
$$\text{Subject to: } S_i = \begin{pmatrix} q+\varepsilon_{i,1} \\ z_i+\varepsilon_{i,2} \end{pmatrix}; \quad \Omega_{\varepsilon i} = \begin{pmatrix} \frac{\sigma_q^2}{2^{\kappa_A}-1} & 0 \\ 0 & \frac{\sigma_z^2}{2^{\kappa_I}-1} \end{pmatrix};$$
$$\kappa_A + \kappa_I \leq \kappa; \quad (\kappa_A;\kappa_I) \in [0,\kappa]^2.$$

Where $\alpha_M = 1 - \frac{\sigma_z}{2^\kappa r \sigma_{q,M}}$ and $\sigma_{q,M} = \frac{2^\kappa r\sqrt{\sigma_{c^*}^2+\rho^2\sigma_{\tilde{p}}^2}-\rho\sigma_z}{2^\kappa r(1-\rho)}$ (because we posited that the firms' equilibrium allocation of attention is interior — see equations (2.82) and (2.84)). Using Lemma 2 again, firm i finds it optimal to pay attention to both aggregate and idiosyncratic shocks if:

$$2^{-\kappa} < [r+(1-r)\alpha_M]\sigma_{q,M}/\sigma_z < 2^\kappa \tag{2.89}$$

Inequalities (2.88) and (2.89) can never be satisfied simultaneously (thus, we never get multiple equilibria for any ρ).

The last thing to show is that inequality (2.89) is satisfied when policy is set optimally ($\rho = \rho^*$). To be more precise, we want to show that:

$$2^{-\kappa} < [r+(1-r)\alpha_M^*]\sigma_{q,M}^*/\sigma_z < 2^\kappa \tag{2.90}$$

Where $\alpha_M^* = 1 - \frac{\sigma_z}{2^\kappa r \sigma_{q,M}^*}$, $\sigma_{q,M}^* = \frac{\sigma_{c^*}^2}{\sqrt{\sigma_{c^*}^2 + \rho^{*2}\sigma_{\tilde{p}}^2}}$ (see (2.87)) and where ρ^* solves (2.86).

Firstly, note that:

$$[r + (1-r)\alpha_M^*]\sigma_q^*/\sigma_z < [r + (1-r)\alpha_H]\sigma_{c^*}/\sigma_z < 2^\kappa \tag{2.91}$$

Where the first inequality holds because $\alpha_M^* < \alpha_H$ and $\sigma_q^* < \sigma_{c^*}$. The second inequality holds because in the absence of policy intervention ($\rho = 0$), the equilibrium allocation of attention is interior.

Secondly, note that:

$$2^{-\kappa} < [r + (1-r)\alpha_M]\sigma_{q,M}/\sigma_z \quad \forall \rho \leq 0 \tag{2.92}$$

From (2.91) and (2.92), it follows that (2.90) is satisfied (hence our guess that the equilibrium allocation of attention is interior is verified, and it follows that the unique equilibrium we have identified always exists). $\square$

3 Brainard Uncertainty and the Signal Value of Prices

3.1 Introduction

One rationale standing behind caution in monetary policy-making is parameter uncertainty in the spirit of Brainard (1967) — in other words, uncertainty regarding the transmission of policy itself.[1] This is also referred to as "multiplicative uncertainty", and is to be distinguished from "additive uncertainty" which is independent of the policymaker's behavior:[2]

"Researchers have generally specified models in which uncertainty is independent of the policymaker's behavior. In these models, the only uncertainty is whether the economy will deviate from the path policy-makers expect on account of what are known as 'additive shocks'. As Theil (1958) showed, the best that policy-makers could do in this case would be to ignore the effects of uncertainty upon the economy. This is known as 'certainty-equivalence'." (Batini et al. (1999))

Note that in Chapter 2, the central bank faces only "additive" uncertainty (concerning the fundamental).[3] In this paper, I introduce parameter uncertainty in a stylized, dynamic version of the model presented in Chapter 2, then I analyse similar questions (concerning the effects of policy intervention on price informativeness, and the consequent implications for optimal policy).

Building on Chapter 2, I first analyse a sequence of static beauty contest games featuring an (infinitely-lived) activist policymaker who learns about fundamentals by observing noisy signals of the agents' average actions, in the spirit of Morris and Shin (2005). Each period, a new generation of (short-lived) rationally inattentive agents are born, who live for one period in which they process a fixed amount of information about their environment, then take their actions.[4] Firstly, the paper argues that the central bank can use the information revealed by the average actions

[1] Throughout the paper I use the terms "parameter uncertainty", "Brainard uncertainty" and "uncertainty regarding the transmission of policy" interchangeably.

[2] More details are presented in section 3.3.

[3] For instance, in section 2.4, it is easy to note that the central bank faces "additive" uncertainty and that its optimal policy displays certainty-equivalence — see, for instance Claim 2.3 and note that the policy instrument is optimally set according to $m = E[c^*|x]$. In contrast, under parameter uncertainty, the central bank no longer acts solely on the basis of expected values, so optimal policy no longer displays certainty-equivalence (for more details, see section 3.3.1).

[4] Each private-sector agent's problem is equivalent to the firm's problem in Chapter 2 (if there are no idiosyncratic shocks in the model of Chapter 2, i.e. if $\sigma_z^2 = 0$).

of agents in past generations to improve the expected welfare of agents in the current generation. Furthermore, it argues that, in such a setting, the policymaker's actions do not affect his information precision about fundamentals, because the central bank can always perfectly disentangle its policy instrument from the average action.[5]

I then relax the assumption that the central bank perfectly knows the parameters governing the transmission of its policy (i.e. I introduce parameter uncertainty in the spirit of Brainard (1967)). I argue that under parameter uncertainty, the central bank can no longer perfectly disentangle the effects of its policy from the average action, because the latter necessarily reflects some of the uncertainty associated with the transmission of policy (alongside information concerning the fundamental). Not perfectly knowing either the fundamental or the transmission of policy, the policymaker learns about both upon observing the agents' average action. This implies that "policy experiments" designed to elicit information concerning policy transmission (as described in Bertocchi and Spagat (1993)) are informationally costly for the central bank in this framework, as they necessarily come at the expense of information concerning fundamentals. Parameter uncertainty complicates the central bank's problem, as a trade-off emerges between learning about the fundamental and learning about the parameter governing the transmission of policy — the central bank has both an incentive to experiment with policy (in order to elicit information about its transmission) and an incentive to be cautious with policy (in order to have more precise information about the fundamental). This resembles the problem of "learning with two unobservable parameters" in Balvers and Cosimano (1994), as I explain in more detail in section 3.3.4.

Related literature: The paper relates to the literature on central bank learning under parameter uncertainty (see, for instance, Bertocchi and Spagat (1993); Balvers and Cosimano (1994); Sack (1998); Ellison and Valla (2001)). I expand on this in section 3.3.3, after I introduce the model. The paper also relates to the literature analysing optimal policy in settings where firms set prices under imperfect common knowledge, and the central bank learns from prices (for instance, Morris and Shin (2005); Baeriswyl (2011); Baeriswyl et al. (2020)). For more details about this strand of the literature, see section 2.2 in Chapter 2.

[5]This is similar to the observation (in Chapter 2) that price informativeness is independent of the central bank's reaction function so long as firms pay a fixed amount of attention to aggregate shocks — see the discussion in section 2.4.2.

3.2 Certainty-equivalent model

Agents and payoffs: Consider a sequence of static beauty contest games in the spirit of Morris and Shin (2005). Time is discrete and indexed by $t \in \{0, 1, 2...\}$. Each period t, a new generation of private-sector agents (indexed by i on the unit interval) play a beauty contest game. More specifically, each agent i in generation t chooses his action $a_{i,t} \in \mathbb{R}$ to match a weighted sum of a Gaussian fundamental θ_t, as well as the average action of the agents in that period $\overline{a}_t = \int_0^1 a_{i,t} di$. Additionally, each agent cares about matching the central bank's policy instrument, denoted by g_t. More precisely, each agent's utility writes as:

$$u_{i,t} = -\left[a_{i,t} - r(\theta_t + g_t) - (1 - r)\overline{a}_t\right]^2 \tag{3.1}$$

Where the degree of strategic complementarity is $(1-r)$. The way in which policy is incorporated in the model is reminiscent of James and Lawler (2011) — "it implies, of course, that appropriate adjustments in g can fully neutralize the consequences of variations in θ. Such a formulation is likely to be especially relevant in a macroeconomic context, where θ might, for example, be taken to correspond to a particular aggregate demand shock realization" (James and Lawler (2011)). Equation (3.1) can also be derived in the context of a microfounded economy, as in Chapter 2.[6]

Initially, the fundamental is drawn from a Gaussian prior $\theta_{-1} \sim N(0, \sigma^2_{\theta,-1})$; then, it evolves according to an AR(1):

$$\theta_t = \alpha\theta_{t-1} + \varepsilon^\theta_t \tag{3.2}$$

Where $\alpha \in [0, 1]$ and the ε^θ_t's are independently and identically distributed according to $\varepsilon^\theta_t \sim N(0, \sigma^2_\theta)$.

As mentioned above, the central bank controls the policy instrument g_t. The policymaker will have some private information (concerning the fundamental) and will choose a reaction function specifying how he sets the policy instrument (g_t) conditional on his information (this is introduced in the next subsection). The policymaker's per-period payoff is given by:

$$u_{cb,t} = -(\overline{a}_t - \theta_t - g_t)^2 - \lambda \int_0^1 (a_{i,t} - \overline{a}_t)^2 di \tag{3.3}$$

[6] Where the agent's utility would correspond to a log-quadratic approximation of a monopolistically competitive firm's profit function in an economy in which the household consumes a composite good a la Dixit-Stiglitz and the central bank controls the money supply.

The central bank's per-period payoff is decreasing in the distance of the average action from the sum of the fundamental and the policy instrument in that period (i.e. the first squared term), as well as in the dispersion of actions from the average action (i.e. the second term). The weight assigned to action dispersion in the welfare function is given by $\lambda > 0$. We will note in what follows that the exact value of λ does not matter for the results[7] — remark however that $\lambda > 1$ if the welfare function is derived from microfoundations, while in a beauty contest formulation of the problem in the spirit of Morris and Shin (2005), $\lambda = 1$. This is also discussed in Baeriswyl et al. (2020) where the same objective function is used for the policymaker.

Information structure: The assumptions concerning the information structure are similar to the ones in Chapter 2. More specifically, I assume that private sector agents are rationally inattentive (and can directly observe the fundamental, as well the setting of the policy instrument), while the central bank learns about the fundamental by observing noisy signals of the agents' average actions (as in Morris and Shin (2005)).

Private sector agents: All agents in all generations have the same finite capacity to process information, denoted by κ. In similar fashion to Chapter 2, agents in each period process information before taking their actions. The setting of the policy instrument in period t (g_t) is observable to private-sector agents in generation t. In what follows we will also note that g_t is Gaussian and each private agent knows the distribution from which it is drawn. Define the vector $X_t = (\theta_t, g_t)^T$ and let $S_{i,t}$ denote the vector of (private) signals observed by agent i in period t. The information processing constraint writes as:

$$I(X_t, S_{i,t}) \leq \kappa$$

Which states that the mutual information between X_t and the vector of signals observed by each agent i in generation t ($S_{i,t}$) must not exceed the agent's capacity to process information (κ) — for more details, see section 2.3 in chapter 2. I assume that each signal observed by agents can be about any linear combination of current shocks (i.e. θ_t and g_t). Hence, we can represent the signal vector of agent i in generation t as:

$$S_{i,t} = F_{i,t} X_t + \varepsilon_{i,t} \tag{3.4}$$

[7] As the central bank will not face a trade-off between reducing the first and second components of the loss function.

Where $F_{i,t}$ is a $K_{i,t}$ x 2 matrix of coefficients and $\varepsilon_{i,t} = (\varepsilon_{1,it}, \varepsilon_{2,it}, ..., \varepsilon_{K_{i,t},it})^T$ is a Gaussian white noise random vector (independent of X_t and all other random variables) with covariance matrix $\Omega_{i,t}$. As previously mentioned, $X_t = (\theta_t, g_t)^T$. As in chapter 2, choosing the properties of the signal vector $S_{i,t}$ entails specifying the number of signals to observe ($K_{i,t}$), the content of these signals ($F_{i,t}$) and the covariance matrix of noise in the signals ($\Omega_{i,t}$).

Taking into account that each agent acts optimally conditional on the information received (i.e. he chooses $a_{i,t}$ to maximize (3.1)), we can write agent i's problem (in generation t) when choosing his allocation of attention as:

$$\begin{aligned} &\min_{K_{i,t},F_{i,t},\Omega_{i,t}} E\left[a^*_{i,t} - r(\theta_t + g_t) - (1-r)\overline{a}_t\right]^2 \\ &\text{Subject to: } a^*_{i,t} = E[r(\theta_t + g_t) + (1-r)\overline{a}_t | S_{i,t}], \\ &\qquad S_{i,t} = F_{i,t} X_t + \varepsilon_{i,t}, \\ &\qquad \varepsilon_{i,t} \sim N(0, \Omega_{i,t}), \\ &\qquad I(S_{i,t}, X_t) \leq \kappa. \end{aligned} \tag{3.5}$$

When choosing his allocation of attention, each agent takes as given the central bank's reaction function (which will imply that each agent takes as given the distribution of g_t, as will be discussed in what follows), as well as the other agents' allocations of attention. He anticipates how he will optimally choose his action conditional on the information he receives[8] and takes into account how his decision regarding the allocation of attention will affect his expected utility.

Central bank: As in Morris and Shin (2005), in each period t, the central bank observes a noisy signal of the agents' average action in period $t-1$. In period t, the policymaker observes the signal:

$$Z_t = \overline{a}_{t-1} + \varepsilon^Z_t \tag{3.6}$$

Where the ε^Z_t's are Gaussian noise terms distributed independently of each other and all other random variables according to $\varepsilon^Z_t \sim N(0, \sigma^2_z)$. The central bank's information set in period t is the collection of all past signals up to period t:

$$I^{cb}_t = \{Z_t, Z_{t-1}, Z_{t-2}, ..., Z_1\}$$

[8]He also anticipates how other agents will choose their actions (conditional on their information) in equilibrium — given the other agents' allocation of attention, each agent can compute $\overline{a}_t$ as a linear combination of θ_t and g_t when choosing his own allocation of attention.

In the spirit of Morris and Shin (2005), we will be interested in the central bank's information precision about the fundamental over time. We will denote by τ_t^{cb} the central bank's information precision about the fundamental in period t, as measured by:

$$\begin{aligned} \tau_t^{cb} &= 1/var(\theta_t | I_t^{cb}) \\ &= 1/(var | Z_t, Z_{t-1}, Z_{t-2}, ...) \end{aligned} \tag{3.7}$$

The central bank chooses a policy rule specifying how it sets the policy instrument g_t conditional on its information in period t. I restrict attention to linear policy rules of the form:

$$\begin{aligned} g_t =& \rho_{t,t} Z_t + \rho_{t,t-1} Z_{t-1} + \rho_{t,t-2} Z_{t-2} + ... + \rho_{t,1} Z_1 \\ =& \sum_{k=0}^{t-1} \rho_{t,t-k} Z_{t-k} \end{aligned} \tag{3.8}$$

Where $\rho_{t,t-k}$ denotes the responsiveness of the policy instrument in period t to the signal observed by the central bank in period $t-k$. Hence in each period t, the central bank's policy rule (denoted $\boldsymbol{\rho}_t$) specifies a vector of coefficients:

$$\boldsymbol{\rho}_t = (\rho_{t,t}, \rho_{t,t-1}, \rho_{t,t-2}, ..., \rho_{t,1}) \tag{3.9}$$

For ease of notation, denote by $S_{cb,t}$ the (column) vector of signals observed by the central bank up to period t:

$$S_{cb,t} = (Z_t, Z_{t-1}, ..., Z_1)^T \tag{3.10}$$

Such that we can write the central bank's policy instrument (in period t) more compactly as:

$$g_t = \boldsymbol{\rho}_t S_{cb,t} \tag{3.11}$$

When choosing its policy rule, the central bank takes into account how this affects the private sector agents' equilibrium behavior in each period t (and thus, expected welfare in period t, as well as the incoming signal in period $t+1$ which is centered on the private-sector agents' average action in period t).

Remark 3.1. *We will note that — given the private sector agents' equilibrium behavior — the central bank indeed finds it optimal to use a policy rule which is* *linear in its private signals.*[9]

In turn, because the central bank sets its policy as a linear function of its signals (which are Gaussian), it follows that g_t is normally distributed each period. As each private agent takes as given the central bank's reaction function (as well as the equilibrium strategies of all agents in previous generations, such that the distributions of $Z_{t-1}, Z_{t-2}, ...$ are common knowledge in period t), each agent in generation t takes as given the distribution of g_t. Furthermore, because g_t and θ_t are Gaussian, and payoffs (3.1) are quadratic, each private agent finds it optimal to learn about shocks by observing Gaussian signals.

In short, the remark states that all agents in the game (referring to both the central bank and private sector agents) will be best-responding against the other players' strategies. If the central bank sets policy as a linear function of its signals, private sector agents find it optimal to learn about shocks by observing Gaussian signals, and, in turn, if private sector agents learn about shocks by observing Gaussian signals (and act optimally given their information), the central bank will find it optimal to set policy as a linear function of its private signals.

Timing: In period $t = 0$, the central bank first chooses its reaction function:[10]

$$\{\boldsymbol{\rho}_t\}_{t=1}^{\infty} = \{(\rho_{t,t}, \rho_{t,t-1}, \rho_{t,t-2}, ..., \rho_{t,1})\}_{t=1}^{\infty}.$$

Then, in each period $t \geq 0$:

1. Each private agent chooses his allocation of attention: $(K_{i,t}, F_{i,t}, \Omega_{i,t})$
2. The central bank observes its private signal (Z_t) and sets policy g_t according to the policy rule ($g_t = \boldsymbol{\rho}_t S_{cb,t}$)
3. The fundamental (θ_t) is realized

[9] More precisely, the central bank will find it optimal to set $g_t = -E[\theta_t | I_t^{cb}]$ and $E[\theta_t | I_t^{cb}]$ will be a linear combination of the central bank's signals ($Z_t, Z_{t-1}, Z_{t-2}, ...$); see Proposition 3.2.

[10] It will turn out to be irrelevant whether we assume that the central bank chooses its reaction function (for all future periods) at time $t = 0$, or if we allow the central bank to choose its reaction function period-by-period. The reason is that the central bank's per-period loss in period t will be proportional to the variance of the sum of the policy instrument and the fundamental in period t (see Corollary 3.1), so the central bank will (optimally) set its policy instrument equal to (minus) its expectation of the fundamental conditional on its information up to period t — in turn, this is a (deterministic) linear combination of its signals up to period t, in which the weight associated with each signal is deterministic (i.e. independent of the realizations of signals up to period t). This happens because in a Gaussian environment, conditional variances (and the weights in the linear rules which agents optimally use to form their expectations, conditional on their Gaussian signals) are independent of signal realizations.

4. Each private agent observes his signal vector ($S_{i,t}$) and takes his action $a_{i,t}$

Equilibrium definition: An equilibrium specifies:

1. The central bank's policy rule $\{\boldsymbol{\rho}_t\}_{t=1}^{\infty}$ (mapping the central bank's signal realizations up to period t to the policy instrument in period t (for all t))

2. An allocation of attention for each agent in each generation $(K_{i,t}^*, F_{i,t}^*, \Omega_{i,t}^*)$,

3. An action rule for each agent in each generation (mapping signal realizations to his action $a_{i,t}^*$)

Such that:

1. The central bank's policy is optimal (i.e. it solves problem (3.12) below)

2. Each agent's allocation of attention maximizes his expected utility subject to the information processing constraint

   (i.e. $(K_{i,t}^*, F_{i,t}^*, \Omega_{i,t}^*)$ solves problem (3.5) for all i and all t)

3. Each agent acts optimally given his information

   (i.e. $a_{i,t}^* = E[r(\theta_t + g_t) + (1-r)\overline{a}_t | S_{i,t}]$ for all i and all t)

As previously mentioned, when choosing his allocation of attention, each agent takes as given the other agents' allocations of attention and the central bank's policy rule, and he anticipates how other agents will act conditional on their information (in step 5 of each period) — this will allow us to get an expression for agent i's expected utility as a function of his allocation of attention (in step 2 of each period).[11] I analyse symmetric equilibria in which all agents in the same generation choose the same allocation of attention:

$$(K_{i,t}^*, F_{i,t}^*, \Omega_{i,t}^*) = (K_t^*, F_t^*, \Omega_t^*) \forall i.$$

Let us now turn our attention to the central bank's policy rule — as previously mentioned, the policymaker takes into account how his reaction function shapes the agents' equilibrium behavior and, in turn, how this affects expected welfare, as well as the central bank's own information in future periods. More specifically, when choosing the policy reaction function (in period $t = 0$), the central bank's problem reads:

[11] Note that there is a unique equilibrium in step 5 for any allocation of attention chosen in step 2. See Chapter 2.

$$\max_{\{\boldsymbol{\rho}_t\}_{t=1}^{\infty}} \quad -\sum_{t=1}^{\infty} \delta^t E\left[(\overline{a}_t - \theta_t - g_t)^2 + \lambda \int_0^1 (a_{i,t}^* - \overline{a}_t)^2 di\right] \tag{3.12}$$

Subject to:

$$g_t = \boldsymbol{\rho}_t S_{cb,t} \tag{3.13}$$

$$S_{cb,t} = (Z_t, Z_{t-1}, ..., Z_1)^T \tag{3.14}$$

$$Z_t = \overline{a}_{t-1} + \varepsilon_t^Z \tag{3.15}$$

$$\overline{a}_t = \int_0^1 a_{i,t}^* di \tag{3.16}$$

$$a_{i,t}^* = E[r(\theta_t + g_t) + (1-r)\overline{a}_t | S_{i,t}] \tag{3.17}$$

$$S_{i,t} = F_{i,t}^* X_t + \varepsilon_{i,t} \tag{3.18}$$

$$\varepsilon_{i,t} \sim N(0, \Omega_{i,t}^*) \tag{3.19}$$

$$(K_{i,t}^*, F_{i,t}^*, \Omega_{i,t}^*) \text{ solves problem (3.5)} \tag{3.20}$$

Where

$$\theta_t = \alpha\theta_{t-1} + \varepsilon_t^\theta \tag{3.21}$$

And where $\delta \in [0, 1]$ denotes the central bank's discount factor.

The policymaker chooses his reaction function $\{\boldsymbol{\rho}_t\}_{t=1}^{\infty}$ to maximize the discounted sum of its expected per-period payoffs (3.12) (where the central bank's per-period payoff is specified in (3.3)). The reaction function maps the central bank's signal vector in period t to the policy instrument in period t as specified in (3.13). (3.14) states that the central bank's signal vector in period t is the collection of all signals observed up to period t, which are centered on the private sector agents' average actions (see (3.15)). Furthermore, the policymaker takes as given the agents' equilibrium behavior (conditional on his reaction function) — more formally, equilibrium conditions 2 and 3 above are taken as constraints in the central bank's optimization problem. (3.17) states that the central bank takes as given that each agent acts optimally conditional on his information, while (3.18)-(3.20) state that the central bank takes as given that each private-sector agent chooses his allocation of attention optimally to maximize his expected utility. (3.21) restates the law of motion for the fundamental.

I first characterise the agents' equilibrium behavior (in Proposition 3.1), then I derive the optimal policy in the steady state (in Proposition 3.2).

Proposition 3.1. *Fix any policy reaction function in period t ($\boldsymbol{\rho}_t$) and any history of play (up to period t)*[12]*; the following constitutes an equilibrium among private-sector agents in generation t:*

1. *Each agent observes one private signal $s_{i,t} = \theta_t + g_t + \varepsilon_{i,t}$, where each $\varepsilon_{i,t}$ is i.i.d. according to $\varepsilon_{i,t} \sim N(0, var(\theta_t + g_t)/(2^{2\kappa} - 1))$,*[13]

2. *Each agent acts according to: $a^*_{i,t} = \gamma s_{i,t}$, where $\gamma = \frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})}$.*

Proof: See Appendix B.1. □

Proposition 3.1 characterises the private sector agents' equilibrium behavior (in each period t), given the central bank's policy rule up to period t. Note that this is similar to the equilibrium behavior of firms in chapter 2. Each private-sector agent in period t observes a noisy private signal centered on the sum of the policy instrument and the fundamental in that particular period, and his action is proportional to the realization of his private signal. Agents react relatively less to their private signals whenever the degree of strategic complementarity is higher, or whenever their capacity to process information is lower. For more details, see Appendix B.1, or chapter 2.

Before characterising the central bank's optimal reaction function, it is useful to make some remarks which will help simplify the problem. Firstly, in Corollary 3.1, we will note that the central bank's per period expected loss in period t is proportional to the variance of the sum of the fundamental and the policy instrument in period t (i.e. proportional to $var(\theta_t + g_t)$). Secondly, in Corollary 3.2, we will note that the central bank's information precision about the fundamental does not depend on its reaction function $\{\boldsymbol{\rho}_t\}_{t=1}^{\infty}$. Together, these will imply that the central bank optimally chooses a reaction function which sets g_t in each period to minimize $var(\theta_t + g_t)$ conditional on the central bank's information up to period t (i.e. $g^*_t = -E[\theta_t | I^{cb}_t]$). The central bank's steady-state information precision about the fundamental is characterised in Corollary 3.2 and the optimal reaction function in the steady-state is then characterised in Proposition 3.2.

[12] Referring to the previous policy rules $\{\boldsymbol{\rho}_t\}_{t=1}^{t-1}$, as well as the equilibrium behavior of agents in previous generations (i.e. their chosen allocations of attention and corresponding equilibrium actions).

[13] i.e. the equilibrium allocation of attention in period t specifies $K^*_t = 1, F^*_t = (1 \ \ 1), \Omega^*_t = (var(\theta_t + g_t)/(2^{2\kappa} - 1))$

Corollary 3.1. *The central bank's expected (per-period) loss in period t is proportional to $var(\theta_t + g_t)$ (i.e. to the variance of the sum of the fundamental and the policy instrument in period t):*

$$E[u_{cb,t}] = -\zeta var(\theta_t + g_t)$$

Where $\zeta = (\gamma - 1)^2 + \frac{\lambda\gamma^2}{2^{2\kappa}-1}$, and γ is characterised in Proposition 3.1.

Proof: see Appendix B.2. □

Corollary 3.2. *The central bank's information precision about the fundamental in any period T (i.e. τ_T^{cb}) is independent of its policy reaction function up to period T (i.e. $\{\boldsymbol{\rho}_t\}_{t=1}^{T}$). The central bank's steady-state information precision (denoted τ_{ss}^{cb}) solves:*

$$\tau_{ss}^{cb}\left[\frac{\alpha^2}{\tau_{ss}^{cb} + \gamma^2/\sigma_Z^2} + \sigma_\theta^2\right] = 1 \tag{3.22}$$

The central bank's steady-state information precision is:

1. *Increasing in the agents' capacity to process information: $\frac{\partial \tau_{ss}^{cb}}{\partial \kappa} > 0$;*
2. *Decreasing in the degree of strategic complementarity $(1-r)$: $\frac{\partial \tau_{ss}^{cb}}{\partial r} > 0$;*
3. *Decreasing in the persistence of the fundamental: $\frac{\partial \tau_{ss}^{cb}}{\partial \alpha} < 0$;*
4. *Decreasing in the variance of the innovation to the fundamental: $\frac{\partial \tau_{ss}^{cb}}{\partial \sigma_\theta^2} < 0$;*
5. *Decreasing in the variance of the central bank's observation noise: $\frac{\partial \tau_{ss}^{cb}}{\partial \sigma_Z^2} < 0$.*

Proof: Firstly, let us argue that the central bank's information precision about the fundamental is independent of its reaction function. If agents act as described in Proposition 3.1, then the average action in period $t-1$ is given by

$$\overline{a}_{t-1} = \gamma(\theta_{t-1} + g_{t-1}) \tag{3.23}$$

For any policy reaction function. Hence, the noisy signal observed by the central bank in period t writes as:

$$Z_t = \overline{a}_{t-1} + \varepsilon_t^Z = \gamma(\theta_{t-1} + g_{t-1}) + \varepsilon_t^Z \tag{3.24}$$

Because the central bank knows g_{t-1} in period t, it can construct the unbiased signal of the fundamental:

$$s_t := \frac{Z_t}{\gamma} - g_{t-1} = \theta_{t-1} + \frac{\varepsilon_t^Z}{\gamma} \tag{3.25}$$

Which reveals the fundamental with noise:

$$\sigma_s^2 := \frac{\sigma_Z^2}{\gamma^2} \tag{3.26}$$

Importantly, note that σ_s^2 does not depend on the central bank's reaction function. Although $\boldsymbol{\rho}_{t-1}$ influences the average action in period $t-1$, it does not influence the informativeness of the noisy signal Z_t for the central bank (because the central bank can perfectly disentangle its policy from the average action).[14] The argument applies to all periods, so changes in the central bank's reaction function $\{\boldsymbol{\rho}_t\}_{t=1}^T$ do not affect the central bank's information precision about the fundamental in period T.

In the spirit of Morris and Shin (2005), we can then derive a recursive expression for the central bank's information precision over time. Consider a generic period $t-1$. Recall that the central bank's information precision (about the fundamental in period $t-1$) after observing the signal Z_{t-1} is denoted by τ_{t-1}^{cb}:

$$\tau_{t-1}^{cb} = 1/var(\theta_{t-1}|I_{t-1}^{cb}) = 1/var(\theta_{t-1}|Z_{t-1}, Z_{t-2}, ...) \tag{3.27}$$

Following play in period $t-1$, the central bank observes the signal Z_t in period t, from which it constructs the unbiased signal s_t of the fundamental θ_{t-1} (with noise $\sigma_s^2 = \frac{\sigma_Z^2}{\gamma^2}$). Hence, the central bank's information precision about θ_{t-1} in period t is:

$$\begin{aligned} 1/var(\theta_{t-1}|I_t^{cb}) &= 1/var(\theta_{t-1}|Z_t, Z_{t-1}, Z_{t-2}, ...) \\ &= \tau_{t-1}^{cb} + \frac{\gamma^2}{\sigma_Z^2} \end{aligned} \tag{3.28}$$

The central bank does not have any additional information concerning θ_t in period t, so its estimate of θ_t is:

$$E[\theta_t|I_t^{cb}] = E[\alpha\theta_{t-1} + \varepsilon_t^\theta] = \alpha E[\theta_{t-1}|I_t^{cb}] \tag{3.29}$$

Which implies that:

$$var(\theta_t|I_t^{cb}) = \alpha^2 var(\theta_{t-1}|I_t^{cb}) + \sigma_\theta^2 \tag{3.30}$$

Noting that $var(\theta_t|I_t^{cb}) = 1/\tau_t^{cb}$, we can substitute (3.28) into (3.30) to obtain a

[14] A similar observation is made in Chapter 2.

recursive expression for the central bank's information precision:

$$\frac{1}{\tau_t^{cb}} = \frac{\alpha^2}{\tau_{t-1}^{cb} + \gamma^2/\sigma_Z^2} + \sigma_\theta^2 \tag{3.31}$$

The steady-state information precision thus solves:

$$\tau_{ss}^{cb}\left[\frac{\alpha^2}{\tau_{ss}^{cb} + \gamma^2/\sigma_Z^2} + \sigma_\theta^2\right] = 1 \tag{3.32}$$

Which is a quadratic equation in τ_{ss}^{cb}. Solving this for τ_{ss}^{cb} (and noting that one root is strictly negative and thus cannot be a solution as $\tau_{ss}^{cb} \geq 0$) gives:

$$\tau_{ss}^{cb} = \frac{\sqrt{\left(\sigma_z^2 - \alpha^2\sigma_z^2 - \gamma^2\sigma_\theta^2\right)^2 + 4\gamma^2\sigma_\theta^2\sigma_z^2} + \sigma_z^2 - \alpha^2\sigma_z^2 - \gamma^2\sigma_\theta^2}{2\sigma_\theta^2\sigma_z^2} \tag{3.33}$$

The comparative statics in Corollary 3.2 follow from (3.33). Note that increasing the agents' capacity to process information (κ), or decreasing the degree of strategic complementarity (i.e. increasing r) prompts agents to respond more strongly to their private signals $s_{i,t}$ (i.e. leads to a higher γ in equilibrium — see Proposition 3.1). This makes the central bank's signal in period t more informative about the fundamental in period $t-1$, and leads to a higher steady-state information precision for the central bank — the intuition is similar to Morris and Shin (2005). □

Proposition 3.2. *In the steady-state, optimal policy specifies:*

$$\boldsymbol{\rho}_t^* = (-\frac{\alpha\phi}{\gamma}, -\frac{\alpha^2\phi}{\gamma}, -\frac{\alpha^3\phi}{\gamma}, ...)$$

i.e.

$$\rho_{t,t-k}^* = -\frac{\alpha^{k+1}\phi}{\gamma} \tag{3.34}$$

where

$$\phi = \frac{\gamma^2/\sigma_z^2}{\gamma^2/\sigma_z^2 + \tau_{ss}^{cb}} \tag{3.35}$$

And where γ is characterised in Proposition 3.1 and τ_{ss}^{cb} is characterised in Corollary 3.2.

Proof: see Appendix B.3. □

3.3 Model featuring Brainard uncertainty

So far we have assumed that the policymaker faces uncertainty only regarding fundamentals. In this section, I relax this assumption and analyse the effects of parameter uncertainty on the optimal policy. By parameter uncertainty, I refer to uncertainty regarding the relationship between the central bank's policy instrument and the agents' optimal action — more specifically, I now assume that each agent's payoff is given by:

$$u_{i,t} = -\left[a_{i,t} - r(\theta_t + b_t g_t) - (1-r)\overline{a}_t\right]^2 \tag{3.36}$$

Where b_t is not perfectly known by the policymaker (whereas in the certainty-equivalent model from before, $b_t = 1 \forall t$). I account for parameter uncertainty by assuming that there is uncertainty regarding the effects of policy intervention (on the agents' optimal action); although the policymaker will not know the exact value of this parameter (b_t), it will know the distribution from which it is drawn.

This is referred to as multiplicative uncertainty (as opposed to additive uncertainty, for instance, about fundamentals θ_t) — the more the policy instrument is used, the more uncertainty is multiplied into the target variable (in this case, the agents' optimal action). In turn, in a static setup such as in Brainard (1967) this calls for less policy intervention. I first show that this also applies to a static version of the beauty contest game from the previous section. Then, I discuss how parameter uncertainty would affect the central bank's learning about the fundamental in a dynamic setting, and the implications of this for optimal policy.

3.3.1 Static setup

Firstly, let us consider a one-shot beauty contest game featuring an activist policymaker and Brainard uncertainty. Similarly to before, agents' payoffs are given by:

$$u_i = -\left[a_i - r(\theta + bg) - (1-r)\overline{a}\right]^2 \tag{3.37}$$

And the policymaker's payoff is:

$$u_{cb} = -(\overline{a} - \theta - bg)^2 - \lambda \int_0^1 (a_i - \overline{a})^2 di \tag{3.38}$$

Which are counterparts of equations (3.36) and (3.3) respectively. In this section only, suppose that $\theta \sim N(0, \sigma_\theta^2)$ and $b \sim N(1, \sigma_b^2)$ (and the prior distributions of θ and b are common knowledge).

As previously mentioned, the policymaker does not know the value of b, but he does know the distribution from which it is drawn; other than this (i.e. the distribution of b), the central bank does not have any other information concerning b. In this section only, I assume that the central bank learns about the fundamental (θ) by observing an exogenous signal $x = \theta + \varepsilon_x$, where $\varepsilon_x \sim N(0, \sigma_x^2)$ is independent of all random variables. As before, the central bank chooses a reaction function mapping its signal realization to the policy instrument $g = \rho x$ (where the central bank chooses ρ to maximize the unconditional expectation of (3.38)).

Unlike the central bank, suppose that private-sector agents know how policy intervention affects their optimal action (more precisely, I assume that the realization of b is common knowledge among private-sector agents when choosing their allocation of attention).[15] Other than this, the private-sector agents' problem is the same as before. To avoid any potential confusion, I restate the timing of the problem in the presence of parameter uncertainty:

1. The central bank chooses its reaction function (ρ)
2. The parameter b is realized
3. Private-sector agents are born (note that b and ρ are common knowledge among private-sector agents)
4. Each private agent chooses his allocation of attention (K_i, F_i, Ω_i)
5. The fundamental (θ) is realized
6. The central bank observes its private signal x and sets g according to the policy rule $g = \rho x$
7. Each private agent observes his signal vector (S_i) and takes his action a_i

The following Proposition characterises the equilibrium of the game.

Proposition 3.3. *Fix any policy reaction function ρ. The following constitutes an equilibrium among private-sector agents:*

1. *Each agent observes one private signal $s_i = \theta + bg + \varepsilon_i$, where each ε_i is i.i.d. according to $\varepsilon_i \sim N(0, var(\theta + bg|b)/(2^{2\kappa} - 1))$,*[16]

[15] I make this assumption in order to preserve the linear-quadratic-Gaussian structure of the agents' problem. I expand on this in the next subsection.

[16] i.e. the equilibrium allocation of attention specifies $K^* = 1, F^* = (1\ \ b), \Omega^* = (var(\theta + bg|b)/(2^{2\kappa} - 1))$

2. Each agent acts according to: $a_i^* = \gamma s_i$*, where* $\gamma = \frac{r(1-2^{-2\kappa})}{1-(1-r)(1-2^{-2\kappa})}$*.*

Optimal policy specifies:

$$\rho^* = -\frac{1}{1+\sigma_b^2}\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_x^2}, \tag{3.39}$$

i.e. the policy instrument is set according to:

$$g^* = -\frac{1}{1+\sigma_b^2}E[\theta|x]. \tag{3.40}$$

Proof: See Appendix B.4. □

Under Brainard uncertainty, the central bank no longer acts solely on the basis of the expected values of random variables, but also makes use of information concerning the variance of b (in other words, optimal policy no longer displays "certainty-equivalence"). Also note that g^* is always lower (in absolute value) than the central bank's expectation of the fundamental. Further, the optimal degree of policy activism is decreasing in σ_b^2. These are all standard features of Brainard conservatism — in fact, optimal policy (3.39) is equivalent to the one in Brainard's "one target-one instrument" case.[17] The more the central bank intervenes to stabilize the shock to the fundamental, the higher is the policy-induced variance in the agents' optimal action and, in turn, this calls for less policy activism.

3.3.2 The signal value of prices under parameter uncertainty

Before reintroducing dynamics, let us analyse how policy intervention affects the information revealed by the noisy signal of the average action in the static setup from the previous subsection. More precisely, suppose that following play of the static beauty contest game (from section 3.3.1), the central bank observes a noisy signal of the agents' average action $Z = \overline{a} + \varepsilon_Z$ (as in section 3.2). Given Proposition 3.3, for any policy reaction function ρ, the average action is given by:

$$\overline{a} = \gamma(\theta + bg) \tag{3.41}$$

So the noisy signal observed by the central bank writes as:

$$Z = \gamma(\theta + bg) + \varepsilon_Z \tag{3.42}$$

[17] See equation (5') in Brainard (1967) and the discussion regarding it (and compare this with equation (3.40) above).

Note that this conveys information to the central bank regarding both the fundamental (θ) and the transmission of policy (b). More specifically, the incremental information value about the fundamental comes from the signal:

$$\hat{\theta} := \frac{Z}{\gamma} - g \tag{3.43}$$

Which reveals the fundamental with noise:

$$\sigma^2_{\hat{\theta}} := E[(\hat{\theta} - \theta)^2] = g^2\sigma^2_b + \frac{\sigma^2_z}{\gamma^2} \tag{3.44}$$

While the incremental information value about the transmission of policy (b) comes from the signal:

$$\hat{b} := \frac{Z}{\gamma g} - \frac{E[\theta|x]}{g} \tag{3.45}$$

Which reveals the transmission of policy with noise:

$$\sigma^2_{\hat{b}} := E[(\hat{b} - b)^2] = \frac{var[\theta|x]}{g^2} + \frac{\sigma^2_z}{\gamma^2 g^2} \tag{3.46}$$

Note that whenever the central bank intervenes more (referring to a higher absolute value of g), it learns more about the transmission of policy, but less about the fundamental (i.e. this leads to a lower $\sigma^2_{\hat{b}}$, but also to a higher $\sigma^2_{\hat{\theta}}$). Absent any policy intervention ($g = 0$), the signal Z contains no information about b ($\sigma^2_{\hat{b}} \to \infty$), but reveals the same amount of information about θ as in the absence of Brainard uncertainty ($\sigma^2_{\hat{\theta}} = \frac{\sigma^2_z}{\gamma^2}$).

The intuition is the following: because agents process information about their optimal action (which is a weighted sum of the fundamental and the policy instrument), the average action reveals this weighted sum (to the extent that agents respond to their private signals) — see (3.41). While in the certainty-equivalent setup from section 3.2, the average action was proportional to the sum of policy and the fundamental,[18] and the policymaker could perfectly disentangle the policy instrument from the average action, this is no longer the case under parameter uncertainty, because the average action necessarily reflects some of the uncertainty associated with the transmission of policy, alongside information concerning the fundamental. Not perfectly knowing either θ or b, the policymaker learns about both upon observing the noisy signal of the average action.

[18] As opposed to the *weighted* sum $\theta + bg$.

The fact that more policy intervention increases the central bank's information precision about policy transmission (but lowers the central bank's information precision about fundamentals) should hopefully be intuitive in light of the above (as more policy intervention leads to more policy-induced variance in the average action). In turn, in a dynamic setup in which there is persistence in both θ and b, such that better information about θ_t (b_t) improves the central bank's estimate of θ_{t+1} (b_{t+1}), the central bank should also take learning into account when setting policy.[19] More specifically, the central bank would now have an incentive to experiment with policy,[20] in order to learn more about its transmission, as well as an incentive to be cautious with policy,[21] in order to have more precise information concerning fundamentals.

3.3.3 Relation to literature on central bank learning under parameter uncertainty

Central bank learning under parameter uncertainty has already been extensively studied — see for instance, Bertocchi and Spagat (1993); Balvers and Cosimano (1994); Sack (1998); Ellison and Valla (2001) among many others. While in the original (static) formulation of Brainard (1967), parameter uncertainty calls for less policy activism from the central bank, this has been disputed as being poor from an informational perspective, as the policymaker learns less about the transmission of policy if this is being used cautiously. In turn, this calls for more activism and "policy experiments" designed to elicit information regarding policy transmission:

"Whether or not the Fed actually does sacrifice short-term goals to carry out experiments, we will show that it should, since the government should in general not behave myopically. While the degree to which it should sacrifice current reward in exchange for information depends on such factors as discount rates and the nature of the uncertainty faced, it is certainly true that large informational gains justify small sacrifices. [...] At each stage of the process, there is a potential trade-off between minimization of output variability and the value of the information which can be obtained through an activist policy. The implications of learning for policy prescriptions are intuitively clear; monetary policy should be activist in the sense that it should be responsive to new information as it arises. In addition, there is a

[19] Because its information precision about both the fundamental and the transmission of policy would depend on its policy reaction function (unlike in the certainty-equivalent setup, see Corollary 3.2).

[20] i.e. to set a higher absolute value of g (in expectation).

[21] i.e. to set a lower value of g (in expectation).

second sense in which monetary policy should be activist. It should actively seek to generate information even if it is costly to do so." (Bertocchi and Spagat (1993))

In a dynamic setup of our model (featuring parameter uncertainty), the central bank should be "activist" in both senses described by Bertocchi and Spagat (1993) — in the first sense, it should change policy over time upon observing its signals and updating its beliefs about the fundamental[22] (and the transmission of policy); in the second sense, it should consider undertaking "policy experiments" (i.e. a more activist policy) seeking to generate information about the transmission of policy. However, it should take into account that such "policy experiments" are informationally costly, as they come at the expense of information concerning the fundamental — more specifically, the central bank loses information concerning the counterfactual of what would have happened in the absence of policy intervention (because the average action now reflects some of the uncertain effects associated with the policy intervention and this distorts the informativeness of the average action about the fundamental).[23]

In turn, this implies that the central bank's problem in a dynamic setting resembles the more general problem of "learning with two unobservable parameters" discussed in Balvers and Cosimano (1994) — I expand on this in the next subsection.

3.3.4 Dynamics

In this section, I will again assume that θ_t is governed by the AR(1) process described in section 3.2. Similarly, I assume that initially b_{-1} is drawn from a common knowledge prior distribution $b_{-1} \sim N(\bar{b}, \sigma^2_{b,-1})$; then, it also evolves according to an AR(1):

$$b_t = \beta b_{t-1} + \varepsilon^b_t \tag{3.47}$$

Where $\beta \in [0, 1]$, and the ε^b_t's are independently and identically distributed according to $\varepsilon^b_t \sim N(0, \sigma^2_b)$.

[22] Note that the central bank in the certainty-equivalent setup from section 3.2 is activist in this sense.

[23] The central bank would only act as described in Bertocchi and Spagat (1993) if it faced no uncertainty concerning the fundamental. For instance, in the example from the previous subsection, consider what happens if $\sigma^2_x = 0$, such that the central bank's private signal about the fundamental (x) is perfectly informative. The central bank would then only use the signal Z to infer information concerning b. In this case, if there is persistence in b (such that learning is valuable for the central bank in future periods), the policymaker should consider a policy which responds more strongly to its private signal x (than the myopic policy (3.39)). The example is in the spirit of Bertocchi and Spagat (1993) — see cases 3-5 in their paper.

Firstly, we will note that a complication arises in the dynamic formulation of the problem. More specifically, the prior distribution of the agents' optimal action (in a generic period t) is no longer Gaussian, unless relatively strong assumptions are imposed.

Recall that in the static setup from section 3.3.1, we assumed that b is common knowledge among private-sector agents (when choosing their allocation of attention), such that the prior distribution of each agent's optimal action (conditional on b) is Gaussian. Given that payoffs are quadratic, this allows us to use the same reasoning as in Proposition 3.1 to characterise the equilibrium allocation of attention. In contrast, if private-sector agents did not know the realization of b, the prior distribution of the optimal action (under perfect information) $\theta + bg$ would no longer be Gaussian because the product of two normal random variables (in this case b and g) is not normally distributed. In turn, this complicates the problem of determining the equilibrium allocation of attention (conditional on the central bank's reaction function) because the linear-quadratic-Gaussian structure of the problem is not preserved.

In a dynamic setting, assuming that agents in each generation t know b_t is no longer sufficient to ensure that the optimal action (under perfect information) in period t (i.e. $\theta_t + b_t g_t$) is Gaussian, even if the central bank uses a linear policy rule because, from the perspective of a private-sector agent born in generation t, g_t is not normally distributed — to see this more clearly, suppose that all agents in previous generations acted in a similar way to the previous subsection, i.e. according to:

$$a_{i,t-k} = \gamma(\theta_{t-k} + b_{t-k} g_{t-k} + \varepsilon_{i,t-k}) \tag{3.48}$$

For all i and k. It follows that in each period $t-k+1$, the central bank observes the signal:

$$Z_{t-k+1} = \gamma(\theta_{t-k} + b_{t-k} g_{t-k}) + \varepsilon^{Z}_{t-k+1} \tag{3.49}$$

Hence, given a linear policy reaction function (such as (3.8)), the policy instrument in period t writes as:

$$\begin{aligned} g_t &= \sum_{k=0}^{t-1} \rho_{t,t-k} Z_{t-k} \\ &= \sum_{k=0}^{t-1} \rho_{t,t-k} [\gamma(\theta_{t-k-1} + b_{t-k-1} g_{t-k-1}) + \varepsilon^Z_{t-k}] \end{aligned} \tag{3.50}$$

So from the perspective of agent i in generation t (who does not know the value of b_{t-k}), the policy instrument in period t is not normally distributed because the terms $b_{t-k} g_{t-k}$ are not normally distributed (as discussed above). We could assume that the whole sequence $\{b_t\}_{t=0}^t$ is common knowledge among private-sector agents in period t (and this would preserve the linear-quadratic-Gaussian structure), but this is not particularly appealing.[24] An additional complication is that in the dynamic setting featuring parameter uncertainty, the central bank may not necessarily want to set policy as a linear function of its signals. Furthermore, even if one is ready to make these strong assumptions,[25] the central bank's problem would resemble to a large extent the "learning with two unobservable parameters" problem in Balvers and Cosimano (1994). Hence, I instead make some simplifying assumptions seeking to draw a more direct parallel with Balvers and Cosimano (1994) — the first one fixes the information structure (in the spirit of the previous sections); the second one restricts private-sector agents to use a linear action rule; the third states that the central bank observes average actions perfectly:

1. Each agent in generation t knows the distribution of the optimal action under perfect information (i.e. the prior distribution of $\theta_t + b_t g_t$) and observes a private signal of it, i.e. each agent in generation t observes the signal $s_{i,t} = \theta_t + b_t g_t + \varepsilon_{i,t}$, where each $\varepsilon_{i,t}$ is i.i.d. according to $\varepsilon_{i,t} \sim N(0, var(\theta_t + b_t g_t)/(2^{2\kappa} - 1))$.

2. Each agent takes his action as a linear function of his signal.

[24] Indeed, even in the static setting, one may wonder whether it is realistic to assume that private-sector agents have perfect knowledge concerning the effects of policy intervention on their optimal action. In the stylised, beauty-contest formulation of the model this may seem reasonable, but in the context of the microfounded economy, this is less so. In the latter context, uncertainty concerning the effects of policy intervention would translate to uncertainty regarding the relationship between the central bank's policy instrument (for instance, the money supply) and nominal demand — it does not seem realistic to assume that private-sector agents are perfectly informed about this relationship.

[25] i.e. restricting the central bank to use a linear policy rule and assuming that private-sector agents in generation t know the whole sequence $\{b_t\}_{t=0}^t$.

3. In each period t, the central bank perfectly observes the agents' average action in period $t-1$ (i.e. I assume that $\sigma_Z^2 = 0$).

Remark 3.2. *Assumptions 1 and 2 above could be motivated*[26] *by assuming that private-sector agents in each generation t use a "simplified view of the world" and misperceive the shape of the distribution of the optimal action in that period — in particular, they believe that the optimal action $\theta_t + b_t g_t$ is Gaussian (which is not the case), but other than this, they correctly perceive the mean and variance of the distribution.*[27] *Note that, in the limit, as $\sigma_b^2 \to 0$, the model collapses to the "certainty-equivalent" model from section 3.2, and the prior distribution of the optimal action in generation t is indeed Gaussian, so conditions 1 and 2 above emerge endogenously (as previously discussed).*

Proposition 3.4. *Suppose conditions 1, 2 and 3 above are satisfied. Then, the central bank's problem is equivalent to the one in Balvers and Cosimano (1994).*[28]

Proof: Given conditions 1 and 2, one can use similar arguments as in the proof of Proposition 3.1 to argue that each agent i finds it optimal to act according to $a_{i,t}^* = \gamma s_{i,t}$, when all other agents $j \neq i$ act according to $a_{j,t}^* = \gamma s_{j,t}$. Hence, $a_{i,t}^* = \gamma s_{i,t}$ specifies an equilibrium action rule for private-sector agents in each generation. Consequently, the average action in period t is given by $\overline{a} = \gamma(\theta_t + b_t g_t)$. Then, the same argument as in the proof of Corollary 3.1 implies that the expected per-period welfare loss in period t is proportional to the variance of $\theta_t + b_t g_t$ — more precisely, expected welfare in period t can be expressed as: $E[u_{cb,t}] = -\zeta var(\theta_t + b_t g_t)$. Hence, the central bank's problem can be expressed as:

$$\max_{\{g_t\}_{t=1}^{\infty}} -\zeta\delta^t \sum_{t=1}^{\infty} var(\theta_t + b_t g_t | I_t^{cb}) \tag{3.51}$$

Where $I_t^{cb} = \{g_s, \overline{a}_s\}_{s=0}^{t-1}$, denotes the central bank's information set in period t (when the central bank knows all previous average actions, as well as the policy instrument in all previous periods). Note that observing the average action from period $t-k$ is

[26] Albeit not in the most elegant way.

[27] More precisely, in each period t, the optimal action $\theta_t + b_t g_t$ will be a well-defined random variable with some mean and variance — if agents use the "simplified view of the world" described in the Remark, they (wrongly) believe that the optimal action is Gaussian, but they (correctly) perceive the first and second moments of its distribution. Note that in this case we would also need to write each agent's information processing constraint in terms of this misperception — more precisely, the information processing constraint would place an upper bound on the expected uncertainty reduction concerning the (misperceived, Gaussian) optimal action.

[28] More specifically, the fundamental (θ) corresponds to the intercept in their model, while the transmission of policy (b) corresponds to the slope parameter in their model.

informationally equivalent to observing the agents' optimal action in period $t-k$ (i.e. observing $\gamma(\theta_{t-k}+b_{t-k}g_{t-k})$ is informationally equivalent to observing $\theta_{t-k}+b_{t-k}g_{t-k}$ because the central bank knows γ). It follows that the central bank's problem is equivalent to the one described in Balvers and Cosimano (1994), section 2.1. $\square$

3.4 Conclusion

The paper analyses a sequence of static beauty contest games featuring an (infinitely-lived) activist policymaker who learns about fundamentals by observing noisy signals of the agents' average actions (as in Morris and Shin (2005)). Each period, a new generation of (short-lived) rationally inattentive agents are born who process a fixed amount of information about their environment, then take their actions. The paper firstly argues that the central bank can exploit the information revealed by the average actions of agents in past generations to improve the expected welfare of agents in the current generation — furthermore, it argues that in such a setting, the policymaker's actions do not affect his own information precision about fundamentals (because the central bank can always perfectly disentangle its policy instrument from the average action).

I then relax the assumption that the central bank perfectly knows the effects of its policy on the agents' optimal action (i.e. I introduce parameter uncertainty in the spirit of Brainard (1967)). I argue that under parameter uncertainty, the central bank can no longer perfectly disentangle its policy from the average action, because the average action necessarily reflects some of the uncertainty associated with the transmission of policy (alongside information concerning the fundamental). This implies that "policy experiments" designed to elicit information concerning policy transmission (as described in Bertocchi and Spagat (1993)) are informationally costly for the central bank, as they necessarily come at the expense of information concerning fundamentals. I then argue that under some simplifying assumptions, the central bank's problem resembles the problem of "learning with two unobservable parameters" in Balvers and Cosimano (1994).

A Proofs and Details for Chapter 1

A.1 Proof of Proposition 1.1

Proof: Given (1.4) the average action in equilibrium writes as:

$$\overline{a} = \int_0^1 a_i^* di$$
$$\overline{a} = \int_{i \in I} a_i^{I*} di + \int_{i \in U} a_i^{U*} di$$
$$\overline{a} = \int_{i \in I} [\phi_1^{I*}(\theta + \varepsilon_i) + \phi_2^{I*} y] di + \int_{i \in U} [\phi^{U*} y] di$$
$$\overline{a} = \lambda(\phi_1^{I*}\theta + \phi_2^{I*} y) + (1-\lambda)\phi^{U*} y$$

So the public signal can be expressed as:

$$y = \overline{a} + \varepsilon_y = [\lambda\phi_1^{I*}]\theta + [\lambda\phi_2^{I*} + (1-\lambda)\phi^{U*}]y + \varepsilon_y \tag{A.1}$$

Denote by s the unbiased signal of the fundamental (θ) contained in y — more specifically, for any $\phi_1^{I*} \neq 0$ and $\lambda\phi_2^{I*} + (1-\lambda)\phi^{U*} \neq 1$,[1] this is defined as:

$$s := y \frac{\left[1 - \lambda\phi_2^{I*} - (1-\lambda)\phi^{U*}\right]}{\lambda\phi_1^{I*}} \tag{A.2}$$

Note from (A.1) and (A.2) that:

$$s = \theta + \frac{\varepsilon_y}{\lambda\phi_1^{I*}} \tag{A.3}$$

Hence the precision[2] of s as an unbiased signal of the fundamental (denoted by τ_s) is equal to:

$$\tau_s = \tau_y(\lambda\phi_1^{I*})^2 \tag{A.4}$$

I also refer to this as the "informativeness" of the public signal about the fundamental. Intuitively, the information (about the fundamental) revealed by the noisy signal of the average action depends on the fraction of agents who are informed (λ) and the response of informed agents' actions to their private signals (ϕ_1^{I*}), alongside the precision of the observation noise τ_y, because it is the aggregation of the private

[1] For now let us just guess and verify that these hold. I argue afterwards that $\lambda\phi_2^{I*} + (1-\lambda)\phi^{U*} \neq 1$ and $\phi_1^{I*} \neq 0$ in any equilibrium (for any parameter values).

[2] i.e. the inverse of the variance of the noise.

signals x_i which reveals the fundamental.

Given the information structure, Bayesian updating implies that informed agents act according to:

$$a_i^{I*} = \frac{\tau_x}{\tau_x + \tau_\theta + (\lambda\phi_1^{I*})^2\tau_y} x_i + \frac{\tau_y(\lambda\phi_1^{I*})^2}{\tau_x + \tau_\theta + \tau_y(\lambda\phi_1^{I*})^2} s \tag{A.5}$$

And uninformed agents act according to:

$$a_i^{U*} = \frac{\tau_y(\lambda\phi_1^{I*})^2}{\tau_\theta + \tau_y(\lambda\phi_1^{I*})^2} s \tag{A.6}$$

Matching coefficients on x_i in (A.5) with (1.4) gives (1.10). Let $f(\phi_1^{I*})$ denote the right-hand side of (1.10) as a function of ϕ_1^{I*}, i.e.:

$$f(\phi_1^{I*}) = \tau_x / \left[\tau_x + \tau_\theta + \tau_y(\lambda\phi_1^{I*})^2\right]$$

Note that:

- $f(0) > 0$
- $f(1) < 1$
- $f(\phi_1^{I*})$ is continuous and strictly decreasing in ϕ_1^{I*}

It follows that there is a unique solution to (1.10).[3]

Intuitively, $f(\phi_1^{*I})$ is the optimal weight an informed agent puts on his private signal when taking his action, given that all other informed agents put weight ϕ_1^{*I} on their private signals (the equilibrium must be symmetric, hence the fixed point problem). $f(0) > 0$ means that when other informed agents disregard their private signals, an informed agent puts a positive weight on his own private signal (the intuition for $f(1) < 1$ is similar). $f(\phi_1^{*I})$ is decreasing, because as other informed agents respond more strongly to their private signals, the public signal becomes more informative (given any finite, positive t_y and any $\lambda \in (0, 1]$), so each informed agent finds it optimal to respond less strongly to his private signal x_i.

Substituting (A.2) into (A.5) and (A.6) and matching coefficients on y with (1.4)

[3] More precisely, (1.10) is a cubic equation (in ϕ_1^{I*}) with one real root and two complex roots. I restrict attention to real roots.

for informed and uninformed agents respectively gives:

$$\begin{cases} \phi_2^{I*} = \frac{\tau_y(\lambda\phi_1^{I*})^2}{\tau_x+\tau_\theta+\tau_y(\lambda\phi_1^{I*})^2} \frac{\left[1-\lambda\phi_2^{I*}-(1-\lambda)\phi^{U*}\right]}{\lambda\phi_1^{I*}} \\ \phi^{U*} = \frac{\tau_y(\lambda\phi_1^{I*})^2}{\tau_\theta+\tau_y(\lambda\phi_1^{I*})^2} \frac{\left[1-\lambda\phi_2^{I*}-(1-\lambda)\phi^{U*}\right]}{\lambda\phi_1^{I*}} \end{cases} \tag{A.7}$$

For ease of notation, let $m = \frac{\tau_y(\lambda\phi_1^{I*})^2}{\tau_\theta+\tau_y(\lambda\phi_1^{I*})^2}$ and $n = \frac{\tau_y(\lambda\phi_1^{I*})^2}{\tau_x+\tau_\theta+\tau_y(\lambda\phi_1^{I*})^2}$. Solving the system of equations (A.7) yields:

$$\phi_2^{I*} = n/[m(1-\lambda) + \lambda(n + \phi_1^{I*})] \tag{A.8}$$

$$\phi^{U*} = m/[m(1-\lambda) + \lambda(n + \phi_1^{I*})] \tag{A.9}$$

Note that there is a unique equilibrium in the second step of the game (i.e. $(\phi_1^{I*}, \phi_2^{I*}, \phi^{U*})$) for any $\lambda \in (0, 1]$ and any parameter values (positive, finite $\tau_x, \tau_y, \tau_\theta$).[4] $\square$

An analytical expression for ϕ_1^{I*} (i.e. the solution to equation (1.10)) is given by:

$$\begin{aligned} \phi_1^{I*} = & \frac{\left[27\lambda^4\tau_x\tau_y^2 + \sqrt{108\lambda^6(\tau_\theta+\tau_x)^3\tau_y^3 + 729\lambda^8\tau_x^2\tau_y^4}\right]^{\frac{1}{3}}}{3\sqrt[3]{2}\lambda^2\tau_y} - \\ & - \frac{\sqrt[3]{2}(\tau_\theta+\tau_x)}{\left[27\lambda^4\tau_x\tau_y^2 + \sqrt{108\lambda^6(\tau_\theta+\tau_x)^3\tau_y^3 + 729\lambda^8\tau_x^2\tau_y^4}\right]^{\frac{1}{3}}} \end{aligned} \tag{A.10}$$

(A.10) can be used to argue that $\phi_1^{I*}(\lambda, \tau_y)$ is continuous in λ and τ_y for any $\lambda \in (0, 1]$ and any positive, finite $\tau_x, \tau_y, \tau_\theta$. Then, it is straightforward to argue that $\tau_s(\lambda, \tau_y)$ is continuous in λ and τ_y, and hence that $\gamma(\lambda, \tau_y)$ is continuous in λ and τ_y which is used in the proof of Proposition 1.2.

A.2 Proof of Corollary 1.1

Proof: From (1.10) it follows that holding everything else constant, a higher τ_y leads to a lower ϕ_1^{I*} in equilibrium (i.e. increasing the precision with which agents observe the average action prompts informed agents to respond less strongly to their private signals (x_i) in the equilibrium of the second step of the game) — by implicitly differentiating (1.10) w.r.t. τ_y it is straightforward to show that:

$$\frac{\partial\phi_1^{I*}(\lambda, \tau_y)}{\partial\tau_y} < 0 \tag{A.11}$$

[4] Also, for any parameter values $\lambda\phi_2^{I*} + (1-\lambda)\phi^{U*} < 1$ and $\phi_1^{I*} > 0$, so $\lambda\phi_2^{I*} + (1-\lambda)\phi^{U*} \neq 1$ and $\phi_1^{I*} \neq 0$ hold in any equilibrium. Thus, our guess in footnote 1 in (Appendix A.1) is verified.

In turn, a lower ϕ_1^{I*} is only consistent with a higher $\tau_s(\lambda, \tau_y)$, i.e. informed agents respond less strongly to their private signals x_i only if the public signal is more informative about the fundamental — note that we can express:[5]

$$\phi_1^{I*}(\lambda, \tau_y) = \frac{\tau_x}{\tau_x + \tau_\theta + \tau_s(\lambda, \tau_y)} \tag{A.12}$$

So indeed:

$$\frac{\partial \phi_1^{I*}(\lambda, \tau_y)}{\partial \tau_s(\lambda, \tau_y)} < 0 \tag{A.13}$$

Using the chain rule to differentiate (A.12) w.r.t. τ_y we can write:

$$\frac{\partial \phi_1^{I*}(\lambda, \tau_y)}{\partial \tau_y} = \frac{\partial \phi_1^{I*}(\lambda, \tau_y)}{\partial \tau_s(\lambda, \tau_y)} \frac{\partial \tau_s(\lambda, \tau_y)}{\partial \tau_y} \tag{A.14}$$

It follows from (A.11), (A.13) and (A.14) that $\frac{\partial \tau_s(\lambda, \tau_y)}{\partial \tau_y} > 0$.

A similar argument can be used to show that $\frac{\partial \tau_s(\lambda, \tau_y)}{\partial \lambda} > 0$. □

A.3 Proof of Corollary 1.2

Proof: If the private signal is free, all agents become informed ($\lambda^* = 1$) so the welfare loss is:

$$L^*_{W|c=0} = \frac{1}{\tau_x + \tau_\theta + \tau_s(1, \tau_y)}$$

It follows from Corollary 1.1 that $\frac{\partial \tau_s(1, \tau_y)}{\partial \tau_y} > 0$, so $\frac{\partial L^*_W}{\partial \tau_y}_{|c=0} < 0$. □

A.4 Proof of Corollary 1.3

Proof: Increasing the informativeness of the public signal lowers the expected losses of both informed and uninformed agents, but uninformed agents benefit relatively more (because they are more poorly informed) — differentiating (1.15) w.r.t. $\tau_s(\lambda, \tau_y)$, it is straightforward to show that:

$$\frac{\partial \gamma(\lambda, \tau_y)}{\partial \tau_s(\lambda, \tau_y)} > 0 \tag{A.15}$$

Which states that as the public signal becomes more informative, the ratio of the expected loss of informed agents to the expected loss of uninformed agents increases.

[5] Equation (A.12) follows from (1.12), (1.4) and (A.5).

Using the chain rule to differentiate (1.15) we can express:

$$\frac{\partial\gamma(\lambda,\tau_y)}{\partial\tau_y} = \frac{\partial\gamma(\lambda,\tau_y)}{\partial\tau_s(\lambda,\tau_y)}\frac{\partial\tau_s(\lambda,\tau_y)}{\partial\tau_y} \tag{A.16}$$

As $\frac{\partial\tau_s(\lambda,\tau_y)}{\partial\tau_y} > 0$ (see Corollary 1.1), it follows from (A.15) and (A.16) that $\frac{\partial\gamma(\lambda,\tau_y)}{\partial\tau_y} > 0$.

A similar argument can be used to show that $\frac{\partial\gamma(\lambda,\tau_y)}{\partial\lambda} > 0$. □

A.5 Proof of Proposition 1.2

Proof: Firstly, note that there are no equilibria in which all agents are uninformed (for any τ_y). It is straightforward to argue this by contradiction. Suppose $\lambda^* = 0$ (and note from the equilibrium conditions that this requires $L_u(0,\tau_y) \leq L_i(0,\tau_y)$). Then, the public signal is uninformative about the fundamental, so (in the absence of any other information source) each agent chooses $a_i = 0$ (which is the prior mean of θ). It follows that $L_u(0,\tau_y) = 1/\tau_\theta$ and $L_i(0,\tau_y) = 1/[\tau_\theta + \tau_x] + c$. Because we assumed that $c < \tau_x/[\tau_\theta(\tau_x + \tau_\theta)]$, it is optimal for an individual agent to become informed when there is no public information, so $L_u(0,\tau_y) > L_i(0,\tau_y)$ (a contradiction). Hence there are no equilibria in which $\lambda^* = 0$.

Secondly, consider equilibria in which all agents become informed. Suppose $\lambda^* = 1$ and note from the equilibrium conditions that this requires $L_u(1,\tau_y) \geq L_i(1,\tau_y)$, or equivalently $\gamma(1,\tau_y) \leq 1$. Note that:

1. $\gamma(1,0) < 1$ (if all agents are informed and there is no public information, each agent finds it optimal to become informed)[6]

2. $lim_{\tau_y\to\infty}\gamma(1,\tau_y) > 1$ (if all agents are informed and public information becomes perfectly precise, each agent finds it optimal to be uninformed)[7]

3. $\gamma(1,\tau_y)$ is continuous and strictly increasing in τ_y[8]

It follows that there is a unique τ_y which solves $\gamma(1,\tau_y) = 1$ — denote the solution by $\hat{\tau_y}$. It also follows (from point 3) that $\gamma(1,\tau_y) \leq 1$ for $\tau_y \leq \hat{\tau_y}$, so for such parameter values there is an equilibrium in which all agents become informed. It is also easy to see that this is the unique equilibrium.[9] This proves the first part of the Proposition.

[6] Because we assumed that $c < \frac{\tau_x}{\tau_\theta(\tau_x+\tau_\theta)}$.

[7] See Remark 1.2.

[8] See Corollary 1.3.

[9] If $\tau_y \leq \hat{\tau_y}$, then $\gamma(\lambda,\tau_y) < 1$ for $\lambda < 1$ (because $\gamma(\lambda,\tau_y)$ is strictly increasing in both arguments, see Corollary 1.3) so there are no equilibria in which some agents are informed and others are uninformed.

Lastly, consider equilibria in which some agents are informed while others are uninformed (i.e. $\lambda^* \in (0,1)$) and note that this requires $L_u(\lambda^*, \tau_y) = L_i(\lambda^*, \tau_y)$, or equivalently $\gamma(\lambda^*, \tau_y) = 1$. We know that there are no such equilibria for $\tau_y \leq \hat{\tau}_y$. Suppose $\tau_y > \hat{\tau}_y$ — note that:

1. $\gamma(0, \tau_y) < 1$ (if no agent is informed, each agent finds it optimal to become informed)[10]

2. $\gamma(1, \tau_y) > 1$ (if all agents are informed, each agent finds it optimal to be uninformed)[11]

3. $\gamma(\lambda, \tau_y)$ is continuous and strictly increasing in λ[12]

It follows that for $\tau_y > \hat{\tau}_y$ there is a unique λ which solves $\gamma(\lambda, \tau_y) = 1$ — denote the solution by λ^*. Note that this implies that for $\tau_y > \hat{\tau}_y$ there is a unique equilibrium in the first step of the game in which the fraction of agents who become informed is equal to $\lambda^* \in (0,1)$. This proves the second part of the Proposition. □

A.6 Proof of Corollary 1.4

The first part of the Corollary is analogous to Corollary 1.2. To prove the second part of the Corollary, we will note that the precision of the unbiased signal of the fundamental contained in y (i.e. $\tau_s(\lambda^*, \tau_y)$) is constant in all equilibria in which $\tau_y > \hat{\tau}_y$ (or $\lambda^* \in (0,1)$) — more specifically, in all such equilibria it must be the case that $L_u(\lambda^*, \tau_y) = L_i(\lambda^*, \tau_y)$. This is equivalent to:

$$\frac{1}{\tau_\theta + \tau_s(\lambda^*, \tau_y)} = \frac{1}{\tau_x + \tau_\theta + \tau_s(\lambda^*, \tau_y)} + c$$

Which is a quadratic equation in $\tau_s(\lambda^*, \tau_y)$. Solving this for $\tau_s(\lambda^*, \tau_y)$ (and noting that the solution needs to be greater than or equal to zero) yields:

$$\tau_s(\lambda^*, \tau_y) = \left[\sqrt{\frac{\tau_x(4 + \tau_x c)}{c}} - \tau_x \right] /2 - \tau_\theta \tag{A.17}$$

Hence in all equilibria in which $\tau_y > \hat{\tau}_y$ the precision of the unbiased signal of the fundamental contained in y is a constant which depends on the cost of private information (c), the precision of the private signal (τ_x) and the precision of the

[10] If no agent is informed (i.e. $\lambda = 0$), then the public signal is uninformative about the fundamental, so each agent finds it optimal to become informed because we assumed that $c < \frac{\tau_x}{\tau_\theta(\tau_x + \tau_\theta)}$.

[11] Because $\tau_y > \hat{\tau}_y$ and $\gamma(1, \tau_y)$ is strictly increasing in τ_y — see Corollary 1.3.

[12] See Corollary 1.3.

fundamental (τ_θ). It follows that for $\tau_y > \hat{\tau}_y$ the expected loss of uninformed agents writes as:

$$L_u(\lambda^*, \tau_y) = \frac{1}{\tau_\theta + \tau_s(\lambda^*, \tau_y)} = \frac{2}{\sqrt{\frac{\tau_x(4+\tau_x c)}{c}} - \tau_x} \tag{A.18}$$

As $L_u(\lambda^*, \tau_y) = L_i(\lambda^*, \tau_y)$ in all equilibria in which $\tau_y > \hat{\tau}_y$, we can express the welfare loss as:

$$L_W^* = \lambda^* L_i(\lambda^*, \tau_y) + (1 - \lambda^*) L_u(\lambda^*, \tau_y) = \frac{2}{\sqrt{\frac{\tau_x(4+\tau_x c)}{c}} - \tau_x} \tag{A.19}$$

Note that this does not depend on τ_y.[13] This proves the second part of Corollary 1.4.

In the main text, I also mention that the responsiveness of informed agents' actions to their private signals is constant in all equilibria where public information is sufficiently precise ($\tau_y > \hat{\tau}_y$) — using (A.12) and (A.17) this can be expressed as:

$$\phi_1^{I*}(\lambda^*, \tau_y) = \frac{\tau_x}{\tau_x + \tau_\theta + \tau_s(\lambda^*, \tau_y)} = \frac{2\tau_x}{\sqrt{\frac{\tau_x(4+\tau_x c)}{c}} + \tau_x} \tag{A.20}$$

It is also claimed that any increase in the precision with which agents observe the average action above $\hat{\tau}_y$ is accompanied by a fall in the fraction of agents who acquire the private signal (i.e. a lower λ^*), such that the informativeness of the public signal remains unchanged. Recall from (A.3) and the discussion following it that the informativeness of the public signal about the fundamental ($\tau_s(\lambda, \tau_y)$) depends on:

1. the fraction of agents who are informed (λ);
2. the responsiveness of informed agents' actions to their private signals (ϕ_1^{I*});
3. the precision with which agents observe the average action (τ_y).

Because $\tau_s(\lambda^*, \tau_y)$ and $\phi_1^{I*}(\lambda^*, \tau_y)$ are constant in all equilibria in which $\lambda^* \in (0, 1)$, it follows that if there is a change in τ_y, the fraction of agents who become informed in equilibrium (λ^*) adjusts such that the informativeness of the public signal ($\tau_s(\lambda^*, \tau_y)$) stays constant. More specifically, using (1.12), the precision of the unbiased signal

[13] Also note that this does not depend on the precision of the prior (τ_θ). If the fundamental is drawn from a more imprecise prior, the fraction of agents who become informed in equilibrium increases. This improves the informativeness of the public signal to the extent that agents remain indifferent between buying and not buying the private signal, and welfare remains unchanged.

of the fundamental contained in y (in equilibrium) writes as:

$$\tau_s(\lambda^*, \tau_y) = \tau_y[\lambda^* \phi_1^{I*}(\lambda^*, \tau_y)]^2 \tag{A.21}$$

This is equivalent to:

$$\lambda^* = \sqrt{\frac{\tau_s(\lambda^*, \tau_y)}{\tau_y \phi_1^{I*}(\lambda^*, \tau_y)^2}} \tag{A.22}$$

It is clear from (A.22), (A.20) and (A.17) that an increase in τ_y prompts a fall in λ^*. Note that (A.22) is equivalent to $\gamma(\lambda^*, \tau_y) = 1$.[14]

□

A.7 Relation to models of endogenous central bank information

In this section only, suppose that the noisy signal of the average action is no longer observable to agents — instead, there is a policymaker who observes the noisy signal $y = \overline{a} + \varepsilon_y$. The policymaker observes the signal y, constructs an unbiased signal of the fundamental (s) and simultaneously discloses a public signal $z = s + \varepsilon_z$ (where $\varepsilon_z \sim N(0, 1/\tau_z)$ is distributed independently of all random variables and τ_z measures the policymaker's degree of transparency).[15]

In such a setting, an equilibrium specifies:

$$(\lambda^*, \phi_1^{I*}, \phi_2^{I*}, \phi^{U*})$$

Such that:

$$\begin{cases} a_i^{I*} = E[\theta|x_i, z] = \phi_1^{I*} x_i + \phi_2^{I*} z \\ a_i^{U*} = E[\theta|y] = \phi^{U*} z \end{cases}$$

$$\begin{cases} \lambda^* = 1 & \text{if } \ L_i/L_u < 1 \\ \lambda^* = 0 & \text{if } \ L_i/L_u > 1 \\ \lambda^* \in [0,1] & \text{if } \ L_i/L_u = 1 \end{cases}$$

Where

$$L_i = E[(\phi_1^{I*} x_i + \phi_2^{I*} z - \theta)^2] + c$$

[14] And because we supposed that $\tau_y > \hat{\tau_y}$, λ^* is guaranteed to lie between 0 and 1 — see the proof of Proposition 1.2.

[15] If the policymaker is perfectly transparent ($\tau_z \to \infty$), the setup is equivalent to the one from before, when firms directly observed the signal y. If the policymaker is perfectly opaque ($\tau_z = 0$), then there is no public signal.

And

$$L_u = E[(\phi^{U*}z - \theta)^2]$$

The average action is given by:

$$\overline{a} = \lambda^*(\phi_1^{I*}\theta + \phi_2^{I*}z) + (1-\lambda^*)\phi^{U*}z$$

So the policymaker observes the signal:

$$y = \lambda^*(\phi_1^{I*}\theta + \phi_2^{I*}z) + (1-\lambda^*)\phi^{U*}z + \varepsilon_y$$

From which he constructs the unbiased signal of the fundamental:

$$s := \frac{y - z\left[\lambda\phi_2^{I*} + (1-\lambda)\phi^{U*}\right]}{\lambda\phi_1^{I*}} = \theta + \frac{\varepsilon_y}{\lambda^*\phi_1^{I*}}$$

With noise equal to $1/[(\lambda^*\phi_1^{I*})^2\tau_y]$. For ease of notation, let $\sigma_y^2 = 1/\tau_y$ and $\sigma_z^2 = 1/\tau_z$ (in this case it is more intuitive to express results in terms of the variance of noise in signals rather than their precision). Then we can write the noise in the central bank's unbiased signal of the fundamental as:

$$\sigma_{cb}^2 := \frac{\sigma_y^2}{(\lambda^*\phi_1^{I*})^2}$$

The public signal observed by firms writes as:

$$z = s + \varepsilon_z = \theta + \frac{\varepsilon_y}{\lambda^*\phi_1^{I*}} + \varepsilon_z$$

And the public signal reveals the fundamental with noise equal to:

$$\sigma_{public}^2 = \frac{\sigma_y^2}{(\lambda^*\phi_1^{I*})^2} + \sigma_z^2 = \sigma_{cb}^2 + \sigma_z^2$$

Denote by σ_{cb}^2 the noise in the policymaker's observation of the fundamental, i.e.:

$$\sigma_{cb}^2 = \frac{\sigma_y^2}{(\lambda^*\phi_1^{I*})^2}$$

And by σ_{public}^2 the variance of the noise with which the public signal (disclosed by the central bank) reveals the fundamental, i.e.:

$$\sigma_{public}^2 = \sigma_{cb}^2 + \sigma_z^2 = \frac{\sigma_y^2}{(\lambda^*\phi_1^{I*})^2} + \sigma_z^2$$

In this setting, I investigate how changes in the degree of transparency (i.e. σ_z^2) affect the informativeness of the central bank's disclosure about the fundamental (σ^2_{public}) and hence expected welfare. Note that because the central bank's information precision about the fundamental is endogenous (as it depends on the agents' equilibrium behavior), changes in the degree of transparency (σ_z^2) have both a direct and an indirect effect on the informativeness of the public signal:

$$\frac{\partial \sigma^2_{public}}{\partial \sigma_z^2} = \underbrace{\frac{\partial \sigma^2_{cb}}{\partial \sigma_z^2}}_{\text{indirect effect}} + \underbrace{1}_{\text{direct effect}}$$

The direct effect reflects the fact that the central bank discloses its information more or less precisely. The indirect effect reflects that the precision of the central bank's private signal of the fundamental is endogenous and depends on the agents' equilibrium behavior (which depends on the central bank's degree of transparency).

Expressed in these terms, the remarks made in section 1.3.1 write as:

Suppose private information is free (c=0). The indirect effect is always negative, but never dominates the direct effect (i.e. $\frac{\partial \sigma^2_{cb}}{\partial \sigma_z^2} \in (-1, 0)$*). Hence, a higher degree of transparency always increases the informativeness of the public signal about the fundamental* ($\frac{\partial \sigma^2_{public}}{\partial \sigma_z^2} > 0$) *— therefore, a higher degree of transparency is always welfare-improving* ($\frac{L_W}{\partial \sigma_z^2} > 0$)*.*

Conversely, suppose private information is costly ($c > 0$)*. Further, suppose that the central bank observes the average action with sufficiently high precision* ($\tau_y > \hat{\tau}_y$*, where* $\hat{\tau}_y$ *is defined in Proposition 1.2). Then, there is a threshold* $\hat{\sigma}_z^2$*, below which a fraction less than one of agents become informed in equilibrium* ($\lambda^* < 1$) *— below this threshold, changes in the degree of transparency no longer affect the informativeness of the central bank's disclosure about the fundamental, because the indirect effect perfectly offsets the direct effect. More specifically,* $\frac{\partial \sigma^2_{cb}}{\partial \sigma_z^2} = -1$*, which implies that* $\frac{\partial \sigma^2_{public}}{\partial \sigma_z^2} = 0$*. Hence, for* $\sigma_z < \hat{\sigma_z}$*, changes in the degree of transparency do not affect welfare* ($\frac{L_W}{\partial \sigma_z^2} = 0$)*.*

It is straightforward to prove the remarks above using similar arguments as in the main text.

B Proofs for Chapter 3

B.1 Proof of Proposition 3.1

Proof: Given the central bank's policy rule in period t (i.e. $\boldsymbol{\rho}_t$) and the history of play up to period t, one can compute the distribution of g_t and argue that this is mean-zero Gaussian $g_t \sim N(0, \sigma^2_{g,t})$. The distribution of g_t is common knowledge among all agents in period t.

Note: The equilibrium behavior of private agents here is similar to the equilibrium behavior of firms in Chapter 2, so the proof is along similar lines.

Firstly, we need to show that each agent acts optimally[1] given his information[2] (i.e. that each agent finds it optimal to act as described in point 2 given that the information structure is the one described in point 1). The argument is similar to the one in Chapter 2, section 2.4. Each agent's optimal action rule specifies:

$$a^*_{i,t} = E[r(\theta_t + g_t) + (1-r)\overline{a}_t | S_{i,t}]$$

We have fixed the allocation of attention such that each agent observes the private signal $s_{i,t} = \theta_t + g_t + \varepsilon_{i,t}$, where each $\varepsilon_{i,t}$ is i.i.d. according to $\varepsilon_{i,t} \sim N(0, var(\theta_t + g_t)/(2^{2\kappa} - 1))$. Note that if all agents ($j \neq i$) act as described in point 2, i.e. if $a^*_{j,t} = \gamma s_{j,t} \forall j \neq i$, then the average action is given by:[3]

$$\overline{a}_t = \int_0^1 a^*_{j,t} dj = \int_0^1 \gamma s_{j,t} dj = \int_0^1 \gamma(\theta_t + g_t + \varepsilon_{j,t}) dj = \gamma(\theta_t + g_t) \tag{B.1}$$

Hence agent i's optimal action rule writes as:

$$\begin{aligned} a^*_{i,t} &= E[r(\theta_t + g_t) + (1-r)\gamma(\theta_t + g_t) | S_{i,t}] \\ a^*_{i,t} &= [r + (1-r)\gamma] E[\theta_t + g_t | S_{i,t}] \end{aligned} \tag{B.2}$$

Given the information structure, Bayesian updating implies that each agent's expectation of $\theta_t + g_t$ is given by:

$$E[\theta_t + g_t | S_{i,t}] = (1 - 2^{-2\kappa}) s_{i,t} \tag{B.3}$$

[1] In step 5 of the timing of events.
[2] Which was chosen in step 2 of the timing of events.
[3] Agent i is infinitesimal, so his action does not affect the average action.

Substituting γ from Proposition 3.1 and (B.3) into (B.2) yields:

$$a_{i,t}^* = \gamma s_{i,t} \tag{B.4}$$

Which proves that each agent i finds it optimal to act according to (B.4) when all other agents $j \neq i$ act according to (B.4), so this indeed specifies an equilibrium action rule given the information structure.[4]

Secondly, we need to show that the allocation of attention described in Proposition 3.1[5] constitutes an equilibrium, i.e. that each agent i in generation t finds it optimal to observe one signal $s_{i,t} = \theta_t + g_t + \varepsilon_{i,t}$, where $\varepsilon_{i,t} \sim N(0, var(\theta_t + g_t)/(2^{2\kappa} - 1))$. Note that if all other agents $j \neq i$ observe the signal $s_{j,t}$, then (as argued above) in step 5 of the game, all other agents ($j \neq i$) act according to $a_{j,t}^* = \gamma s_{j,t}$ and the average action is given by (B.1). Hence, agent i's optimal allocation of attention must solve:

$$\begin{aligned} \min_{K_{i,t}, F_{i,t}, \Omega_{i,t}} \; & E\left[a_{i,t}^* - (r + \gamma - \gamma r)(\theta_t + g_t)\right]^2 \\ \text{Subject to: } & a_{i,t}^* = E[(r + \gamma - \gamma r)(\theta_t + g_t)|S_{i,t}], \\ & S_{i,t} = F_{i,t} X_t + \varepsilon_{i,t}, \\ & \varepsilon_{i,t} \sim N(0, \Omega_{i,t}), \\ & I(S_{i,t}, X_t) \leq \kappa. \end{aligned} \tag{B.5}$$

Which is equivalent to:

$$\begin{aligned} \min_{K_{i,t}, F_{i,t}, \Omega_{i,t}} \; & (r + \gamma - \gamma r)^2 var(\theta_t + g_t|S_{i,t}) \\ \text{Subject to: } & S_{i,t} = F_{i,t} X_t + \varepsilon_{i,t}, \\ & \varepsilon_{i,t} \sim N(0, \Omega_{i,t}), \\ & I(S_{i,t}, X_t) \leq \kappa. \end{aligned} \tag{B.6}$$

Given that uncertainty is Gaussian, it should be without loss of generality to restrict attention to one-dimensional signal vectors centered on the optimal action, i.e. $K_{i,t}^* = 1, F_{i,t}^* = (1 \;\; 1)$ (see the discussion in Chapter 2). In this case, the information processing constraint writes as $\frac{1}{2} log_2 \left[\frac{var(\theta_t + g_t)}{var(\theta_t + g_t|S_{i,t})}\right] \leq \kappa$. Because the objective is strictly increasing in $var(\theta_t + g_t|S_{i,t})$ and the agent wants to minimize it, the information processing constraint binds and an optimal signal vector must lead to

[4] This is also the unique equilibrium given the information structure, see the argument in section 2.8.3.

[5] i.e. $K_t^* = 1, F_t^* = (1 \;\; 1), \Omega_t^* = (var(\theta_t + g_t)/(2^{2\kappa} - 1))$.

$var(\theta_t + g_t|S_{i,t}) = var(\theta_t + g_t)/2^{2\kappa}$. Note that the signal $s_{i,t}$ attains this and is thus optimal.[6] Hence, each agent i finds it optimal to observe the signal $s_{i,t}$ when all other agents $j \neq i$ observe the signal $s_{j,t}$, so the purported allocation of attention[7] indeed constitutes an equilibrium. □

B.2 Proof of Corollary 3.1

Proof: Recall that the central bank's per-period payoff is given by:

$$u_{cb,t} = -(\overline{a}_t - \theta_t - g_t)^2 - \lambda \int_0^1 (a_{i,t} - \overline{a}_t)^2 di \tag{B.7}$$

If agents act as described in Proposition 3.1, then the average action is given by (B.1), so the expectation of the first component of the loss function writes as:

$$E[(\overline{a}_t - \theta_t - g_t)^2] = E[(\gamma - 1)^2(\theta_t + g_t)^2] = (\gamma - 1)^2 var(\theta_t + g_t)$$

While the second component of the expected loss writes as:

$$\begin{aligned} E[\lambda \int_0^1 (a_{i,t} - \overline{a}_t)^2 di] &= \lambda E[\int_0^1 (\gamma s_{i,t} - \gamma\theta_t - \gamma g_t)^2 di)] \\ &= \lambda E[\int_0^1 (\gamma \varepsilon_{i,t})^2 di] \\ &= \lambda\gamma^2 var(\theta_t + g_t)/(2^{2\kappa} - 1) \end{aligned} \tag{B.8}$$

Where the last equality holds because the processing errors are i.i.d. and have variance equal to $var(\theta_t + g_t)/(2^{2\kappa} - 1)$ (see Proposition 3.1). Hence, the central bank's expected payoff in period t is:

$$E[u_{cb,t}] = -\underbrace{\left[(\gamma - 1)^2 + \frac{\lambda\gamma^2}{2^{2\kappa} - 1}\right]}_{\zeta} var(\theta_t + g_t) \tag{B.9}$$

□

[6] Because the signal $s_{i,t}$ is centered on $\theta_t + g_t$ and because the variance of the noise term $\varepsilon_{i,t}$ is equal to $\frac{var(\theta_t+g_t)}{2^{2\kappa}-1}$, it follows that $var(\theta_t + g_t|s_{i,t}) = \frac{var(\theta_t+g_t)}{2^{2\kappa}}$.

[7] i.e. $K_t^* = 1, F_t^* = (1 \ \ 1), \Omega_t^* = (var(\theta_t + g_t)/(2^{2\kappa} - 1))$.

B.3 Proof of Proposition 3.2

Proof: Using Corollary 3.1 and (B.1), we can express the central bank's problem as:

$$\max_{\{\boldsymbol{\rho}_t\}_{t=1}^{\infty}} \zeta\delta^{t-1} var(\theta_t + g_t) \tag{B.10}$$

Subject to:

$$g_t = \boldsymbol{\rho}_t S_{cb,t} \tag{B.11}$$

$$S_{cb,t} = (Z_t, Z_{t-1}, ..., Z_1)^T \tag{B.12}$$

$$Z_t = \gamma(\theta_{t-1} + g_{t-1}) + \varepsilon_t^Z \tag{B.13}$$

Given the Gaussian information structure, the central bank's expectation of θ_t in period t will be a (deterministic) linear function of the signals observed up to period t $(Z_t, Z_{t-1}, Z_{t-2}, ...)$:

$$E[\theta_t|I_{cb,t}] = \eta_0 Z_t + \eta_1 Z_{t-1} + \eta_2 Z_{t-2} + ... \tag{B.14}$$

Where the weight associated with each signal does not depend on the signal realizations (see footnote 10). It is straightforward to note that a myopic policy specifies:

$$g_t^{myopic} = -E[\theta_t|I_t^{cb}]$$

And achieves a per-period expected loss of:

$$E[u_{cb,t}] = -\zeta var(\theta_t|I_t^{cb}) = -\zeta\frac{1}{\tau_t^{cb}} \tag{B.15}$$

Then, note from Corollary 3.2 that the myopic policy is optimal (because the central bank's information precision about the fundamental in period T is independent of its policy reaction function up to period T and there are no other intertemporal links). Hence, the central bank sets g_t in each period to minimize $var(\theta_t + g_t)$ conditional on its information in period t, i.e. it sets:

$$g_t^* = -E[\theta_t|I_{cb,t}] = -\eta_0 Z_t - \eta_1 Z_{t-1} - \eta_2 Z_{t-2} + ... \tag{B.16}$$

We can use the central bank's steady-state information precision (characterised in Corollary 3.2) to determine the coefficients $(\eta_0, \eta_1, \eta_2, ...)$ in the steady-state. Consider a generic period t and suppose that the central bank's information precision has reached the steady-state. Recall from (3.28) that the central bank's expectation of θ_t in period t is:

$$E[\theta_t|I_t^{cb}] = \alpha E[\theta_{t-1}|I_t^{cb}] \tag{B.17}$$

In turn, the central bank's expectation of θ_{t-1} in period t is a weighted sum of its expectation of θ_{t-1} in period $t-1$ and the unbiased signal of θ_{t-1} constructed by the central bank from Z_t in period t (which was denoted by s_t, see (3.25)):

$$\begin{aligned} E[\theta_{t-1}|I_t^{cb}] &= (1-\phi)E[\theta_{t-1}|I_{t-1}^{cb}] + \phi s_t \\ &= (1-\phi)E[\theta_{t-1}|I_{t-1}^{cb}] + \phi(\frac{Z_t}{\gamma} - g_{t-1}) \end{aligned} \tag{B.18}$$

Because the central bank's information precision has reached the steady-state:

$$var(\theta_{t-1}|I_{t-1}^{cb}) = 1/\tau_{ss}^{cb} \tag{B.19}$$

Also, note that the noise with which s_t reveals the fundamental is independent of the central bank's estimate of the fundamental up to period t and has variance σ_Z^2/γ^2 (see (3.26)), so Bayesian updating implies that:

$$\phi = \frac{\gamma^2/\sigma_z^2}{\gamma^2/\sigma_z^2 + \tau_{ss}^{cb}} \tag{B.20}$$

We are looking for optimal policy in the steady-state, so using the fact that g_{t-1} was set optimally in period $t-1$ (i.e. $g_{t-1} = g_{t-1}^* = -E[\theta_{t-1}|I_{t-1}^{cb}]$), we can write (B.18) as:

$$E[\theta_{t-1}|I_t^{cb}] = E[\theta_{t-1}|I_{t-1}^{cb}] + \phi\frac{Z_t}{\gamma} \tag{B.21}$$

Furthermore, because the central bank's information precision has reached the steady-state, the central bank used the same linear rule (B.14) to form its expectation of θ_{t-1} in period $t-1$:

$$E[\theta_{t-1}|I_{t-1}^{cb}] = \eta_0 Z_{t-1} + \eta_1 Z_{t-2} + \eta_2 Z_{t-3} + ... \tag{B.22}$$

Substituting (B.22) into (B.21) and the resulting expression into (B.17) yields:

$$E[\theta_t|I_t^{cb}] = \alpha\left[\frac{\phi}{\gamma}Z_t + \eta_0 Z_{t-1} + \eta_1 Z_{t-2} + \eta_2 Z_{t-3} + ...\right] \tag{B.23}$$

Matching coefficients in (B.23) with (B.14) gives:

$$
\begin{aligned}
\eta_0 &= \frac{\alpha\phi}{\gamma} \\
\eta_1 &= \alpha\eta_0 = \frac{\alpha^2\phi}{\gamma} \\
\eta_2 &= \alpha\eta_1 = \frac{\alpha^3\phi}{\gamma} \\
&\vdots
\end{aligned} \tag{B.24}
$$

Which given (B.16) implies that we can write optimal policy in period t as:

$$
g_t^* = -\frac{\alpha\phi}{\gamma} Z_t - \frac{\alpha^2\phi}{\gamma} Z_{t-1} - \frac{\alpha^3\phi}{\gamma} Z_{t-2} - ... \tag{B.25}
$$

Matching coefficients in (B.25) with the central bank's policy rule (3.8) yields Proposition 3.2. □

B.4 Proof of Proposition 3.3

Proof: The first part of the proof (concerning the agents' equilibrium behavior conditional on the central bank's reaction function) is analogous to the proof of Proposition 3.1.

In a similar fashion to the proof of Corollary 3.1, it is then straightforward to argue that the central bank's expected payoff is decreasing in the variance of $\theta + bg$ — as agents take their actions according to $a_i^* = \gamma s_i$, it follows that the average action is given by $\overline{a} = \gamma(\theta + bg)$, so the expectation of the first component of the central bank's loss function again writes as:

$$
E[(\overline{a} - \theta - bg)^2] = (\gamma - 1)^2 var(\theta + bg)
$$

While the second component of the expected loss writes as:

$$
\begin{aligned}
E[\lambda \int_0^1 (a_i^* - \overline{a})^2 di] &= \lambda E[\int_0^1 (\gamma s_i - \gamma\theta - \gamma bg)^2 di)] \\
&= \lambda E[\int_0^1 (\gamma\varepsilon_i)^2 di] \\
&= \lambda E[\gamma^2 var(\theta + bg|b)/(2^{2\kappa} - 1)] \\
&= \lambda\gamma^2 var(\theta + bg)/(2^{2\kappa} - 1)
\end{aligned} \tag{B.26}
$$

Where the penultimate equality holds because the processing errors are i.i.d. and have variance equal to $var(\theta + bg|b)/(2^{2\kappa} - 1)$, while the last equality follows because $E[var(\theta + bg|b)] = var(\theta + bg)$.[8] It follows that the expected welfare loss is proportional to $var(\theta + bg)$, so we can express the central bank's problem as:

$$\max_{\rho} -\zeta var(\theta + bg) \tag{B.27}$$

Subject to:

$$g = \rho x \tag{B.28}$$

Where ζ is characterised in Corollary 3.1. After substituting the constraint into the objective function and simplifying, the central bank's problem is equivalent to:

$$\max_{\rho} \sigma_\theta^2[(1+\rho)^2 + \rho^2\sigma_b^2] + \sigma_x^2\rho^2(1+\sigma_b^2) \tag{B.29}$$

Maximizing this readily yields:

$$\rho^* = -\frac{1}{1+\sigma_b^2}\frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_x^2} \tag{B.30}$$

Noting that $E[\theta|x] = \frac{\sigma_\theta^2}{\sigma_\theta^2+\sigma_x^2}x$, the policy instrument is optimally set according to:

$$g^* = -\frac{1}{1+\sigma_b^2}E[\theta|x]. \tag{B.31}$$

□

[8] It is straightforward to prove the last statement, for instance, by using the law of total variance to express $var(\theta + bg) = E[var(\theta + bg)|b] + var[E(\theta + bg|b)]$. Because θ and g are mean-zero (for any policy reaction function ρ), $E[\theta + bg|b] = 0$, so $var(\theta + bg) = E[var(\theta + bg)|b]$.

References

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

Afrouzi, H. and C. Yang (2021). Dynamic rational inattention and the phillips curve.

Aoki, K. (2003). On the optimal monetary policy response to noisy indicators. Journal of Monetary Economics 50(3), 501–523.

Baeriswyl, R. (2011). Endogenous central bank information and the optimal degree of transparency. International Journal of Central Banking 7(2), 85–111.

Baeriswyl, R. and C. Cornand (2011). Transparency and monetary policy effectiveness. Annals of Economics and Statistics (103/104), 175–194.

Baeriswyl, R., C. Cornand, and B. Ziliotto (2020). Observing and shaping the market: The dilemma of central banks. Journal of Money, Credit and Banking 52(8), 1973–2005.

Ball, L., G. Mankiw, and R. Reis (2005). Monetary policy for inattentive economies. Journal of Monetary Economics 52(4), 703–725.

Balvers, R. J. and T. F. Cosimano (1994). Inflation variability and gradualist monetary policy. The Review of Economic Studies 61(4), 721–738.

Batini, N., B. Martin, and C. Salmon (1999). Monetary policy and uncertainty. Bank of England Quarterly Bulletin.

Bernanke, B. S. and M. Woodford (1997). Inflation forecasts and monetary policy. Journal of Money, Credit and Banking 29(4), 653–684.

Bertocchi, G. and M. Spagat (1993). Learning, experimentation and monetary policy. Journal of Monetary Economics 32, 169–183.

Bond, P. and I. Goldstein (2015). Government intervention and information aggregation by prices. Journal of Finance 70(6), 2777–2812.

Brainard, W. C. (1967). Uncertainty and the effectiveness of policy. American Economic Review 57(2), 411–425.

Bru, L. and X. Vives (2002). Informational externalities, herding, and incentives. Journal of Institutional and Theoretical Economics (JITE) / Zeitschrift für die gesamte Staatswissenschaft 158(1), 91–105.

Burguet, R. and X. Vives (2000). Social learning and costly information acquisition. Economic Theory 15(1), 185–205.

Ellison, M. and N. Valla (2001). Learning, uncertainty and central bank activism in an economy with strategic interactions. Journal of Monetary Economics 48(1), 153–171.

Fulton, C. (2017). Mechanics of linear quadratic gaussian rational inattention tracking problems. Finance and economics discussion series, Board of Governors of the Federal Reserve System (U.S.).

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. American Economic Review 70(3), 393–408.

Hayek, F. A. (1945). The use of knowledge in society. American Economic Review 35(4), 519–530.

Iovino, L., J. La'O, and R. Mascarenhas (2021). Optimal monetary policy and disclosure with an informationally-constrained central banker. Journal of Monetary Economics.

James, J. G. and P. Lawler (2011). Optimal policy intervention and the social value of public information. American Economic Review 101(4), 1561–74.

Judd, K. L. (1985). The law of large numbers with a continuum of iid random variables. Journal of Economic Theory 35(1), 19–25.

Kacperczyk, M., S. V. Nieuwerburgh, and L. Veldkamp (2016). A rational theory of mutual funds' attention allocation. Econometrica 84(2), 571–626.

Maćkowiak, B., F. Matějka, and M. Wiederholt (2018). Dynamic rational inattention: Analytical results. Journal of Economic Theory 176, 650–692.

Maćkowiak, B. and M. Wiederholt (2007). Optimal sticky prices under rational inattention. CEPR Discussion Papers 6243, C.E.P.R. Discussion Papers.

Maćkowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. American Economic Review 99(3), 769–803.

Morris, S. and H. S. Shin (2002). Social value of public information. American Economic Review 92(5), 1521–1534.

Morris, S. and H. S. Shin (2005). Central bank transparency and the signal value of prices. Brookings Papers on Economic Activity 2005(2), 1–43.

Morris, S. and H. S. Shin (2018). Central bank forward guidance and the signal value of market prices. AEA Papers and Proceedings 108, 572–77.

Nimark, K. (2008). Monetary policy with signal extraction from the bond market. Journal of Monetary Economics 55(8), 1389–1400.

Paciello, L. and M. Wiederholt (2013). Exogenous information, endogenous information, and optimal monetary policy. The Review of Economic Studies 81(1), 356–388.

Sack, B. P. (1998). Uncertainty, learning, and gradual monetary policy. Finance and Economics Discussion Series 1998-34, Board of Governors of the Federal Reserve System (U.S.).

Samuelson, P. (1994). Goals, guidelines and constraints facing monetary policymakers. Federal Reserve Bank of Boston Conference Series (38) panel discussion.

Sims, C. A. (2003). Implications of rational inattention. Journal of Monetary Economics 50(3), 665–690.

Sims, C. A. (2010). Chapter 4 - rational inattention and monetary economics. Volume 3 of Handbook of Monetary Economics, pp. 155–181. Elsevier.

Song, W. and S. Stern (2022). Firm inattention and the efficacy of monetary policy: A text-based approach. Staff Working Papers 22-3, Bank of Canada.

Tenreyro, S. and G. Thwaites (2016, October). Pushing on a string: US monetary policy is less powerful in recessions. American Economic Journal: Macroeconomics 8(4), 43–74.

Theil, H. (1958). Economic forecasts and policy. North-Holland Amsterdam.

Uhlig, H. (1996). A law of large numbers for large economies. Economic Theory 8(1), 41–50.

Vives, X. (1996). Social learning and rational expectations. European Economic Review 40(3-5), 589–601.

Vives, X. (1997). Learning from others: A welfare analysis. Games and Economic Behavior 20(2), 177–200.

Vives, X. (2010). Information and learning in markets: the impact of market microstructure. Princeton University Press.

Woodford, M. (2001). Imperfect common knowledge and the effects of monetary policy. NBER Working Papers 8673, National Bureau of Economic Research, Inc.