Policy with Dispersed Information[*]

George-Marios Angeletos

MIT and NBER

Alessandro Pavan

Northwestern University

September 5, 2008

Abstract

Information regarding aggregate economic fundamentals is widely dispersed in society, is only imperfectly aggregated through prices or other indicators of aggregate activity, and can not be centralized by the government or any other institution. In this paper we seek to identify policies that can improve the decentralized use of such dispersed information without requiring the government to observe this information. We show that this can be achieved by appropriately designing the contingency of taxation on ex post public information regarding the realized fundamentals and aggregate activity. When information is common (as in the Ramsey literature) or when agents have private information regarding only idiosyncratic shocks (as in the Mirrlees literature), the contingency on fundamentals alone suffices for efficiency. When instead agents have private information regarding aggregate shocks, the contingency on aggregate activity becomes crucial. An appropriate combination of the two contingencies then permits the government to achieve the following goals: (i) dampen the impact of noise on equilibrium activity, and hence reduce non-fundamental volatility, without also dampening the impact of fundamentals; (ii) induce agents to internalize informational externalities, and hence improve the speed of social learning; (iii) restore a certain form of constrained efficiency in the decentralized use of information; and (iv) guarantee that welfare increases with the provision of any additional information.

JEL codes: C72, D62, D82.

Keywords: Optimal policy, private information, complementarities, information externalities, social learning, efficiency.

1 Introduction

Information regarding commonly-relevant fundamentals—such as aggregate productivity and demand conditions over the business cycle, or the profitability of a new technology—is widely dispersed and only imperfectly aggregated in society. As emphasized by Hayek (1945), such information can not be centralized by the government; instead, society must rely on decentralized mechanisms for the utilization of such information. This, however, does not mean that the decentralized use of information is necessarily the one that best serves social interests.

Long ago, Keynes (1936) argued that financial markets are excessively volatile because professional investors are more concerned with second-guessing the demands of one another, and hence with forecasting the forecasts of others, than with forecasting the fundamental value of the assets they trade.[1] More recently, Morris and Shin (2002) used this line of reasoning to argue that financial markets overreact to noisy public news because they help forecast one another's actions; when this is the case, the provision of public information—e.g., via more transparency in central-bank communications—can reduce welfare. (See also Angeletos and Pavan (2007) for a more extensive analysis of the social value of information.) Turning attention to the business cycle, the latter may be driven, not only by variation in fundamentals, but also by noise in the agents' expectations about these fundamentals as well as about the choices of other agents, possibly leading to excessive non-fundamental volatility. Finally, individuals are unlikely to internalize how their own choices affect the information of others through financial prices, macroeconomic data, and other forms of social learning; if they could be persuaded to base their decisions more on their idiosyncratic sources of information, social learning could become more efficient, leading to less noise and higher welfare. Banerjee (1992) and Vives (1993, 1997) were among the first to emphasize how the failure to internalize such informational externalities can lead to excessive herding and suboptimal social learning. (See also Chamley (2004).) Chari and Kehoe (2003) study how this failure can amplify volatility in financial markets, while Amador and Weill (2007) show how it can also make the provision of public information have a negative effect on welfare by slowing down social learning.

Motivated by these observations, this paper seeks to identify policies that help the government control how agents utilize their dispersed sources of information regarding aggregate fundamentals. Of course, this goal could be achieved easily if the government could observe these sources of information, for it could then impose *direct* taxes on their utilization; but such direct taxes are clearly not plausible. We thus seek to identify policies that achieve the same goal in an *indirect* way, without requiring the government to observe these sources of information. Our contribution is to show that this can be done by appropriately designing the contingency of taxation on public information regarding the realized fundamentals and, more crucially, the realized aggregate activity.

---

[1]Elements of this "beauty-contest" character of financial markets have been formalized recently in Allen, Morris and Shin (2003), Angeletos, Lorenzoni and Pavan (2007), and Bacchetta and Wincoop (2005).

The logic is simple. The anticipation of these contingencies affects the incentives the agents face when they decide how to use their dispersed sources of information. The contingency on fundamentals has a symmetric effect across all sources of information: the more an agent expects marginal taxes to increase with fundamentals, the less he responds to any information about the fundamentals. In contrast, the contingency on aggregate activity has an asymmetric effect: it penalizes the agents relatively more when they react to signals whose errors are highly correlated across the agents. This is because the contingency on aggregate activity, unlike the one on fundamentals, affects the degree of strategic complementarity featured in equilibrium. Indeed, the more marginal taxes are expected to increase with realized aggregate activity, the weaker the complementarity agents perceive in their choices, and hence the weaker the incentive to react to sources of information that help forecast one another's beliefs and actions; and because it is precisely sources of information with highly correlated noise that are relatively better predictors of others' beliefs and actions, this contingency penalizes relatively more the use of such sources of information. Importantly, this is achieved only in an indirect way: despite the fact that direct taxes on the use of the different sources of information are not feasible, the contingency of taxes on realized aggregate activity provides similar incentives.

An appropriate design of the two contingencies thus permits the government to dampen the impact of noise without also dampening the impact of fundamentals; to improve the speed of social learning; and, in overall, to restore efficiency in the decentralized use of information. This in turn also guarantees that welfare increases with any additional information, whether private or public, thus helping the government bypass the complications considered, inter alia, by Morris and Shin (2002), Angeletos and Pavan (2007) and Amador and Weill (2007).

These insights are not limited to any specific application. In this paper we thus opt to provide some general lessons that may be useful across a variety of applications. We start in the next section by illustrating the incentive effects of the aforementioned contingencies within a simple investment example. We then proceed to an abstract framework that allows, subsequently, for both rich payoff interactions and informational externalities across the agents. These two steps shed further light on the generality of the insights, on the circumstances under which it is important for optimality to use the aforementioned policy contingencies, and on the novelty of our policy exercise.

Related Literature. Although there is a long history in studying informational frictions in macroeconomics (e.g., Phelps, 1970; Lucas, 1972; Townsend, 1983; Woodford, 2002), to the best of our knowledge this paper is the first one to study optimal taxation with dispersed information on aggregate shocks. This is unlike either the Ramsey literature (e.g., Barro, 1979; Lucas and Stokey, 1983; Chari, Christiano and Kehoe, 1994), which does not allow for any private information, or the Mirrlees/New Dynamic Public Finance literature (e.g., Kocherlakota, 2005), which allows private information only on idiosyncratic shocks. By ruling out private information on aggregate shocks, these literatures have also ruled out the type of inefficiencies and policy objectives that we consider.

Complementary in this respect are Angeletos and La'O (2008), Lorenzoni (2008), and Angeletos, Lorenzoni and Pavan (2007). The first two papers study micro-founded business-cycle models that introduce dispersed information about aggregate productivity; the third one studies the interaction between real investment and financial markets in an economy where entrepreneurs and financial traders have dispersed information about the profitability of a new technology. The policy results in these more applied works verify that our methodology and key policy insights are not limited to the particular reduced-form framework employed in this paper, nor to taxation as the only relevant policy instrument. Indeed, the broader lesson from this paper is how the contingencies of macroeconomic policies on ex post information regarding the realized fundamentals and, crucially, the realized aggregate activity can improve efficiency in the decentralized use of information.

Related is also the literature on efficient implementation with correlated information and inter-dependent valuations (see, among others, Cremer and McLean, 1985, McAffee and Reny, 1992, Jehiel and Moldovanu, 2001, McLean and Postlewaite, 2002, 2003, 2004). While the information structures considered are similar, the efficiency concept we employ here is very different in that we do not allow the planner to transfer information across agents or, equivalently, to send recommendations to an agent that depend on the information received from the other agents. We find this restriction appropriate when studying the properties of optimal tax schemes in large economies: while it seems plausible that the government could use the contingency of taxes on aggregate outcomes to manipulate the way information is used in equilibrium, it does seem plausible that the government is able to consult with all agents and transfer information across them before the latter make their investment, production, and consumption decisions. Our efficiency concept thus shares with Hayek (1945) and Radner (1962) the idea that information is dispersed and cannot be communicated to a "center." Similar efficiency concepts have been used to study the welfare properties of large Cournot games (Vives, 1988), of social learning (Vives, 1997), and of rational-expectations equilibria (Laffont, 1985; Messner and Vives, 2005).

2 An example and a preview

We start with an example that illustrates how the anticipation of the aforementioned policy contingencies affects the incentives agents face when deciding how to react to their dispersed sources of information regarding the underlying common fundamentals.

There is a large number of risk-neutral agents, each choosing how much to invest in a technology with unknown productivity θ (the commonly-relevant fundamental for this example). Investing k units costs $\frac{1}{2}k^2$ of the consumable good in one period and delivers θk in the next period, so that agent i's payoff is

$$u_i = (e - \tfrac{1}{2}k_i^2) + \beta(\theta k_i - \tau_i),$$

where k_i denotes his investment, e his endowment of the consumable good in the first period, $\beta \in (0,1)$ his discount factor, and τ_i his tax payments. Each agent receives two private signals about θ, one with only idiosyncratic noise and another with partly common noise: $x_i = \theta + \xi_i$ and $y_i = \theta + \varepsilon + \epsilon_i$, where the noises ξ_i and ϵ_i are independent across agents while ε is common. The productivity θ and all these noises are Normally distributed, and independent of one another.

The government has no information at the time agents make their choices, nor can it collect the information that is dispersed among them. It can only commit to tax schedules that are contingent on information that will become publicly available in the second period, *after* agents have made their choices. Suppose that both the fundamentals and the agents' actions become public information at that stage and consider tax schedules of the form $\tau_i = tk_i - L$, where L is a lump-sum transfer and $t = t_\theta \theta + t_K K$ is a proportional tax. The coefficients t_θ and t_K parameterize the contingencies of the tax on the realizations of the fundamental and aggregate investment. Imposing budget balance gives $L = tK - G$, where G is the exogenous level of government spending.

Let $\mathbb{E}_i$ denote the expectation of agent i conditional on his two signals, x_i and y_i. In the absence of policy, his optimal investment would have been $k_i = \beta \mathbb{E}_i \theta$; now it is given by

$$k_i = \beta \mathbb{E}_i[(1 - t_\theta)\theta - t_K K]. \tag{1}$$

Because of the linearity of this condition and the Gaussian specification of the information, an educated guess is that the equilibrium investment of an agent is linear in his two signals. Thus suppose there exist coefficients $(\gamma_0, \gamma_x, \gamma_y)$ such that the equilibrium investment strategy is

$$k_i = \gamma_0 + \gamma_x x_i + \gamma_y y_i. \tag{2}$$

Aggregate investment is then given by $K = \gamma_0 + (\gamma_x + \gamma_y)\theta + \gamma_y \varepsilon$. Substituting the latter into condition (1) gives the best response to the strategy specified in (2). Requiring that the two coincide, so that the strategy specified in (2) is indeed an equilibrium, gives the following equilibrium values for the coefficients γ_x and γ_y:

$$\gamma_x = (1 - t_\theta)(1 + t_K \rho_y)\pi_x \Gamma(t_K) \quad \text{and} \quad \gamma_y = (1 - t_\theta)\pi_y \Gamma(t_K),$$

where $\Gamma(t_K)$ is a decreasing function of t_K, π_x and π_y are the precisions of the two signals, and ρ_y is the correlation across agents of the noises in the second signal.[2]

Note that a higher t_θ reduces γ_x and γ_y proportionally; this is because the more agents expect their marginal taxes to increase with realized productivity, the less their incentive to react to any source of information regarding productivity. In contrast, a higher t_K has an asymmetric effect,

---

[2]Formally, $\rho_y = Corr(y_i - \theta, y_j - \theta) \ \forall \ i \neq j$. Also, the formula for $\Gamma(t_K)$ and all the results of this section can be obtained as a special case of the more general results in the proof of Proposition 3.

reducing γ_y more so than γ_x; this is because each agent has an incentive to react relatively less to sources of information that have a lot of common noise when he expects marginal taxes to be positively correlated with realized aggregate activity, which in turn is positively correlated, in equilibrium, with such common noise. It follows that the government can control the reaction of investment to the noise and the fundamentals by appropriately designing the two contingencies. In particular, suppose the government sets $t_K > 0$ so as to reduce γ_y/γ_x, while also setting $t_\theta < 0$ so as to keep $\gamma_x + \gamma_y$ constant. This ensures that agents rely less on the signal with the most correlated noise (y) and more on the signal with the least correlated noise (x), so that at the end equilibrium investment reacts less to the underlying common noise even though it reacts the same to the underlying fundamentals. In contrast, if the government could use only the contingency on the fundamentals (the contingency that is more familiar from the pertinent literature), then it could reduce the impact of noise only at the expense of reducing also the impact of fundamentals.

To further appreciate the distinctive role of the two contingencies and how they may affect the speed of social learning, consider the "signal-to-noise" ratio in aggregate investment (that is, the ratio of the volatility that is caused by variation in the underlying fundamentals over the volatility that is caused by variation in the underlying noise):

$$\frac{Var(K|\varepsilon)}{Var(K|\theta)} = \left(1 + \frac{\gamma_x}{\gamma_y}\right)^2 \frac{Var(\theta)}{Var(\varepsilon)}.$$

Clearly, this signal-to-noise ratio is independent of t_θ but increases with t_K. This also suggests that the latter contingency is an important instrument through which the government may be able to control how much agents (and the government itself) can learn about the underlying fundamentals from indicators of aggregate economic activity.

To recap, this example illustrates how the aforementioned policy contingencies can affect the decentralized use of information. However, this example does not help understand when it may be desirable to do so, nor when their combination is essential. Moreover, this example rules out any payoff or informational interactions among the agents, such as the ones that obtain through trading in markets or other forms of social interaction. Not only are such payoff and informational interactions central to applications, but also their absence would eliminate any reason for policy intervention. To address these issues, we proceed as follows in the rest of the paper.

We start in Section 3 with an abstract framework that rules out informational externalities but allows for rich payoff interactions. This framework is flexible enough to capture, in reduced form, the role played by dispersed information in a variety of applications. It thus helps identify some general principles regarding the impact of dispersed information on equilibrium, efficiency, and policy—principles that are likely to hold across a variety of contexts.

We used a close variant of this framework in Angeletos and Pavan (2007) to study the social value

of information under a particular Gaussian specification for the information structure. Whereas that paper abstracted from policy, the contribution of the present paper is precisely the policy exercise. To highlight the novelty of this exercise relatively to the pertinent policy literature, we now allow for both idiosyncratic and aggregate shocks, as well as for arbitrary information structures.

In Section 4 we revisit the characterization of equilibrium and efficient allocations for the more general structure of shocks and information considered in the present paper. This permits us to formalize a particular form of inefficiency that can emerge when, and only when, agents have private information regarding aggregate shocks. This inefficiency has been ruled out by the pertinent policy literature, shares certain features with the one conjectured by Keynes, and manifests itself as excessive non-fundamental volatility (low signal-to-noise ratio) in aggregate activity.

In Section 5 we turn to policy. We first show, for an arbitrary policy, how thecontingencies of the tax schedule on realized aggregate outcomes affects the incentives agents face when deciding how to react to different sources of information; this generalizes the insights of the investment example considered above. We then proceed to study optimal policy; we do so by identifying the policies that implement the efficient use of information as an equilibrium. This approach may prove useful for studying optimal policy in a variety of applications that feature dispersed information on aggregate shocks. Here, we use it to prove a simple but important policy principle: the contingency on aggregate activity is essential for achieving efficiency—or, more generally, for implementing any feasible allocation—*only* when agents have dispersed information regarding aggregate shocks. When, instead, agents have either no private information (as in the Ramsey literature) or private information regarding only idiosyncratic shocks (as in the Mirrlees literature), it suffices to make the tax schedule contingent on the aggregate fundamentals.

In Section 6 we extend the analysis to a dynamic setting where information is imperfectly aggregated through signals of aggregate activity; these are short-cuts for financial prices, macro data, and other sources of social learning. A new inefficiency emerges as agents do not internalize how their own choices affect the quality of information contained in those signals: social learning would be faster if equilibrium activity reacted more to fundamentals and less to noise. Once again, this inefficiency relies on private information regarding aggregate shocks: had the agents had private information only on idiosyncratic shocks, then nothing new could be learned about aggregate shocks from aggregating their private information. Our key result is that the contingency of the policy on realized aggregate activity is instrumental for correcting this type of inefficiency as well.

Finally, in Section 7 we discuss how the policies identified here also guarantee that equilibrium welfare necessarily increases with any additional information that the government may be able to collect and disclose to the market. In contrast, without the policies we identify here, one could not guarantee that the government should provide the market with more information, even if it were costless to collect such information.

3 The baseline framework

Actions and payoffs. The economy is populated by a continuum of risk-neutral agents of measure one, indexed by $i \in [0, 1]$, each choosing an action $k_i \in \mathbb{R}$. In addition, there is a government, which imposes a tax $\tau_i \in \mathbb{R}$ on each agent i, subject to the usual budget constraint. The agent's payoff depends on his own action, the actions of others, the tax he pays, and exogenous fundamentals:

$$u_i = V(k_i, K, \sigma_k, \theta_i, \bar{\theta}) - \tau_i, \tag{3}$$

where K and σ_k denote, respectively, the average and the dispersion of this action in the cross-section of the population (i.e., the first moment and the square root of the second moment of the cross-sectional distribution of k), $\theta_i \in \Theta \subseteq \mathbb{R}$ is an exogenous fundamental ("shock") specific to agent i, and $\bar{\theta}$ is the average shock in the population. For concreteness, we can think of k_i as investment, θ_i as individual productivity, and $\bar{\theta}$ as aggregate productivity; however, the interpretation will vary from application to application. Finally, $V : \mathbb{R}^2 \times \mathbb{R}_+ \times \Theta^2 \to \mathbb{R}$ is a strictly concave quadratic polynomial and its derivatives satisfy $V_\sigma(\cdot) = V_{\sigma\sigma}\sigma$, $V_{kk} + V_{\sigma\sigma} < 0$, and $V_{kK} < -V_{kk}$. These properties ensure existence and uniqueness of equilibrium and efficient allocations; they also keep the analysis tractable by ensuring that the first-order conditions that characterize these allocations are linear in $(k, K, \theta, \bar{\theta})$ and independent of σ_k. We let $\mathcal{V}$ denote the set of payoff functions V that satisfy these properties.

Interpretation. This game is meant to be a reduced-form representation of richer applications: a variety of market interactions may be "hidden" behind our reduced-form game, so that the payoff interdependences embedded in V may originate, not only in direct externalities in preferences or technologies, but also in pecuniary externalities, monopoly power, credit frictions, and the like.

To illustrate, consider the following example. There is a continuum of households, each consisting of a consumer and a producer, and two goods, one of which could be interpreted as leisure. Let q_{1i} and q_{2i} denote the respective quantities consumed by household i. His utility is given by $u_i = v(q_{1i}, \theta_i) + q_{2i}$, where $v(q, \theta) = \theta q - q^2/2$. The term θ_i represents a taste shock in the relative demand for the two goods. His budget is $pq_{1i} + q_{2i} = pe_1 + e_2 + \pi_i$, where p is the price of good 1 relative to good 2, e_1 and e_2 are exogenous endowments, and π_i are the profits of the producer living in household i. These profits are given by $\pi_i = pk_i - C(k_i)$, where k_i is the quantity of good 1 produced and $C(k) = k^2/2$ its cost in terms of good 2. While production may take place under dispersed information about the taste shocks, exchange and consumption decisions take place after these shocks have been revealed. At that point, household i's demand for good 1 is $q_{1i} = \theta_i - p$ and the corresponding aggregate demand is $Q_1 = \bar{\theta} - p$. Since market clearing imposes $Q_1 = K + e_1$, the equilibrium price must satisfy $p = P(K, \bar{\theta}) \equiv \bar{\theta} - K - e_1$. Using this result, we have that the

utility of household i utility reduces to

$$u_i = V(k_i, K, \theta_i, \bar{\theta}) \equiv v(\theta_i - P(K, \bar{\theta}), \theta_i) + e_2 + P(K, \bar{\theta})[(k_i - K) - (\theta_i - \bar{\theta})] - C(k_i),$$

which is readily nested in our framework. Clearly, in this example the interdependence of payoffs emerges in trading (and the associated pecuniary externalities), not any direct technological or preference externality.

As an alternative example, consider a Dixit-Stiglitz monopolistic economy of the type that is now standard in business-cycle theory. Such an economy can often be reduced to a game in the pricing or production choices of different firms. Within the context of business cycles, one can thus interpret k as the level of employment or the price set by a firm. The interdependence of payoffs then typically emerges from two channels: the fact that the demand for the product of the individual firm depends on aggregate demand; and the fact that the wage depends on aggregate income. Of course, the primitive model will typically feature rich micro-foundations and the reduced-form game may be linear-quadratic only after a certain log transformation. Nevertheless, these complications need not crucially affect the applicability of our main policy results. Indeed, as mentioned in the Introduction, the results in Angeletos and La'O (2008) and Lorenzoni (2008) show that both the methodology and the key policy insights of this paper can be adapted to fully micro-founded business-cycle economies.

Timing. There are three stages. In stage 1, the government announces a policy rule T that specifies how taxes will be collected in stage 3 as a function of the information that will be public by then. In stage 2, agents simultaneously choose their actions k_i under the information structure described below. Finally, in stage 3, individual actions and the average fundamental $\bar{\theta}$ are publicly revealed, taxes are collected according to T, payoffs are realized, and the game ends.

Information Structure. Let Ω denote a set of possible "signals" or "types" for each agent i, F a set of probability distributions over $\Theta \times \Omega$ and $\mathcal{P}$ a probability measure over F.[3] Nature first draws f from F using the probability measure $\mathcal{P}$ and then uses f to make independent draws of pairs $(\theta, \omega) \in \Theta \times \Omega$, one for each agent. Given f, let $h \in H$ denote the corresponding marginal distribution of θ and $\phi \in \Phi$ the corresponding marginal distribution of ω. We assume that the probability distribution f coincides with the distribution of (θ, ω) in the cross-section of the population; the average shock is thus equal to $\bar{\theta} = \int \theta dh(\theta)$. Furthermore, given any $f \in F$ and any (measurable) strategy $k : \Omega \to \mathbb{R}$ for the agents, we assume that the the cross-sectional mean and dispersion of activity induced by $k(\cdot)$ are respectively $K(\phi) = \int k(\omega)d\phi(\omega)$ and $\sigma_k(\phi) = [\int (k(\omega) - K(\phi))^2 d\phi(\omega)]^{\frac{1}{2}}$.[4] In the following,

---

[3]While we are not imposing any specific structure on the set Ω, we are implicitly assuming it is a well-behaved set over which probability measures are well defined; for concreteness, one can think of Ω as a subset of $\mathbb{R}^n$ ($n \in \mathbb{N}$).

[4]These assumptions are standard in games with a continuum of players. In certain cases, such as the linear-Gaussian example described below, these assumptions can be justified through generalizations of the Strong Law of Large Numbers (see the technical Appendix in Vives, 2008, for a discussion). Also note that, while we are restricting

we refer to f as the "aggregate state of the world," h as the "aggregate economic fundamentals," and ϕ as the " distribution of information in society." The description of the information structure is then completed by assuming that, while all the facts described above are common knowledge, the particular realization of the aggregate state f need not be; instead, each agent privately observes his own ω and then uses this observation along with the aforementioned facts to form posterior beliefs about both his own shock θ and the underlying aggregate state f.

Note that this formalization is highly flexible: ω_i can encode arbitrary information about i's own productivity and about the joint distribution of productivities and information in the population. To illustrate, consider the following Gaussian example which is often assumed in applications. Agent i's productivity is given by $\theta_i = \bar{\theta} + \varsigma_i$, where $\bar{\theta}$ is Normally distributed with mean μ_θ and variance σ_θ^2, while ς_i is i.i.d. across i, independent of $\bar{\theta}$, Normally distributed with zero mean and variance σ_ς^2. Each agent i's information $\omega_i = (z_i, x_i, y_i)$ consists of a private signal $z_i = \theta_i + \zeta_i$ about own productivity, a private signal $x_i = \bar{\theta} + \xi_i$ about aggregate productivity, and a public signal $y_i = y = \bar{\theta} + \varepsilon$ about aggregate productivity. The idiosyncratic noises ζ_i and ξ_i are i.i.d. across i, Normally distributed with zero mean and variances σ_z^2 and σ_x^2, respectively, whereas the common noise ε is Normally distributed with zero mean and variance σ_y^2; all these noises are independent of one another, as well as of $\bar{\theta}$ and of ς_i. In this example, $(\mu_\theta, \sigma_\theta, \sigma_z, \sigma_\varsigma, \sigma_x, \sigma_y)$ are fixed parameters, $\Theta = \mathbb{R}$, $\Omega = \mathbb{R}^3$, and, for any given $(\bar{\theta}, \varepsilon)$, f is a multivariate Normal distribution over $\mathbb{R}^4$ with mean $(\bar{\theta}, \bar{\theta}, \bar{\theta}, \bar{\theta} + \varepsilon)$ and covariance matrix

$$\begin{pmatrix} \sigma_\varsigma^2 & \sigma_\varsigma^2 & 0 & 0 \\ \sigma_\varsigma^2 & \sigma_\varsigma^2 + \sigma_z^2 & 0 & 0 \\ 0 & 0 & \sigma_x^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

One can then conveniently index each $f \in F$ by the pair $(\bar{\theta}, \varepsilon) \in \mathbb{R}^2$ and recast the information structure as follows: Nature first draws $(\bar{\theta}, \varepsilon)$ from a bivariate Normal distribution with mean $(\mu_\theta, 0)$ and covariance matrix $\begin{pmatrix} \sigma_\theta^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix}$ and then uses the resulting distribution f to make independent draws of $(\theta_i, z_i, x_i, y_i)$, one for each i. The aforementioned bivariate Normal distribution of $(\bar{\theta}, \varepsilon)$ then plays the same role as $\mathcal{P}$ in the general formalization. Moreover, given any (measurable) strategy $k : \mathbb{R}^3 \to \mathbb{R}$, the mean and the dispersion of activity in the cross-section of the population can be expressed directly as a functions of $(\bar{\theta}, \varepsilon)$ rather than ϕ.

The special case where agents know their own shocks but not the aggregate shocks can then be nested by letting $\sigma_z = 0$ and $\sigma_\varsigma, \sigma_x, \sigma_y > 0$, while the special case in which there are no

attention to symmetric strategy profiles, this restriction is without loss of generality in our environment due to the symmetry and concavity of the payoff structure and the symmetry of the information structure. These properties guarantee that any two agents with the same information ω necessarily take the same action in equilibrium and are dictated the same action along the efficient allocation.

idiosyncratic shocks (as in the example in the previous section) can be nested by letting $\sigma_\varsigma = 0$. More generally, the aforementioned description of the information structure allows for private information regarding both idiosyncratic and aggregate shocks, about only one of the two, or about neither of these shocks. As it will become clear, the key for our results will be whether agents have private information regarding "aggregate shocks." To fix language, by "aggregate shocks" we mean the realized distribution h of these shocks in the cross section, while by "idiosyncratic shocks" we mean the realized fundamental θ_i that is specific to agent i.

Equilibrium, Efficiency, and Policy. A strategy is a mapping $k : \Omega \to \mathbb{R}$ that specifies an action for all possible signals $\omega \in \Omega$. Our equilibrium concept is standard Perfect Bayesian Equilibrium. Our efficiency concept, on the other hand, is the following: an efficient allocation (or "efficient use of information") is a strategy $k : \Omega \to \mathbb{R}$ that maximizes ex-ante utility.

Definition. *An efficient strategy is a mapping $k : \Omega \to \mathbb{R}$ that maximizes*

$$\mathbb{E}u = \int_F \int_{\Omega,\Theta} V(k(\omega), K(\phi), \sigma_k(\phi), \theta, \bar{\theta}) df(\omega, \theta) d\mathcal{P}(f)$$

subject to

$$K(\phi) = \int_\Omega k(\omega) d\phi(\omega) \; \forall \phi \in \Phi \quad and \quad \sigma_k(\phi)^2 = \int_\Omega [k(\omega) - K(\phi)]^2 d\phi(\omega) \; \forall \phi \in \Phi.$$

As anticipated in the Introduction, this efficiency concept is a constrained one only in the sense that the "planner" cannot transfer information from one agent to another: the action prescribed to an agent cannot depend on the private information of other agents. This concept thus bypasses the details of specific policy instruments and instead identifies directly the strategy that maximizes welfare under the restriction that information cannot be centralized. As it will be illustrated in the next section, the optimal policy can then be characterized by finding the tax schedule that implements the efficient use of information as an equilibrium.

Qualification. To avoid a number of distracting technical complications in the characterization of the equilibrium and efficient strategies, all proofs restrict Θ, Ω, and F to be finite sets. However, nothing substantial in the reasoning depends on this restriction. For example, the proofs can be extended to the case of multivariate Gaussian information structures, like the one described above, following similar steps as those in the Technical Appendix of Vives (2008). For this reason, the notation throughout the paper and the statement of the results do not take a stand on whether the aforementioned sets are finite or not.

Notation. To simplify, throughout the main text we use $\mathbb{E}_i[\cdot]$, k_i, K, and σ_k as short hands for $\mathbb{E}[\cdot|\omega_i]$, $k(\omega_i)$, $K(\phi)$, and $\sigma_k(\phi)$.

4   Decentralized use of information

In this section we characterize equilibrium and efficient allocations. Towards this goal, note that strategic uncertainty (uncertainty regarding one another's actions) can emerge only when agents have private information on ϕ; when instead ϕ is common knowledge, then for any given strategy the distribution of actions is also common knowledge. Moreover, because no signal ω can contain more information about the distribution of fundamentals in the population (h) than the entire distribution of signals (ϕ), common knowledge of ϕ implies common information regarding h. We conclude that common knowledge of ϕ is synonymous with both absence of strategic uncertainty and absence of private information regarding the underlying aggregate shocks.

To isolate the impact of private information regarding aggregate shocks (which is the case of interest for us), we first consider the structure of equilibrium and efficient allocations in the absence of such private information, that is, when ϕ is common knowledge.[5] This case nests two important benchmarks: (i) no private information on anything, as in the Ramsey literature; and (ii) private information only on idiosyncratic shocks, as in the Mirrlees literature.

Proposition 1. *Suppose ϕ is common knowledge, so that agents have no private information about aggregate shocks and face no strategic uncertainty. There exist coefficients $(\kappa_0, \kappa_1, \kappa_2)$ and $(\kappa_0^*, \kappa_1^*, \kappa_2^*)$ such that the equilibrium is given by*

$$k_i = \kappa \left(\mathbb{E}_i \theta_i, \mathbb{E}_i \bar{\theta} \right) \equiv \kappa_0 + \kappa_1 \mathbb{E}_i \theta_i + \kappa_2 \mathbb{E}_i \bar{\theta}, \tag{4}$$

while the efficient allocation is given by

$$k_i = \kappa^* \left(\mathbb{E}_i \theta_i, \mathbb{E}_i \bar{\theta} \right) \equiv \kappa_0^* + \kappa_1^* \mathbb{E}_i \theta_i + \kappa_2^* \mathbb{E}_i \bar{\theta}. \tag{5}$$

The coefficients $(\kappa_0, \kappa_1, \kappa_2)$ and $(\kappa_0^*, \kappa_1^*, \kappa_2^*)$ depend on the payoff structure V. Understanding their specific values is certainly important within the context of any particular application. For our purposes, however, what is important is only to register the following simple but general principle: as long as agents have either no private information at all or private information regarding only idiosyncratic shocks, then the equilibrium and efficient actions for an agent are merely functions of the agent's forecasts of his own fundamental θ_i and the aggregate fundamental $\bar{\theta}$.

Consider now the case that agents have private information regarding aggregate shocks. Because ϕ is not common knowledge, aggregate activity is not common knowledge either. As a result, equilibrium and efficient allocations depend, not only on the agents' forecasts of the fundamentals, but also on their forecasts of one another's actions and beliefs. This insight is formalized in the following proposition, which extends related results from Angeletos and Pavan (2007) to the more

---

[5]Clearly, whether ϕ is common knowledge or not is a restriction on $\mathcal{P}$.

general setting of this paper.[6]

Proposition 2. *Suppose ϕ is not common knowledge, so that agents have private information on aggregate shocks and face strategic uncertainty. There exist coefficients $\alpha < 1$ and $\alpha^* < 1$ such that the following are true:*

(i) The equilibrium satisfies

$$k_i = \kappa\left(\mathbb{E}_i\theta_i, \mathbb{E}_i\bar{\theta}\right) + \alpha \cdot \mathbb{E}_i\left[K - \kappa\left(\bar{\theta}, \bar{\theta}\right)\right], \tag{6}$$

while the efficient allocation satisfies

$$k_i = \kappa^*\left(\mathbb{E}_i\theta_i, \mathbb{E}_i\bar{\theta}\right) + \alpha^* \cdot \mathbb{E}_i\left[K - \kappa^*\left(\bar{\theta}, \bar{\theta}\right)\right]. \tag{7}$$

(ii) Let $\vartheta^1 \equiv \int \mathbb{E}[\theta|\omega]d\phi(\omega)$ and $\bar{\vartheta}^1 \equiv \int \mathbb{E}[\bar{\theta}|\omega]d\phi(\omega)$ denote, respectively, the average expectation of one's own fundamental and of the aggregate fundamental; for any $n \geq 2$, let $\vartheta^n \equiv \int \mathbb{E}[\vartheta^{n-1}|\omega]d\phi(\omega)$ and $\bar{\vartheta}^n \equiv \int \mathbb{E}[\bar{\vartheta}^{n-1}|\omega]d\phi(\omega)$ denote the corresponding n-th order average expectations; finally, let $\bar{\vartheta}^0 \equiv \bar{\theta}$. The equilibrium is given by

$$k_i = \kappa_0 + (\kappa_1 + \kappa_2)\mathbb{E}_i\left\{\sum_{n=0}^{\infty}(1-\alpha)\alpha^n\bar{\vartheta}^n\right\} + \kappa_1\mathbb{E}_i\left\{(\theta_i - \bar{\theta}) + \sum_{n=1}^{\infty}\alpha^n(\vartheta^n - \bar{\vartheta}^n)\right\}, \tag{8}$$

while the efficient allocation is given by

$$k_i = \kappa_0^* + (\kappa_1^* + \kappa_2^*)\mathbb{E}_i\left\{\sum_{n=0}^{\infty}(1-\alpha^*)\alpha^{*n}\bar{\vartheta}^n\right\} + \kappa_1^*\mathbb{E}_i\left\{(\theta_i - \bar{\theta}) + \sum_{n=1}^{\infty}\alpha^{*n}(\vartheta^n - \bar{\vartheta}^n)\right\}. \tag{9}$$

Part (i) highlights the dependence of equilibrium and efficient allocations on beliefs regarding aggregate activity (forecasts of the actions of others). To understand condition (6), recall that $\kappa\left(\mathbb{E}_i\theta_i, \mathbb{E}_i\bar{\theta}\right)$ is the action agent i would have taken in equilibrium had information about aggregate shocks been common. How much an agent deviates from this benchmark when information is dispersed depends on $\mathbb{E}_i\left[K - \kappa\left(\bar{\theta}, \bar{\theta}\right)\right]$, which is his forecast of the deviation of the other agents' average action from this benchmark, weighted by the coefficient α. In this sense, the coefficient α captures how much each *individual* cares about aligning his action with that of others, or equivalently the private value of forecasting one another's actions; it identifies the degree of strategic complementarity featured in equilibrium. Similarly, the coefficient α^* in condition (7) captures how much *society* would like agents to align their choices, or equivalently the social value of forecasting one another's actions; it identifies the degree of complementarity featured in efficient allocation.

Part (ii) then translates the result in terms of the hierarchy of beliefs (forecasts of the forecasts

---

[6]When Ω has the cardinality of the continuum, the efficient allocation is determined only for $\mathcal{P}$-almost all ω; we ignore this qualification in the rest of the paper.

of others). To better understand this result, consider the special case where the shocks are perfectly correlated ($\theta_i = \bar{\theta}$ for all i), in which case the last term in (8) and (9) disappears. If ϕ had been common knowledge, the equilibrium and efficient allocations would have been, respectively, $k_i = \kappa_0 + (\kappa_1 + \kappa_2)\mathbb{E}_i\bar{\theta}$ and $k_i = \kappa_0^* + (\kappa_1^* + \kappa_2^*)\mathbb{E}_i\bar{\theta}$. Now that ϕ is not common knowledge, the equilibrium and efficient allocations have the same structure, except that now $\mathbb{E}_i\bar{\theta}$ has been replaced by a weighted average of the entire hierarchy of beliefs about the underlying aggregate shocks. This is because an agent's first-order belief of the aggregate shock is no longer sufficient to forecast aggregate activity; the agent needs to forecast the forecasts of others. The terms α and α^* then determine, respectively, the sensitivity of equilibrium and efficient allocations to higher-order beliefs: the higher the degree of complementarity, the stronger the impact of higher-order beliefs relative to first-order beliefs.[7]

These results permit us to formalize a type of inefficiency that resembles the one alluded by Keynes in his beauty-contest metaphor for financial markets—an inefficiency that can emerge only when agents face strategic uncertainty and that is thus ruled out by either the Ramsey or the Mirrlees literature.[8]

Corollary 1. *When agents have private information about aggregate shocks, then, and only then, an inefficiency can emerge due to the discrepancy between the private and the social value of forecasting one another's actions and beliefs.*

To further appreciate this inefficiency, it is useful to spell out its implications for the reaction of the economy to the underlying fundamentals and noise. Suppose that productivities are perfectly correlated and information is Gaussian: $\theta_i = \bar{\theta} \sim \mathcal{N}\left(\mu, \sigma_\theta^2\right)$ and $\omega_i = (x_i^1, ..., x_i^n)$, where $x_i^s = \bar{\theta} + \xi_i^s$, $s \in \{1, ..., n\}$, and where the noises $\xi_i^s \sim \mathcal{N}\left(0, \sigma_s^2\right)$, $s \in \{1, ..., n\}$, are independent of one another and of $\bar{\theta}$, but can be correlated across agents (so that there is some common noise). Next, let $\hat{K}$ denote the projection of the equilibrium K on $\bar{\theta}$. Then, $Var(\hat{K})$ measures the component of volatility that is driven by fundamentals ("fundamental volatility"); $Var(K - \hat{K})$ measures the component that is driven by noise ("non-fundamental volatility"); and the ratio of the two defines the "signal-to-noise ratio" in aggregate activity. Because a stronger complementarity tilts the equilibrium towards the signals with relatively more correlated noise, the following is true.

---

[7]Note that, for all n, ϑ^n and $\bar{\vartheta}^n$ are measurable in ϕ. When ϕ is common knowledge, then $\vartheta^n = \mathbb{E}[\theta|\phi] = \mathbb{E}[\bar{\theta}|\phi] = \bar{\vartheta}^n = \mathbb{E}_i[\bar{\theta}]$ for all n and all i; this is simply because all information regarding aggregate shocks is common. This explains why, when ϕ is common knowledge, (8) and (9) reduce, respectively, to (4) and (5).

[8]To rule out degenerate cases that render the degree of complementarity irrelevant for behavior even when ϕ is not common knowledge, we henceforth assume that the information structure $(\Omega, F, \mathcal{P})$ is "regular" in the following sense: for any two payoff structures V and V' that lead to the same κ but different α, whenever ϕ is not common knowledge, the equilibrium of the economy $(V, \Omega, F, \mathcal{P})$ is different than that of the economy $(V', \Omega, F, \mathcal{P})$ for a non-zero-measure subset of Ω. When shocks θ_i are perfectly correlated, a sufficient condition for this is that there exists a subset $\bar{\Omega}$ of Ω, with non-zero probability measure under $\mathcal{P}$, such that $\mathbb{E}[\bar{\vartheta}^n|\omega] > \mathbb{E}[\bar{\vartheta}^n|\omega']$ for all n and for all $\omega, \omega' \in \bar{\Omega}$. For Gaussian information structures like the one described below, this condition is trivially satisfied—indeed with $\bar{\Omega} = \Omega$—unless the prior is completely uninformative and there is no public signal.

Proposition 3. *Consider the Gaussian information structure described above. The equilibrium signal-to-noise ratio in aggregate activity is inefficiently high if and only if $\alpha > \alpha^*$.*

We conclude that the condition $\alpha > \alpha^*$ is synonymous to any of the following: (i) excessive concern for forecasting the forecasts of others; (ii) overreaction to sources of information with highly correlated noise; and (iii) excessive non-fundamental volatility. Of course, how α and α^* compare, and hence whether the aforementioned inefficiency is present, depends on the details of the application. Therefore, one cannot fully appreciate this inefficiency without a specific context. (See also the discussion in Section 5.4.) However, for the purposes of this paper, we can bypass the details of the origins, and the precise interpretation, of this inefficiency and instead focus on its potential policy implications.

5 Policy

We now turn to policy. We first study how different policies affect the incentives agents face when they decide how to use their different sources of information; this part generalizes the insights illustrated in Section 2. We then identify the policy (or policies) that implement the efficient use of information as an equilibrium; this part establishes that the contingency on realized aggregate activity is essential for optimality *only* when agents have private information regarding aggregate shocks, further highlighting the contribution of our paper vis-a-vis the pertinent literature.

5.1 The equilibrium impact of the policy contingencies

Consider the following class of (possibly non-linear) tax-schedules that are contingent on ex post information about realized aggregate activity and fundamentals:

$$\tau_i = T\left(k_i, K, \sigma_k, \bar{\theta}\right),$$

where $T : \mathbb{R}^2 \times \mathbb{R}_+ \times \Theta \to \mathbb{R}$. Without loss of optimality (as it will be clear soon), we restrict T to be such that the policy-induced payoff function $\tilde{V} \equiv V - T$ remains in $\mathcal{V}$; we further impose $T\left(K, K, 0, \bar{\theta}\right) = G$ for all $(K, \bar{\theta})$ and $T_{kk} + T_{\sigma\sigma} = 0$, which is necessary and sufficient for the policy to be budget-balanced for any possible strategy. Denoting the class of policies that satisfy these properties by $\mathcal{T}$, we have the following result. (To simplify the formulas, we normalize payoffs so that $V_{kk} = -1$; see the Appendix for the more general case.)

Proposition 4. *Given any tax scheme $T \in \mathcal{T}$, let*

$$\tilde{\alpha} \equiv \frac{\alpha - T_{kK}}{1 + T_{kk}}, \quad \tilde{\kappa}_0 \equiv \frac{(1-\alpha)\kappa_0 - T_k(0,0,0)}{1 - \alpha + T_{kk} + T_{kK}}, \quad \tilde{\kappa}_1 \equiv \frac{1}{1 + T_{kk}}\kappa_1, \quad \tilde{\kappa}_2 \equiv \frac{(1-\alpha)\left(\kappa_1 + \kappa_2\right) - T_{k\bar{\theta}}}{1 - \alpha + T_{kk} + T_{kK}} - \tilde{\kappa}_1.$$

When ϕ is common knowledge, the equilibrium is given by

$$k_i = \tilde{\kappa}\left(\mathbb{E}_i\theta, \mathbb{E}_i\bar{\theta}\right) \equiv \tilde{\kappa}_0 + \tilde{\kappa}_1\mathbb{E}_i\theta + \tilde{\kappa}_2\mathbb{E}_i\bar{\theta}.$$

When instead ϕ is not common knowledge, the equilibrium is given by

$$k_i = \tilde{\kappa}\left(\mathbb{E}_i\theta, \mathbb{E}_i\bar{\theta}\right) + \tilde{\alpha} \cdot \mathbb{E}_i\left[K - \tilde{\kappa}\left(\bar{\theta}, \bar{\theta}\right)\right]. \tag{10}$$

There are four instruments that permit the government to influence equilibrium allocations: the level of taxation (as parametrized by $T_k(0,0,0)$); the non-linearity of the tax system (T_{kk}); the contingency of marginal taxes on realized aggregate productivity ($T_{k\bar{\theta}}$); and their contingency on realized aggregate activity (T_{kK}). While all these instruments matter for equilibrium outcomes, each one has a distinctive role. T_{kk} is the only instrument that permits the government to control $\tilde{\kappa}_1$, the sensitivity of the agents' actions to their information about their *own* productivity shocks. For given T_{kk}, the only instrument that permits the government to control $\tilde{\alpha}$, the degree of complementarity is T_{kK}, the contingency on aggregate activity. For given T_{kk} and T_{kK}, the instrument that permits the government to control $\tilde{\kappa}_2$, the sensitivity of individual actions to variations in aggregate productivity, is $T_{k\bar{\theta}}$. Finally, $T_k(0,0,0)$ controls merely the average level of activity.

These results help generalize the insights delivered in the investment example of Section 2. Note that a higher $T_{k\bar{\theta}}$ reduces $\tilde{\kappa}_2$ but does not affect $\tilde{\alpha}$, whereas a higher T_{kK} reduces both $\tilde{\kappa}_2$ and $\tilde{\alpha}$. This means that the contingency on the realized fundamentals has a symmetric effect across all sources of information, whereas the contingency on realized aggregate activity has an asymmetric effect: a higher T_{kK} penalizes the agents relative more when they react to sources of information that are relatively better predictors of aggregate activity (or, equivalently, when they react to variation in higher-order beliefs). In other words, the contingency of the tax schedule on the realized aggregate activity replicates the same incentive effects as a direct preferential tax on certain sources of information.

These results hold for arbitrary information structures, but they are most clearly illustrated in the case of a Gaussian information structure.

Proposition 5. *Consider the Gaussian information structure described in Section 4. Other things equal, a higher $T_{k\bar{\theta}}$ reduces fundamental and non-fundamental volatility proportionally, whereas a higher T_{kK} has a disproportional effect on non-fundamental volatility. The signal-to-noise ratio in aggregate activity is thus independent of $T_{k\bar{\theta}}$ but increases with T_{kK}.*

These results suggest that the contingency on aggregate activity is the key to correcting the particular type of inefficiency that we documented in Section 4. We verify this intuition next.

5.2 Implementation of the efficient decentralized use of information

We now turn to the existence and characterization of a policy $T^* \in \mathcal{T}$ that implements the efficient allocation as an equilibrium. Whenever such a policy exist, the very definition of the efficient allocation guarantees that there is no other policy that can improve upon T^*. This is true even for policies that violate budget balance, or even if one allows the agents to send arbitrary messages to the planner and the planner to make transfers contingent on these messages; what is essential is only that the planner does not send any information to the agents before they make their choices.

Proposition 6. *(i) When ϕ is common knowledge, the efficient allocation can always be implemented with a policy that is contingent only on the fundamentals: it is without loss of optimality to set $T_{kK} = 0$. (ii) When instead ϕ is not common knowledge, the contingency on aggregate activity becomes essential for efficiency: the optimal T_{kK} is uniquely determined for all economies and it is non-zero for all but a zero-measure set of payoff functions.*

The proof of this result follows from Proposition 4. First, note that there exists a unique T_{kk} such that $\tilde{\kappa}_1 = \kappa_1^*$. Given this T_{kk}, there exists a unique T_{kK} such that $\tilde{\alpha} = \alpha^*$. But then there also exist a unique $T_{k\bar{\theta}}$ such that $\tilde{\kappa}_2 = \kappa_2^*$ and a unique $T_k(0,0,0)$ such that $\tilde{\kappa}_0 = \kappa_0^*$. All other parameters of the policy are then pinned down by budget balance. Next, note that, when ϕ is common knowledge, the policy implements the efficient allocation if and only if it induces $\tilde{\kappa} = \kappa^*$; the degree of complementarity $\tilde{\alpha}$ is irrelevant. In this case there is one degree of indeterminacy in the optimal policy in the sense that there are multiple combinations of T_{kK} and $T_{k\bar{\theta}}$ that induce $\tilde{\kappa} = \kappa^*$. It is thus without any loss to set $T_{kK} = 0$. In fact, this is true for *any* implementable allocation, not just the efficient one: when ϕ is common knowledge, any allocation that can be implemented with a policy that has $T_{kK} \neq 0$ can also be implemented with a policy that has $T_{kK} = 0$.

When, instead, ϕ is not common knowledge, the policy implements the efficient allocation if and only if it induces both $\tilde{\kappa} = \kappa^*$ and $\tilde{\alpha} = \alpha^*$.[9] In this case the optimal policy is uniquely determined and, for all but a zero (Lebesgue) measure set of payoff functions, it features $T_{kK} \neq 0$. Moreover, the optimal T_{kK} tends to increase with the gap between α and α^*, which means that the optimal T_{kK} is higher the more severe the inefficiency in the signal-to-noise ratio in equilibrium activity.

We conclude that, whereas the efficient allocation can always be implemented with a tax schedule that is contingent merely on the realized aggregate fundamentals when information regarding aggregate shocks is common (as in the Ramsey and Mirrlees literatures), the contingency on realized aggregate activity becomes necessary for optimality once such information is dispersed. This is because, when information regarding the aggregate fundamentals is common, aggregate activity can be a function of only this information, and hence a contingency of the policy on aggregate activity has exactly the same incentives effects as a contingency on the aggregate fundamentals. In

---

[9]That efficiency obtains only if $\tilde{\alpha} = \alpha^*$ is true under the regularity condition introduced in footnote 8.

contrast, when information regarding the aggregate fundamentals is dispersed, the contingency on aggregate activity has a differential effect than the contingency on aggregate fundamentals, for it is this contingency, and only this one, that can correct any excessive sensitivity of the equilibrium to noise *relative* to the fundamentals.

5.3   Implementation with measurement error

The preceding analysis has assumed that the government can perfectly observe the agents' activity and aggregate productivity at the time taxes are collected. We now consider a variant that introduces measurement error; apart from being more realistic, this will prove useful in the dynamic extension of Section 6, where activity and fundamentals are observed with noise in each period.

We consider both additive and multiplicative measurement error. In the additive case, the government's measurement of agent i's activity is $\tilde{k}_i = k_i + \eta + \nu_i$, while it's measurement of the aggregate fundamental is $\tilde{\theta} = \bar{\theta} + \varsigma$, where η and ς are common noise, while ν_i is idiosyncratic noise. In the multiplicative case, the respective signals are $\tilde{k}_i = k_i(1 + \eta + \nu_i)$ and $\tilde{\theta} = \bar{\theta}(1 + \varsigma)$. In either case, we let $\tilde{K}$ and $\tilde{\sigma}_k$ denote, respectively, the cross-sectional average and dispersion of $\tilde{k}$. We then consider tax schedules of the form $\tau_i = T(\tilde{k}_i, \tilde{K}, \tilde{\sigma}_k, \tilde{\theta})$, where the function T is assumed to satisfy the same properties as before.

Proposition 7. *Propositions 5 and 6 are robust to measurement error.*

To understand this result, note that $\mathbb{E}_i T(\tilde{k}_i, \tilde{K}, \tilde{\sigma}_k, \tilde{\theta}) = \mathbb{E}_i T(k_i, K, \sigma_k, \bar{\theta}) + SOT$, where SOT are second-order terms that capture the impact of the risk introduced by measurement error. (The first-order terms vanish because these errors have zero means.) When the measurement error is additive, SOT is independent of k_i, which means that such measurement error does not interfere at all with the incentives provided by the tax system. When, instead, the measurement error is multiplicative, it does impact incentives. However, by appropriately adjusting the policy, the government can fully undo the incentive effects of the noise. The details of the optimal policy then depend on the measurement error, but the efficient allocation remains implementable and the contingency on K remains essential only when agents have private information on aggregate shocks.

5.4   Discussion/applications

The analysis has established three key results that provide general guidance about the role of policy in environments with dispersed information on aggregate shocks:

- The contingency of the policy on aggregate fundamentals has a symmetric effect on the use of all sources of information, while the contingency on aggregate activity has an asymmetric effect, penalizing relatively more the use of those sources that have highly correlated noise.

- The dispersion of information regarding aggregate shocks can introduce an inefficiency that is absent in either the Ramsey or the Mirrlees literature; this inefficiency manifests itself as excessive non-fundamental volatility.

- The contingency of the policy on aggregate activity is essential for optimality only in the presence of the aforementioned inefficiency.

Within our framework, the aforementioned inefficiency can be formalized by the gap between two coefficients. The first one, α, identifies the degree of complementarity featured in equilibrium; the second, α^*, identifies the degree of complementarity featured in efficient allocations. In the Appendix we further show that one can interpret a lower α^* as a higher social aversion to non-fundamental volatility. A more complete characterization and interpretation of these coefficients requires restricting attention to a specific application: one needs to look at the primitive preferences, technologies, and market interactions that are hidden behind our reduced-form payoff V.

For example, consider the competitive economy discussed in Section 3. In this economy, production choices are strategic substitutes ($\alpha < 0$). this is because a higher aggregate production of good 2 reduces the equilibrium price of that good, which in turn reduces the individual incentive to produce. At the same time, the absence of monopolistic power and of any other friction than the dispersion of information guarantees that there is no inefficiency in the use of information ($\alpha^* = \alpha$), thereby leaving no room for policy intervention. In contrast, in the beauty-contest model of Morris and Shin (2002), agents engage in a game that induces a positive complementarity in equilibrium ($\alpha > 0$). Because this game is zero-sum, this complementarity is not warranted from a social perspective ($\alpha^* = 0$). It then follows from our results that, for that model, the optimal policy features a positive contingency on aggregate activity ($T_{kK} > 0$).

Next, consider Angeletos, Lorenzoni and Pavan (2007). That paper studies the two-way interaction between the real and the financial sector of a micro-founded economy in which entrepreneurs and financial traders have dispersed information regarding the profitability of a new technology. Because high aggregate investment is "good news" for profitability, asset prices increase with aggregate investment. Because firms' incentives to invest in turn increase with asset prices, an endogenous complementarity emerges in the investment decisions of the entrepreneurs. In effect, the entrepreneurs play of reduced-form game in which $\alpha > 0$.[10] Angeletos, Lorenzoni and Pavan (2007) proceed to show that this endogenous complementarity is a source of inefficiency; in effect, they show that $\alpha > \alpha^*$ in their economy. They then adapt the results of this paper to identify the policy that can correct this inefficiency.

---

[10]The payoff structure of that game is endogenous; importantly, because the equilibrium asset prices depend on the information structure, the complementarity also depends on the information structure. Nevertheless, this property does not pose any difficulty for our policy exercise. Indeed, none of our results is affected if we let the reduced-form payoff function V—and hence also the coefficients α and α^*—depend on the underlying information structure.

Finally, consider Angeletos and La'O (2008), Hellwig (2005), and Lorenzoni (2008). These papers consider dispersed-information variants of the class of Dixit-Stiglitz monopolistic economies that is now canonical for business-cycle theory. These models cannot be directly nested in our linear-quadratic framework. Nevertheless, the equilibrium and efficient allocations of this class of models has a log-linear structure quite similar to the one considered here. In fact, an analogue of Proposition 2 continues to hold once one makes two adjustments: first, k_i must now be interpreted as the *logarithm* of the relevant production or pricing decision; second, the coefficients κ_0 and κ_0^* now depend on the level of uncertainty, due to risk aversion. The coefficients α and α^* are then determined by the elasticity of substitution across different goods, the curvature of the production function, the Frisch elasticity of labor supply, and the coefficient of relative risk aversion. Angeletos and La'O (2008) establish that $\alpha = \alpha^*$ as long as nominal prices are flexible or monetary policy replicates the flexible-price allocations. It follows that there is no inefficiency in the equilibrium use of information, and hence no room for policy intervention, as long as information is exogenous.

To recap, the micro-foundations of any particular application are essential for understanding the determinants of the equilibrium and efficient allocations and hence for fully appreciating what determines the need for policy intervention in the first place. However, the general principles derived in this paper do not hinge on the specific micro-foundations of the application under examination.

We conclude with a qualification of our result regarding the necessity of making the policy contingent on aggregate activity: if ϕ, or any sufficient statistic of it became common knowledge at the time taxes are collected, the efficient allocation could also be implemented with a tax schedule that is contingent on ϕ, or the sufficient statistic, rather than K. This observation is obvious but important to keep in mind when studying applications. In applied work, it is often convenient to assume a very stark Gaussian information structure, one with a *purely private* signal (whose noise is uncorrelated across agents) and a *purely public* signal (which is common knowledge). In this case, aggregate activity is only a function of the true fundamental θ and the public signal. It then follows that there is no loss of optimality in making the policy contingent on θ and the public signal, instead of making it contingent on θ and K. However, note that once one perturbs the information structure so as to introduce some unobserved common noise in the agents' private information, then one can no more replicate the contingency on K with a contingency on the public signal. Moreover, it is unclear how one could measure the public signal in practice; this signal is just a convenient modeling device meant to capture a variety of common (or correlated) sources of information that may be available to the agents but not necessarily to the government. In this sense, the implementations we have considered in this paper are both more robust to the details of the information structure and easier to use in practice.

6 A dynamic extension with endogenous learning

In this section we consider a dynamic variant of the baseline framework. This serves two goals. First, it brings the framework closer to macro applications. Second, and most importantly for our purposes, it lets agents observe signals of the aggregate activity in the economy. These signals are proxies for macro data, financial prices, and other channels of social learning. This extension thus permit us to study how the policies we have identified can improve the efficiency of social learning.

6.1 Set up

Time is discrete, indexed by $t \in \{0, 1, ..., N\}$, for arbitrary N. In each period t, each agent i chooses a level of consumption, $c_{i,t}$, a position in a riskless discount bond, $b_{i,t}$, and some action, $k_{i,t}$; the latter, which we can interpret as effort or investment in a risky technology, is the key economic decision. Let K_t and σ_t denote the mean and the dispersion of activity in period t, and $\tilde{\theta}_t$ the productivity in period t (which, for simplicity, is henceforth assumed to be identical across agents). The latter is given by $\tilde{\theta}_t = \theta + a_t$, where $\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$ is a permanent component and $a_t \sim \mathcal{N}(0, \sigma_a^2)$ is a transitory component (i.i.d. across time).

At the beginning of each period t, agents publicly observe $\tilde{\theta}_t$, but cannot tell apart the permanent and the transitory component. In addition, agents observe noisy public signals of past activity, namely $\tilde{K}_{t-1} = K_{t-1} + \eta_t$ and $\tilde{\sigma}_{t-1} = \sigma_{t-1} + \sigma_{v,t}$, where $\eta_t \sim \mathcal{N}(0, \sigma_{\eta,t}^2)$ is a common measurement error and $\sigma_{v,t}$ is the dispersion of idiosyncratic measurement errors; these measurement errors are independent across time, and independent of any other random variable.[11] Finally, agents receive an exogenous private signal $x_{i,t} = \theta + \xi_{i,t}$ and an exogenous common signal $y_t = \theta + \varepsilon_t$, where the noises $\xi_{i,t} \sim \mathcal{N}(0, \sigma_{x,t}^2)$ and $\varepsilon_t \sim \mathcal{N}(0, \sigma_{y,t}^2)$ are independent of any other random variable.

The entire information of agent i in period t is then summarized in a vector $\omega_{i,t}$, which can be recursively constructed by $\omega_{i,t} = (x_{i,t}, \tilde{\theta}_t, \tilde{K}_{t-1}, \tilde{\sigma}_{t-1}; \omega_{i,t-1})$. We let Ω_t denote the set of all possible realizations of $\omega_{i,t}$, ϕ_t the cross-sectional distribution of $\omega_{i,t}$, and Φ_t the set of all possible such distributions. For any strategy $k_t : \Omega_t \to \mathbb{R}$ followed in period t, aggregate activity is given by $K_t(\phi_t) = \int_{\Omega_t} k_t(\omega_t) d\phi_t(\omega_t)$ and its dispersion by $\sigma_t(\phi_t) = \int_{\Omega_t} (k_t(\omega_t) - K_t(\phi))^2 d\phi_t(\omega_t)$. To economize on notation, we once again suppress the dependence of $k_{i,t}$ on $\omega_{i,t}$ and that of K_t and σ_t on ϕ_t, and let $\mathbb{E}_{i,t}$ denote the expectation conditional on $\omega_{i,t}$. Finally, we note that social learning can obtain only because of the dispersion of information regarding aggregate shocks: if ϕ_t had been common knowledge, nothing could be learned from signals of past activity.

---

[11]The analysis easily extends to the case where agents observe private signals of aggregate activity in addition to the aforementioned public ones; we drop the private signals only for expositional simplicity.

Turning to payoffs, the intertemporal preferences of agent i are given by

$$\mathcal{U}_i = \sum_{t=1}^{N+1} \beta^{t-1} U\left(c_{i,t}, k_{i,t}\right).$$

where $U(\cdot)$ is a real-valued function and where $\beta \in (0,1)$ is the discount factor. (We allow U to depend on $k_{i,t}$ to capture cases where the latter represents effort.) The agent's period$-t$ budget, on the other hand, is given by

$$c_{i,t} + G\left(k_{i,t}\right) + q_t b_{i,t} = F(k_{i,t-1}, \tilde{K}_{t-1}, \tilde{\sigma}_{t-1}, \tilde{\theta}_t) + b_{i,t-1} - \tau_{i,t},$$

where q_t denotes the period-t price of discount bonds (the reciprocal of the risk-free rate) and $\tau_{i,t}$ denotes the period-t taxes the agent pays to the government.[12] The function G can be interpreted as the cost of period-t investment, while the function F can be interpreted as the income received in period t. The latter is allowed to depend on others activity, capturing the same kind of external payoff effects as in the baseline framework. This dependence is "noised-up" through the measurement errors in $\tilde{K}_{t-1}$ and $\tilde{\sigma}_{t-1}$ only to ensure that the observation of own income does not perfectly reveal past activity and thereby θ.[13]

This framework is quite flexible. For example, a stylized version of the neoclassical growth model with no labor and with convex investment costs is nested by letting $U\left(c, k\right) = c$, $F\left(k, K, \sigma, \theta\right) = \theta k$, and $G\left(k\right) = k + \chi k^2$, for some constant $\chi > 0$. In this particular case, informational externalities would be the only source of inefficiency. Adding external payoff effects through F may then stylize a variety of market or non-market interactions, as in the baseline framework.[14] The informational role of any prices or other forms of social learning that may be associated with such interactions are then mimicked by the information role of $\tilde{K}_t$.

To keep the analysis tractable, we assume that U is linear in consumption: $U(c, k) = c - H(k)$, for some function H. This restriction rules out redistributive concerns. It further ensures that the bond market clears if and only if $q_t = \beta$ (in which case the demand for the risk-free bond is indeterminate) and that the life-time utility of agent i (in the absence of taxes) reduces to

$$\mathcal{U}_i = \sum_{t=1}^{N} \beta^{t-1} V\left(k_{i,t}, \tilde{K}_t, \tilde{\sigma}_t, \tilde{\theta}_{t+1}\right),$$

---

[12]For period $t = N + 1$, we impose that $k_{i,N+1} = b_{i,N+1} = 0$.

[13]We could remove these measurement errors and still guarantee that aggregate activity does not perfectly reveal θ by introducing private signals with correlated rather than purely idiosyncratic noises. As anticipated earlier, the contingency of policy on aggregate activity is the key instrument for controlling the signal-to-noise ratio in aggregate activity; the precise source of noise need to be crucial.

[14]Clearly, we could also allow the functions H and G to depend on (K, σ, θ), so as to capture externalities in leisure, pecuniary externalities in the cost of investment, and so on.

where the function V is now given by

$$V\left(k, K, \sigma, \theta\right) \equiv -\left[G\left(k\right) + H(k)\right] + \beta F\left(k, K, \sigma, \theta\right).$$

We also assume that V is a quadratic function, satisfying the same properties with respect to (k, K, σ, θ) as the function V in the baseline framework. Along with the assumption that all exogenous random variables are Gaussian, this will guarantee that all endogenous signals are also Gaussian in equilibrium (or, more generally, for any linear strategy), which is essential for maintaining the analysis tractable.[15]

6.2 Equilibrium

The essential difference between the economy of this section and the one of the baseline framework is the endogeneity of information: the strategy agents follow in period t determines how much information about θ is contained in $\tilde{K}_t$, which in turn affects behavior and welfare in periods $t + 1$ on. However, this informational externality does not alter private incentives. The following then is a direct extension of the equilibrium results of the benchmark framework.

Proposition 8. *There exist a linear function $\kappa : \mathbb{R} \to \mathbb{R}$ and a coefficient α such that the equilibrium satisfies*

$$k_{i,t} = \kappa\left(\mathbb{E}_{i,t}\theta\right) + \alpha \cdot \mathbb{E}_{i,t}\left[K_t - \kappa\left(\theta\right)\right] \tag{11}$$

This result does not require the information structure to be Gaussian. However, once we restrict θ and the exogenous noises to be Gaussian, this result ensures that the information contained in the signals of past activity is also Gaussian. All the information—exogenous and endogenous—that is available in any given period can then be summarized in two sufficient statistics, one for the private and the other for the public signals; the dynamics of these two statistics admit a simple recursive structure; and the equilibrium strategy reduces to an affine combination of the two.

Proposition 9. *The equilibrium strategy is given by*

$$k_{i,t} = \kappa\left(\gamma_t X_{i,t} + \left(1 - \gamma_t\right) Y_t\right),$$

---

[15]The framework we have introduced in this section is essentially a hybrid of our baseline framework, which allowed for payoff interactions but abstracted from social learning, and the model in Vives (1997), which allowed for social learning but abstracted from payoff interactions. In particular, we can nest Vives's framework, and its close cousin in Amador and Weill (2007), by setting $V(k, K, \sigma, \theta) = -(\theta - k)^2$. Combining social learning with payoff interactions is *not* merely for the sake of generality; it can be crucial for the normative properties of the class of environments that we are interested in. For example, the key result in Amador and Weill (2007, 2008) is that public information can reduce welfare by reducing the speed of social learning; this result relies on having a positive social value for social learning in the first place, which need not be the case once one allows for payoff interactions (see also Chapter 3 in Vives, 2008, for the same point). In contrast, as it will become clear, our policy results do not hinge on the details of the underlying payoff structure.

with

$$\gamma_t = \frac{(1-\alpha)\,\pi_t^x}{(1-\alpha)\,\pi_t^x + \pi_t^y}. \tag{12}$$

The variables $X_{i,t}$ and Y_t are sufficient statistics for all the private and public information about θ that is available to agent i in period t, while π_t^x and π_t^y are their respective precisions. The sufficient statistics are given recursively by

$$X_{i,t} = \frac{\pi_{t-1}^x}{\pi_t^x} X_{i,t-1} + \frac{\sigma_{x,t}^{-2}}{\pi_t^x} x_{i,t} \quad and \quad Y_t = \frac{\pi_{t-1}^y}{\pi_t^y} Y_{t-1} + \frac{\sigma_{y,t}^{-2}}{\pi_t^y} y_t + \frac{\sigma_a^{-2}}{\pi_t^y} \tilde{\theta}_t + \frac{(\kappa')^2 \gamma_{t-1}^2 \sigma_{\eta,t}^{-2}}{\pi_t^y} \tilde{y}_t \tag{13}$$

where

$$\tilde{y}_t \equiv \frac{\tilde{K}_{t-1} - \kappa((1-\gamma_{t-1})\,Y_{t-1})}{\kappa' \gamma_{t-1}} \tag{14}$$

is a linear transformation of the signal of past activity and $\kappa' \in \mathbb{R}$ is the derivative of κ. Similarly, the precisions π_t^x and π_t^y are given recursively by

$$\pi_t^x = \pi_{t-1}^x + \sigma_{x,t}^{-2} \quad and \quad \pi_t^y = \pi_{t-1}^y + \sigma_{y,t}^{-2} + \sigma_{a,t}^{-2} + (\kappa')^2 \gamma_{t-1}^2 \sigma_{\eta,t}^{-2}. \tag{15}$$

Finally, the initial conditions are $X_{i,0} = 0$, $Y_{i,0} = \mu$, $\gamma_0 = 0$, $\pi_0^x = 0$ and $\pi_0^y = \sigma_\theta^{-2}$.

The logic behind condition (12) is simple. Note that γ_t represents the relative sensitivity of the equilibrium strategy to the sufficient statistic $X_{i,t}$ of the private information of an agent. For given degree of complementarity α, this sensitivity increases with the precision of private information and decreases with the precision of public information. At the same, for given precisions, a higher α tilts the equilibrium strategy away from private information and towards public information, as agents find it optimal to better align their choices. Conditions (13)-(15), on the other hand, describe the dynamics of information: agents update the sufficient statistics inherited from period $t-1$ with the new exogenous and endogenous signals observed in period t. Naturally, how much these signals are weighted depends on their respective precisions; and for the same reasons as in the benchmark models, it also depends on the degree of complementarity, α, featured in equilibrium.

The key novel property then to notice is that the precision of information available in one period depends on the strategy followed in previous periods. In particular, for all t, the precision of the endogenous signal $\tilde{y}_t$ and thereby the precision π_t^y of the sufficient statistic Y_t is increasing in γ_{t-1}. This is because the informative content of the signals of aggregate activity is higher the more sensitive the strategies of the agents to their private information. This is an important informational externality that the equilibrium fails to internalize in the absence of policy intervention.

6.3 Efficiency

We now seek to identify the strategy that maximizes ex-ante utility taking into account the afore-mentioned informational externality. Unlike the case with exogenous information considered in the benchmark model, here we have to restrict attention to strategies that are linear in the history of available private signals;[16] without this restriction, the endogenous signals are no longer Gaussian and the analysis becomes intractable. We can then characterized the efficient allocation as follows.

Proposition 10. *There exists a linear function $\kappa^* : \mathbb{R} \to \mathbb{R}$ and a scalar α^* such that the efficient linear allocation is given by*

$$k_{i,t} = \kappa^* \left(\gamma_t^{**} X_{i,t} + (1 - \gamma_t^{**}) Y_t \right),$$

where

$$\gamma_t^{**} = \frac{(1 - \alpha^*) \pi_t^x}{(1 - \alpha^*) \pi_t^x + \pi_t^y - \beta (1 - \alpha^*) \left(1 - \gamma_{t+1}^{**}\right)^2 \pi_t^x \pi_t^y \left(\pi_{t+1}^y\right)^{-2} (\kappa^{*\prime})^2 \sigma_{\eta,t+1}^{-2}} > 0$$

for all $t < N$, while $\gamma_N^{**} = (1 - \alpha^*) \pi_N^x / \left[(1 - \alpha^*) \pi_N^x + \pi_N^y \right]$. $X_{i,t}$ and Y_t are sufficient statistics for all the private and public information available to agent i in period t, while π_t^x and π_t^y are their respective precisions; these sufficient statistics are obtained recursively using (13)-(15), replacing γ_t with γ_t^{**} and κ with κ^*.

To appreciate this result, consider, as a reference point, what the efficient allocation would have been in period t if all information in period $t + 1$ had been exogenous. This case is nested here by letting $\sigma_{\eta,t+1} = \infty$ (infinite measurement error in the signal of aggregate activity observed during period $t + 1$). The optimal weight on $X_{i,t}$ is then given by

$$\gamma_t^* = \frac{(1 - \alpha^*) \pi_t^x}{(1 - \alpha^*) \pi_t^x + \pi_t^y}.$$

This is the same as with equilibrium, except that α has being replaced with α^*. Any difference between the equilibrium and the efficient use of information could then originate only in payoff effects, as in the benchmark model.[17]

Consider now the implications of informational externalities ($\sigma_{\eta,t+1} < \infty$). The optimal weight

---

[16]By this we mean the following: let h_t denote the public history in period t, which is constructed recursively by $h_t = (\tilde{\theta}_t, \tilde{K}_{t-1}, \tilde{\sigma}_{t-1}; h_{t-1})$; we impose that, for all t and all all $\omega_{i,t}$, $k_t (\omega_{i,t}) = P_t (h_t) + \sum_{\tau=1}^{t} Q_{t,\tau} x_{i,\tau}$, for some deterministic function P_t and some deterministic coefficients $\{Q_{t,\tau}\}_{\tau=1}^{t}$. Note that we do not impose linearity in the public signals, nor linearity in the sufficient statistics; the property stated in Proposition 10 that the efficient strategy can be expressed as a linear function of the sufficient statistics $X_{i,t}$ and Y_t is a result, not an assumption.

[17]Another way to see this is the following: if information had been exogenous, the efficient allocation would satisfy $k_{i,t} = \kappa^* (\mathbb{E}_{i,t}\theta) + \alpha^* \cdot \mathbb{E}_{i,t} [K_t - \kappa^* (\theta)]$, where both the function κ and the coefficient α^* are determined by the payoff function V, as in the benchmark model.

on $X_{i,t}$ now satisfies

$$\gamma_t^{**} = \frac{(1-\alpha^*)\,\pi_t^x}{(1-\alpha^*)\,\pi_t^x + \pi_t^y - \beta\Delta_{t+1}}, \tag{16}$$

where $\Delta_{t+1} \equiv (1-\alpha^*)\left(1-\gamma_{t+1}^{**}\right)^2 \pi_t^x \pi_t^y \left(\pi_{t+1}^y\right)^{-2} (\kappa^{*\prime})^2 \sigma_{\eta,t+1}^{-2} > 0$.[18] It follows that $\gamma_t^{**} > \gamma_t^*$. That is, relative to the case with exogenous information, the efficient use of information is now titled away from the public signals (here summarized in Y_t) and towards the private signals (here summarized in $X_{i,t}$). This is intuitive. Increasing the sensitivity of actions to $X_{i,t}$ increases the precision of the information contained in the signals of past aggregate activity in period $t+1$. Along the efficient allocation, this necessarily increases the present-value welfare from period $t+1$ onward.[19] Of course, doing so comes at a welfare cost during period t: the weight that maximizes the period$-t$ flow welfare is simply γ_t^*. However, the envelope theorem guarantees that, starting from this reference point, the marginal cost of increasing γ_t in terms of period$-t$ welfare is zero, while the marginal benefit in terms of welfare from period $t+1$ and on is strictly positive. Together with the concavity of the planner's problem, this implies that the new optimum is achieved at a point γ_t^{**} strictly higher than γ_t^*. The "wedge" $\beta\Delta_{t+1}$ in (16) captures precisely the (discounted) social benefit of doing so, that is, the informational externality.

Here, it is important to note the following. Although the efficient allocation is necessarily more sensitive to private information than what it would have been if learning was absent, this does not mean that it is also more sensitive than the equilibrium allocation. In other words, the presence of informational externalities does not necessarily mean that the equilibrium features too little learning—whether this is the case depends on the payoff structure. Indeed, in environments where there are no payoff interdependencies (e.g., Banerjee, 1992; Vives, 1997; Amador and Weill, 2007), the equilibrium sensitivity to private information is efficient when information is exogenous and, by implication, it is inefficiently low when information is endogenous. However, in environments with payoff interdependencies, the equilibrium sensitivity to private information can be too high when information is exogenous and, by implication, can remain too high when information is endogenous. When this is the case, the equilibrium features too much, not too little, social learning.[20]

Finally, note that the efficient allocation might feature $\gamma_t^{**} > 1$ if Δ_{t+1} is sufficiently high. That is, a sufficiently strong informational externality could, not only reduce the sensitivity of activity to public information, but even make it change sign. We can easily accommodate this possibility in Propositions 11 and 12 below, but it would complicate the exposition because we would have to consider two cases depending on whether the sensitivity to public information changes sign.

---

[18]The second-order conditions of the efficient allocation guarantee that Δ_{t+1} is small enough so that the denominator of γ_t is also positive.

[19]Here it is important to note that increasing the precision of either exogenous or endogenous signals need not be welfare-improving under the *equilibrium* allocation, but it is always so under the *efficient* allocation; we will come back to this point in Section 7.

[20]Moreover, in this case the failure to internalize the informational externalities, other things equal, improves the efficiency of the equilibrium, for it balances the underlying payoff externalities.

Moreover, we do not expect this possibility to be relevant for applications. We thus opt to rule it out, without any serious loss of generality.

Assumption. *The efficient allocation features $\gamma_t^{**} \in (0, 1)$.*

The following alternative representation of the efficient allocation then helps translate the impact of the informational externality in terms of an implicit desired degree of complementarity in the agents' choices.

Proposition 11. *There exists a unique sequence $\{\alpha_t^{**}\}_{t=1}^N$, with $\alpha_t^{**} < \alpha^*$ for all $t < N$, such that the efficient allocation satisfies*

$$k_{i,t} = \kappa^* (\mathbb{E}_{i,t}\theta) + \alpha_t^{**} \cdot \mathbb{E}_{i,t} [K_t - \kappa^* (\theta)]. \tag{17}$$

As in the case without informational externalities, the weight α_t^{**} in condition (17) summarizes how much society would like the agents to factor their expectations of other agents' choices in their own choices. Unlike the case without informational externalities, this weight now depends on the information structure. Nevertheless, condition (17) remains a valid and insightful representation of the optimal strategy: the result that $\alpha_t^{**} < \alpha^*$ highlights that having the agents internalize the informational externality is isomorphic to having them perceive a lower complementarity in their actions than the one they should have perceived had information been exogenous.

6.4   Policy

We are now ready to characterize optimal policy. We consider tax schemes that make the tax paid by an agent in each period contingent on public information regarding the realized aggregate activity and the realized aggregate fundamentals. In particular, the tax agent i pays in period $t+1$ on the investment he made in period t is contingent on $\tilde{K}_t, \tilde{\sigma}_t$, and $\tilde{\theta}_{t+1}$, all of which are public information in period $t + 1$:

$$\tau_{i,t+1} = T_{t+1}(k_{i,t}, \tilde{K}_t, \tilde{\sigma}_t, \tilde{\theta}_{t+1}), \tag{18}$$

where T_{t+1} is a quadratic function satisfying the same restrictions as those specified in the baseline framework. The existence and uniqueness of the optimal policy then follows essentially from the same argument as the one with measurement error considered in the baseline framework. Along with the fact that $\alpha_t^{**} < \alpha^*$ in all periods (except the very last one), this gives the following result.

Proposition 12. *There exists a unique policy that implements the efficient allocation. The optimal T_{kK} is higher than what it would have been in the absence of informational externalities.*

Once again, the optimal policy does not require any informational advantage on the side of the government: it merely depends on the agents anticipating when they make their decisions that the

marginal tax they will pay in the future will be contingent on public information about aggregate economic conditions.[21] The final goal may now be different, but the key instrument is the same: by effectively subsidizing the use of sources of information that have little correlated noise (the sufficient statistic $X_{i,t}$ here), a higher contingency of the tax schedule on the realized aggregate activity now also guarantees faster social learning. The type of policies we have identified in this paper thus permit to correct, not only the inefficiency in non-fundamental volatility that we documented in the baseline framework, but also the inefficiency that emerges when agents fail to internalize how their choices affect the aggregation of information in the economy.

7   Implications for the social value of information

Throughout the analysis, we have ruled out policies that convey information to the agents. However, because one of the roles of the government is precisely to collect information that is not readily available to the market (think, e.g., of the macroeconomic data collected by the Bureau of Labor Statistics, the US Census Bureau, or the Federal Reserve Banks), it is important to understand whether, and when, it is socially desirable to reveal such information to the market.

The answer to this question is non-trivial: in general, additional information may reduce equilibrium welfare. However, this can not be the case if policy restores efficiency in the equilibrium use of information. This is because the equilibrium then coincides with the solution to a planning problem where the planner directly controls how agents use their available information and can thus guarantee that any additional information will be used at society's best interest.

Proposition 13. *In general, more precise information can reduce equilibrium welfare. However, policies that restore efficiency in the decentralized use of information also guarantee a positive social value for any information disseminated by policy makers or other institutions.*

This result gives guidance on how one can overcome—or at least alleviate—the kind of problems considered in Morris and Shin (2002), Angeletos and Pavan (2007), and Amador and Weill (2007,

---

[21]There is a slight imprecision here. To implement (18), the tax authorities must observe $k_{i,t}$ for each i. But then, in the absence of other frictions, the government could perfectly uncover K_t. We can bypass this uninteresting complication in at least three ways. First, we can assume that there is a large number of tax bureaucrats, each of whom is effective in monitoring the choices or incomes of specific individuals, so that they can collect taxes according to (18), but are not good in aggregating and communicating information to the central tax authority, so that the latter only gets to observe K_t with measurement error. Moreover, a random amount of total tax revenue is lost, so that total tax revenue does not perfectly reveal K_t. Alternatively, we can introduce some "noise" agents, who choose their $k_{i,t}$ in a completely random way. Provided that the tax authorities can not tell apart these agents from the "rational" ones, we can reinterpret K_t as the (unobserved) activity of the rational agents and $\tilde{K}_t$ as (observed) total activity. Finally, we could have the tax paid by agent i be $\tau_{i,t} = T(\tilde{k}_{i,t}, \tilde{K}_t, \tilde{\sigma}_t, \tilde{\theta}_{t+1})$, where $\tilde{k}_{i,t} = k_{i,t} + \eta_t + \nu_{i,t}$ and where ν_{it} is idiosyncratic noise, and let the agent learn about the common measurement error η_t from the observation of his own $\tilde{k}_{i,t}$. We would then have to adjust some of the analysis in order to incorporate this additional source of learning, but the key insights would remain largely unaffected. Indeed, as the variance of $\nu_{i,t}$ converges to infinity *relatively* to that of all other noises, this source of learning becomes irrelevant and the results remain unaffected.

2008). These papers have identified situations in which equilibrium welfare may decrease with the provision of public information; some have then used this possibility to make a case against transparency in central bank communication. By restoring efficiency in the decentralized use of information, the policies we have identified here help guarantee that welfare increases with more information. This is true no matter whether the initial inefficiency originated in payoff interactions (as in Morris-Shin and Angeletos-Pavan) or informational externalities (as in Amador-Weill). Moreover, whereas the pertinent literature has studied the optimality of central-bank transparency largely in isolation from the corrective role of monetary policy, this result indicates that those two aspects of policy making are far from orthogonal to one another.

8 Concluding remarks

In this paper we sought to identify policies that can control how agents use their dispersed sources of information regarding commonly-relevant fundamentals, without requiring the government to observe or collect the information that is dispersed in the economy. Our key result was that this goal can be achieved by appropriately designing the contingencies of marginal taxes on public information regarding the realized aggregate fundamentals and, most importantly, the realized aggregate activity. While the former contingency has a symmetric effect across all sources of information, the latter contingency has an asymmetric effect: it penalizes the agents relatively more when they react to sources of information that have highly correlated noise. An appropriate design of the two contingencies then helps the government dampen the impact of noise without also dampening the impact of fundamentals; improve how much the agents (or the government itself) can learn through prices, macro data, and other indicators of aggregate activity; and guarantee that welfare will increase with the provision of any additional information.

By introducing dispersed private information on aggregate shocks, our policy exercise made an important methodological deviation from both the Ramsey tradition (which rules out any private information) and the Mirrlees tradition (which allows for private information only about idiosyncratic shocks). To highlight this, we showed that the contingency on realized aggregate activity is essential for restoring efficiency in our class of economies *only* when agents have dispersed private information regarding aggregate shocks; when, instead, information regarding aggregate shocks is common, it suffices to make the tax schedule contingent on the aggregate fundamentals alone.

To isolate the particular type of inefficiencies and policy objectives that we were interested in, we ruled out any redistributive goal for taxation. In many applications, redistributive concerns may interact with the policy objectives we studied. For example, the contingencies we have studied affect how much idiosyncratic risk agents are exposed to; conversely, the progressivity of taxation affects how much agents react to their private information regarding aggregate shocks. It is thus an

important direction for future research to extend our analysis to environments that allow for risk aversion and redistributive concerns. However, this need not affect the key insights of the paper: even when agents are risk averse, the contingency of taxes on aggregate activity remains a powerful implicit tax on the use of sources of information that have highly correlated noise.[22]

In conclusion, the more general contribution of the paper is not the implementation of efficient allocations for a particular class of economies; rather, it is the identification of a simple, but powerful, combination of policy contingencies that can help the government manipulate the decentralized use of information and thereby to control the non-fundamental volatility in aggregate activity and the speed of social learning. This insight may be particularly relevant for the business cycle: not only is it likely that a significant component of the business cycle is driven by correlated errors in the information regarding aggregate productivity and demand conditions that is dispersed among the firms and consumers in the economy, but also the aggregation of this information through prices and macro data may be far from perfect. Finally, this insight is clearly not limited to taxation: the contingencies of monetary policy on realized macroeconomic outcomes could serve a similar role as the tax contingencies studied in this paper. Further exploring how the policy objectives we have identified in this paper filter into the design of optimal fiscal and monetary policies over the business cycle is a promising direction for further research.[23]

Appendix A: Proofs

Proof of Propositions 1 and 2. Step 1 proves the efficiency results in Proposition 1 and in part (i) of Proposition 2. Step 2 proves the corresponding equilibrium results. Finally, Step 3 proves part (ii) of Proposition 2.

Step 1. The efficient strategy can be represented as the solution to the following optimization

---

[22]To see this, consider a risk-averse variant of the example of Section 2: each agent chooses k_i so as to maximize $\mathbb{E}_i U(c_i)$, where $c_i = \theta k_i - \frac{1}{2}k_i^2 - T_i$ and where U is a CARA utility. It is then easy to check that the contingency of the tax on $\bar{\theta}$ continues to have a symmetric effect on the sensitivity of investment to the available signals, while the contingency on K continues to have an asymmetric effect; thus, once again, the latter contingency is the key to controlling the signal-to-noise ratio in aggregate activity. See also Angeletos and La'O (2008) for an application in which consumers have CRRA preferences and yet risk aversion does not interfere with the policy exercise of interest; this is because dispersed information impact production choices without inducing idiosyncratic consumption risk.

[23]See Angeletos and La'O (2008) for some work in this direction.

problem (the "planner's problem"): choose a function $k : \Omega \to \mathbb{R}$ so as to maximize

$$\mathbb{E}u = \int_F \int_{\Omega,\Theta} V(k(\omega), K(\phi), \sigma_k(\phi), \theta, \bar{\theta}) df(\omega, \theta) d\mathcal{P}(f) \tag{19}$$

subject to

$$K(\phi) = \int_\Omega k(\omega) d\phi(\omega) \ \forall \phi \in \Phi \qquad \text{and} \qquad \sigma_k(\phi)^2 = \int_\Omega [k(\omega) - K(\phi)]^2 d\phi(\omega) \ \forall \phi \in \Phi. \tag{20}$$

The strict concavity of V ensures that a solution to this problem exists and is unique. Moreover, the solution can be characterized with standard Lagrangian methods. Let $G(\phi)$ denote the marginal distribution of $\mathcal{P}$ over Φ and, for any $\phi \in \Phi$, let $Z(\theta, \bar{\theta}|\phi)$ denote the distribution of $(\theta, \bar{\theta})$ obtained from the distribution $\mathcal{P}$ conditioning on the event that the cross-sectional distribution of information is ϕ. The Lagrangian for this problem can be written as follows:

$$
\begin{aligned}
\Lambda = \ & \int_\Phi \int_{\Theta^2} \int_\Omega V(k(\omega), K(\phi), \sigma_k(\phi), \theta, \bar{\theta}) d\phi(\omega) dZ(\theta, \bar{\theta}|\phi) dG(\phi) \\
& + \int_\Phi \lambda(\phi) \left[K(\phi) - \int_\Omega k(\omega) d\phi(\omega) \right] dG(\phi) \\
& + \int_\Phi \eta(\phi) \left[\sigma_k^2(\phi) - \int_\Omega [k(\omega) - K(\phi)]^2 d\phi(\omega) \right] dG(\phi)
\end{aligned}
$$

where $\lambda(\phi)$ and $\eta(\phi)$ are the Lagrange multipliers associated with the constraints in (20). The first order conditions with respect to $K(\phi)$, $\sigma_k(\phi)$, and $k(\omega)$ are given by the following:[24]

$$\int_{\Theta^2} \int_\Omega V_K(k(\omega), K(\phi), \theta, \bar{\theta}) d\phi(\omega) \, dZ(\theta, \bar{\theta}|\phi) + \lambda(\phi) = 0 \quad \forall \phi \tag{21}$$

$$\int_{\Theta^2} \int_\Omega V_\sigma(k(\omega), K(\phi), \sigma_k(\phi), \theta) d\phi(\omega) \, dZ(\theta, \bar{\theta}|\phi) + 2\eta(\phi)\sigma_k(\phi) = 0 \quad \forall \phi \tag{22}$$

$$\int_{\Theta^2 \times \Phi} \left[V_k(k(\omega), K(\phi), \theta, \bar{\theta}) - \lambda(\phi) - 2\eta(\phi)(k(\omega) - K(\phi)) \right] dP(\theta, \bar{\theta}, \phi|\omega) = 0 \quad \forall \omega \tag{23}$$

where $P(\theta, \bar{\theta}, \phi|\omega)$ denotes the distribution of $(\theta, \bar{\theta}, \phi)$ conditional on ω (i.e., the posterior of an agent about $\theta, \bar{\theta}$ and ϕ). Using the facts that V_K is linear, that $K(\phi) = \int_\Omega k(\omega) d\phi(\omega)$, and that $V_\sigma(k, K, \sigma_k, \theta) = V_{\sigma\sigma}\sigma_k$, conditions (21) and (22) reduce to $\eta(\phi) = -\frac{1}{2}V_{\sigma\sigma}$ and

$$\lambda(\phi) = -\int_{\Theta^2} V_K(K(\phi), K(\phi), \theta, \bar{\theta}) dZ(\theta, \bar{\theta}|\phi) = -V_K(K(\phi), K(\phi), \mathbb{E}[\bar{\theta}|\phi], \mathbb{E}[\bar{\theta}|\phi]).$$

Substituting the above into (23) and noting that $\mathbb{E}[\bar{\theta}|\phi] = \mathbb{E}[\bar{\theta}|\phi, \omega]$ and hence $\mathbb{E}[\mathbb{E}[\bar{\theta}|\phi]|\omega] = \mathbb{E}[\bar{\theta}|\omega]$, we conclude that the strategy $k : \Omega \to \mathbb{R}$ is efficient if and only if it satisfies the following condition for all $\omega \in \Omega$:

$$\mathbb{E} \left[V_k(k(\omega), K(\phi), \theta, \bar{\theta}) + V_K(K(\phi), K(\phi), \bar{\theta}, \bar{\theta}) + V_{\sigma\sigma}(k(\omega) - K(\phi)) \, \bigg| \, \omega \right] = 0. \tag{24}$$

---

[24] Recall that, because $V_\sigma(\cdot) = V_{\sigma\sigma}\sigma$, the derivatives V_k and V_K do not depend on σ.

Finally, by the linearity of V_k, we have that $V_k(k, K, \theta, \bar{\theta}) = V_k(K, K, \bar{\theta}, \bar{\theta}) + V_{kk}(k - K) + V_{k\theta}(\theta - \bar{\theta})$ and hence the above can be restated as

$$\mathbb{E}_i \Big[V_k(K, K, \bar{\theta}, \bar{\theta}) + V_K(K, K, \bar{\theta}, \bar{\theta}) + V_{k\theta}(\theta_i - \bar{\theta}) + (V_{kk} + V_{\sigma\sigma})(k - K) \Big] = 0. \qquad (25)$$

Consider first the case where ϕ is common knowledge. Condition (25) reduces to

$$V_K(K, K, \mathbb{E}_i\bar{\theta}, \mathbb{E}_i\bar{\theta}) + V_K(K, K, \mathbb{E}_i\bar{\theta}, \mathbb{E}_i\bar{\theta}) + V_{k\theta}(\mathbb{E}_i\theta_i - \mathbb{E}_i\bar{\theta}) + (V_{kk} + V_{\sigma\sigma})(k_i - K) = 0. \qquad (26)$$

Let $\vartheta^1 \equiv \int \mathbb{E}[\theta|\omega]d\phi(\omega)$ denote the cross-sectional average of $\mathbb{E}_i\theta_i$. Because ϕ is common knowledge, ϑ^1 is also common knowledge and $\mathbb{E}_i\bar{\theta} = \vartheta^1$ for all i.[25] Hence, aggregating (26) across agents gives

$$V_k(K, K, \vartheta^1, \vartheta^1) + V_K(K, K, \vartheta^1, \vartheta^1) = 0. \qquad (27)$$

Condition (26) then reduces to

$$k_i = K + \frac{V_{k\theta}}{V_{kk} + V_{\sigma\sigma}}(\mathbb{E}_i\theta_i - \mathbb{E}_i\bar{\theta}). \qquad (28)$$

Let

$$\kappa_0^* \equiv \frac{V_k(0,0,0,0) + V_K(0,0,0,0)}{-(V_{kk} + 2V_{kK} + V_{KK})}, \quad \kappa_1^* \equiv \frac{V_{k\theta}}{-(V_{kk} + V_{\sigma\sigma})}, \quad \kappa_2^* \equiv \frac{V_{k\theta} + V_{k\bar{\theta}} + V_{K\theta} + V_{K\bar{\theta}}}{-(V_{kk} + 2V_{kK} + V_{KK})} - \kappa_1^*. \qquad (29)$$

Solving (27) and (28) then gives $K = \kappa_0^* + (\kappa_1^* + \kappa_2^*)\vartheta^1$ and $k_i = K + \kappa_1^*(\mathbb{E}_i\theta_i - \mathbb{E}_i\bar{\theta}) = \kappa_0^* + \kappa_1^*\mathbb{E}_i\theta_i + \kappa_2^*\mathbb{E}_i\bar{\theta}$, which gives the efficiency result of Proposition 1.

Next, consider the case where ϕ is not common knowledge. Because both V_k and V_K are linear, the first two terms in condition (25) can be rewritten as

$$V_k(\kappa^*(\bar{\theta}, \bar{\theta}), \kappa^*(\bar{\theta}, \bar{\theta}), \bar{\theta}, \bar{\theta}) + V_K(\kappa^*(\bar{\theta}, \bar{\theta}), \kappa^*(\bar{\theta}, \bar{\theta}), \bar{\theta}, \bar{\theta}) + (V_{kk} + 2V_{kK} + V_{KK})(K - \kappa^*(\bar{\theta}, \bar{\theta})).$$

By the definition of κ^*, the first two terms are zero. It follows that condition (25) reduces to

$$(V_{kk} + 2V_{kK} + V_{KK})\mathbb{E}_i(K - \kappa^*(\bar{\theta}, \bar{\theta})) + V_{k\theta}(\mathbb{E}_i\theta_i - \mathbb{E}_i\bar{\theta}) + (V_{kk} + V_{\sigma\sigma})(k_i - \mathbb{E}_iK) = 0. \qquad (30)$$

Rearranging, and letting

$$\alpha^* \equiv 1 - \frac{V_{kk} + 2V_{kK} + V_{KK}}{V_{kk} + V_{\sigma\sigma}}, \qquad (31)$$

gives condition (7). Finally, $V_{kk} + 2V_{kK} + V_{KK} < 0$ by the concavity of V, while $V_{kk} + V_{\sigma\sigma} < 0$ by

---

[25]Note that when ϕ is common knowledge, $\vartheta^1 \equiv \mathbb{E}[\mathbb{E}[\theta|\omega]|\phi] = \mathbb{E}[\mathbb{E}[\theta|\omega, \phi]|\phi] = \mathbb{E}[\theta|\phi]$; furthermore, because ω cannot contain more information about h (and hence about $\bar{\theta}$) than ϕ, $\mathbb{E}[\bar{\theta}|\omega] = \mathbb{E}[\bar{\theta}|\phi] = \mathbb{E}[\mathbb{E}[\theta|h, \phi]|\phi] = \mathbb{E}[\theta|\phi]$.

assumption. This guarantees that $\alpha^* < 1$ and completes the proof of the efficiency result in part (i) of Proposition 2.

Step 2. Because V is strictly concave in k, the best response of agent i is pinned down by the first-order condition $\mathbb{E}_i V_k \left(k_i, K, \theta_i, \bar{\theta} \right) = 0$. Furthermore, because V is quadratic in (k, K, θ), this first-order condition reduces to

$$V_k (0,0,0,0) + V_{kk} k_i + V_{kK} \mathbb{E}_i K + V_{k\theta} \mathbb{E}_i \theta_i + V_{k\bar{\theta}} \mathbb{E}_i \bar{\theta} = 0. \tag{32}$$

Consider first the case where ϕ is common knowledge. Aggregating (32) across all i gives

$$V_k (0,0,0,0) + (V_{kk} + V_{kK})K + (V_{k\theta} + V_{k\bar{\theta}})\vartheta^1 = 0. \tag{33}$$

Condition (32) then reduces to

$$k_i = K + \frac{V_{k\theta}}{V_{kk}}(\mathbb{E}_i \theta_i - \mathbb{E}_i \bar{\theta}). \tag{34}$$

Let

$$\kappa_0 \equiv \frac{V_k (0,0,0,0)}{-(V_{kk} + V_{kK})}, \quad \kappa_1 \equiv \frac{V_{k\theta}}{-V_{kk}}, \quad \kappa_2 \equiv \frac{V_{k\theta} + V_{k\bar{\theta}}}{-(V_{kk} + V_{kK})} - \kappa_1. \tag{35}$$

Solving (33) and (34) for k_i and K then gives $K = \kappa_0 + (\kappa_1 + \kappa_2)\vartheta^1$ and $k_i = K + \kappa_1(\mathbb{E}_i \theta_i - \mathbb{E}_i \bar{\theta}) = \kappa_0 + \kappa_1 \mathbb{E}_i \theta_i + \kappa_2 \mathbb{E}_i \bar{\theta}$. This establishes existence and uniqueness of equilibrium and gives condition (4) of Proposition 1.

Next consider the case that ϕ is not common knowledge. By the definition of κ,

$$V_k (0,0,0,0) + V_{kk}\kappa(\theta, \bar{\theta}) + V_{kK}\kappa(\bar{\theta}, \bar{\theta}) + V_{k\theta}\theta + V_{k\bar{\theta}}\bar{\theta} = 0.$$

Using the above, condition (32) reduces to

$$V_{kk}\mathbb{E}_i \left[k_i - \kappa(\theta, \bar{\theta}) \right] + V_{kK}[K - \kappa(\bar{\theta}, \bar{\theta})] = 0.$$

Letting

$$\alpha \equiv -\frac{V_{kK}}{V_{kk}} \tag{36}$$

and rearranging gives condition (6).

Clearly, the above argument establishes that (6) is both necessary and sufficient for any equilibrium. What then remains to prove is that the equilibrium exists and is unique. This can be done with the help of Step 1. First, note that an economy is parameterized by $\mathbf{e} \equiv (V, \Omega, F, \mathcal{P})$. Next, note that for every $V \in \mathcal{V}$ there exists a $V' \in \mathcal{V}$ such that such that the κ^* and α^* corresponding to V' coincide with the κ and α corresponding to V. By comparing conditions (6) and (7), it is immediate that the set of equilibrium strategies for the economy $\mathbf{e} = (V, \Omega, F, \mathcal{P})$ coincides with the

set of efficient strategies for the economy $\mathbf{e}' = (V', \Omega, F, \mathcal{P})$. The existence and uniqueness of the equilibrium for economy $\mathbf{e}$ then follows from the existence and uniqueness of the efficient allocation of the economy $\mathbf{e}'$, which we established in Step 1.

Step 3. Consider part (ii) of Proposition 2. We prove the result for the equilibrium; the proof for the efficient allocation is analogous. From (6),

$$k_i = (1-\alpha)(\kappa_0 + (\kappa_1 + \kappa_2)\mathbb{E}_i\bar{\theta}) + \kappa_1(\mathbb{E}_i\theta_i - \mathbb{E}_i\bar{\theta}) + \alpha\mathbb{E}_iK, \tag{37}$$

and therefore

$$K = (1-\alpha)(\kappa_0 + (\kappa_1 + \kappa_2)\bar{\vartheta}^1) + \kappa_1(\vartheta^1 - \bar{\vartheta}^1) + \alpha\int\mathbb{E}[K|\omega]d\phi(\omega).$$

Iterating, we obtain

$$K = \kappa_0 + (\kappa_1 + \kappa_2)\sum_{n=1}^{\infty}(1-\alpha)\alpha^{n-1}\bar{\vartheta}^n + \kappa_1\sum_{n=1}^{\infty}\alpha^{n-1}(\vartheta^n - \bar{\vartheta}^n),$$

Substituting the above into (37) and rearranging gives the result. ∎

Proof of Proposition 3. For any signal $s \in \{1, ..., n\}$, let $\pi_s \equiv \sigma_s^{-2}$ denote its precision; let $\rho_s \equiv Corr(\xi_{si}, \xi_{sj})$, for $i \neq j$, denote the correlation of the noise across any pair of agents; let $\bar{\xi}_s$ denote the mean realization of ξ_{si} in the population (the common component of noise). We order the signals so that $\rho_1 > \rho_2 > ... > \rho_s$ and let $\pi_0 \equiv \sigma_\theta^2$ denote the precision of the prior. By standard Gaussian updating, the posterior of agent i about $\bar{\theta}$ satisfies

$$\mathbb{E}_i\bar{\theta} = \delta_0\mu + \sum_s \delta_s x_{si},$$

where $\delta_s = \pi_s/\pi$ for $s \in \{0, 1, ..., n\}$ and $\pi = \pi_0 + \sum_s \pi_s$, while his posterior about the common noise in the s-th signal satisfies

$$\mathbb{E}_i\bar{\xi}_s = \rho_s(x_{is} - \mathbb{E}_i\theta).$$

Finally, from Proposition 2 we know that a strategy is an equilibrium if and only if it satisfies

$$k_i = (1-\alpha)\,\mathbb{E}_i\kappa(\bar{\theta}, \bar{\theta}) + \alpha\mathbb{E}_iK. \tag{38}$$

We now guess and verify that the equilibrium strategy is linear in the available signals.

Suppose there exist coefficients $(\beta_0, \beta_1, ..., \beta_n)$ such that

$$k_i = \beta_0 + \sum_s \beta_s x_{si}. \tag{39}$$

It then follows that

$$K = \beta_0 + \left(\sum_s \beta_s\right)\bar\theta + \sum_s \left(\beta_s \bar\xi_s\right), \tag{40}$$

and therefore $\mathbb{E}_i K = \beta_0 + \left(\sum_s \beta_s(1-\rho_s)\right)\mathbb{E}_i\theta + \sum_s \left(\beta_s\rho_s x_s\right) + \beta_2\rho_2 x_2$. Substituting the latter into (38) and requiring that the resulting expression coincides (39) for all realizations of the signals, we conclude that the coefficients $(\beta_0, \beta_1, ..., \beta_n)$ must solve the following system:

$$\left(\beta_0 - \kappa_0\right)\left(1 - \alpha\right) = \left\{(1-\alpha)\left(\kappa_1 + \kappa_2\right) + \alpha\sum_{s'}\beta_{s'}(1-\rho_{s'})\right\}\delta_0\mu \tag{41}$$

$$\beta_s\left(1 - \alpha\rho_s\right) = \left\{(1-\alpha)\left(\kappa_1 + \kappa_2\right) + \alpha\sum_{s'}\beta_{s'}(1-\rho_{s'})\right\}\delta_s, \;\; \forall s \in \{1, ..., n\} \tag{42}$$

The unique solution to this system gives

$$\beta_s = (\kappa_1 + \kappa_2)\frac{\delta_s\prod_{s'\neq s}(1-\alpha\rho_{s'})}{denom} \;\; \forall s \in \{1, ..., n\},$$

where $denom \equiv \delta_0\prod_s(1-\alpha\rho_s) + (1-\alpha)\sum_s\delta_s\prod_{s'\neq s}(1-\alpha\rho_{s'})$. It is then immediate that, for any $s, s' \in \{1, ..., n\}$,

$$\frac{\beta_s}{\beta_{s'}} = \frac{\delta_s}{\delta_{s'}}\frac{1-\alpha\rho_{s'}}{1-\alpha\rho_s},$$

which implies that $\beta_s/\beta_{s'}$ increases with α if and only if $\rho_s > \rho_{s'}$; that is, a higher α tilts the equilibrium use of information towards the signal with the most correlated noise.

Consider now the signal-to-noise ratio in aggregate activity. From (40), the component of aggregate activity that is explained by fundamentals is $\hat{K} = \mathbb{E}[K|\bar\theta] = \beta_0 + \sum_s\beta_s\theta$; the residual, $K - \hat{K} = \sum_s\beta_s\bar\xi_s$, gives the non-fundamental component. It follows that the equilibrium signal-to-noise ratio is given by

$$R \equiv \frac{Var(\hat{K})}{Var(K - \hat{K})} = \frac{(\sum_s\beta_s)^2 Var(\bar\theta)}{\sum_s\beta_s^2 Var(\xi_s)} = \frac{\left(\sum_s\left\{\prod_{s'\neq s}(1-\alpha\rho_{s'})\right\}\delta_s\right)^2}{\delta_0\left(\sum_s\left\{\prod_{s'\neq s}(1-\alpha\rho_{s'})^2\right\}\rho_s\delta_s\right)}.$$

The latter is independent of κ and is decreasing in α:

$$\frac{\partial R}{\partial\alpha} = -\frac{2\left(\sum_s\left\{\prod_{s'\neq s}(1-\alpha\rho_{s'})\right\}\delta_s\right)}{\delta_0\left(\sum_s\left\{\prod_{s'\neq s}(1-\alpha\rho_{s'})^2\right\}\rho_s\delta_s\right)^2}\left(\sum_{s,s'}\left\{\prod_{s''\neq s,s'}(1-\alpha\rho_{s''})^3\right\}\delta_s\delta_{s'}\left(\rho_s - \rho_{s'}\right)^2\right) < 0. \tag{43}$$

Finally, since the signal-to-noise ratio along the efficient allocation is given by the same formula replacing α with α^*, it is immediate that the equilibrium ratio is inefficiently high if and only if $\alpha > \alpha^*$. ∎

34

Proof of Proposition 4. Given any policy $T \in \mathcal{T}$, let

$$\tilde{V}\left(k, K, \sigma_k, \theta, \bar{\theta}\right) \equiv V\left(k, K, \sigma_k, \theta, \bar{\theta}\right) - T(k, K, \sigma_k, \bar{\theta})$$

denote an agent's payoff, net of taxes. The restrictions we have imposed on T guarantee that $\tilde{V} \in \mathcal{V}$. The characterization of the equilibrium then follows directly from the proofs of Propositions 1 and 2, replacing the function V with the function $\tilde{V}$. Using the formulas for (35) and (36) thus gives

$$\tilde{\kappa}_0 \;=\; \frac{\tilde{V}_k\left(0,0,0,0\right)}{-\tilde{V}_{kk} - \tilde{V}_{kK}} = \frac{V_k\left(0,0,0,0\right) - T_k\left(0,0,0\right)}{-V_{kk} - V_{kK} + T_{kk} + T_{kK}} = \frac{(1-\alpha)\kappa_0 + \frac{1}{V_{kk}}T_k\left(0,0,0\right)}{1 - \alpha - \frac{1}{V_{kk}}T_{kk} - \frac{1}{V_{kk}}T_{kK}} \tag{44}$$

$$\tilde{\kappa}_1 \;=\; \frac{\tilde{V}_{k\theta}}{-\tilde{V}_{kk}} = \frac{V_{k\theta}}{-V_{kk} + T_{kk}} = \frac{1}{1 - \frac{1}{V_{kk}}T_{kk}}\kappa_1 \tag{45}$$

$$\tilde{\kappa}_2 \;=\; \frac{\tilde{V}_{k\theta} + \tilde{V}_{k\bar{\theta}}}{-\tilde{V}_{kk} - \tilde{V}_{kK}} - \tilde{\kappa}_1 = \frac{V_{k\theta} + V_{k\bar{\theta}} - T_{k\bar{\theta}}}{-V_{kk} - V_{kK} + T_{kk} + T_{kK}} - \tilde{\kappa}_1 = \frac{(1-\alpha)\left(\kappa_1 + \kappa_2\right) + \frac{1}{V_{kk}}T_{k\bar{\theta}}}{1 - \alpha - \frac{1}{V_{kk}}T_{kk} - \frac{1}{V_{kk}}T_{kK}} - \tilde{\kappa}_1 \tag{46}$$

$$\tilde{\alpha} \;=\; \frac{\tilde{V}_{kK}}{-\tilde{V}_{kk}} = \frac{V_{kK} - T_{kK}}{-V_{kk} + T_{kk}} = \frac{\alpha + \frac{1}{V_{kk}}T_{kK}}{1 - \frac{1}{V_{kk}}T_{kk}}. \tag{47}$$

Normalizing $V_{kk} = -1$ then gives the formulas in the proposition. $\blacksquare$

Proof of Proposition 5. By the same argument as in the proof of Proposition 3, the equilibrium signal-to-noise ratio induced by the policy is given by

$$R = \frac{\left(\sum_s \left\{\prod_{s'\neq s}\left(1 - \tilde{\alpha}\rho_{s'}\right)\right\}\delta_s\right)^2}{\delta_0\left(\sum_s \left\{\prod_{s'\neq s}\left(1 - \tilde{\alpha}\rho_{s'}\right)^2\right\}\rho_s\delta_s\right)}.$$

This depends on the policy only through $\tilde{\alpha}$. It is thus independent of $T_{k\bar{\theta}}$ and increasing in T_{kK}. $\blacksquare$

Proof of Proposition 6. We prove the result in reverse ordering.

Part (ii). Consider the case that ϕ is not common knowledge. Provided that the information structure is regular in the sense of footnote 8, the policy implements the efficient allocation if and only if it induces $\tilde{\kappa}_0 = \kappa_0^*$, $\tilde{\kappa}_1 = \kappa_1^*$, $\tilde{\kappa}_2 = \kappa_2^*$, and $\tilde{\alpha} = \alpha^*$. It thus suffices to prove that there exists a policy $T^* \in \mathcal{T}$ that does so, and that this policy is unique. This is easily shown from conditions (44)-(46) and (47). First, note that $\tilde{\kappa}_1 = \kappa_1^*$ if and only if

$$T_{kk} = V_{kk}(1 - \kappa_1/\kappa_1^*) = -V_{\sigma\sigma}. \tag{48}$$

Along with the assumption that $V_{kk} + V_{\sigma\sigma} < 0$, this also guarantees that $V_{kk} - T_{kk} < 0$. Next, note

that, since $T_{kk} = -V_{\sigma\sigma}$, $\tilde{\alpha} = \alpha^*$ if and only

$$T_{kK} = -V_{kk}(\alpha - \alpha^*) - T_{kk}\alpha^* = -V_{kk}\alpha + (V_{kk} + V_{\sigma\sigma})\alpha^* = V_{\sigma\sigma} - V_{kK} - V_{KK}. \qquad (49)$$

It is then immediate that $T_{kK} \neq 0$ for all but a (Lebesgue) measure zero set of payoff functions V for which $V_{\sigma\sigma} - V_{kK} - V_{KK} = 0$. With (T_{kk}, T_{kK}) thus determined, it is then immediate that there exist a unique $T_{k\bar{\theta}}$ such that $\tilde{\kappa}_2 = \kappa_2^*$ and a unique $T_k(0,0,0)$ such that $\tilde{\kappa}_0 = \kappa_0^*$; these are given by

$$\begin{aligned} T_{k\bar{\theta}} &= V_{kk}(1-\alpha)\left[(\kappa_1^* + \kappa_2^*) - (\kappa_1 + \kappa_2)\right] - (T_{kk} + T_{kK})(\kappa_1^* + \kappa_2^*) = -V_{K\theta} - V_{K\bar{\theta}}, (50) \\ T_k(0,0,0) &= V_{kk}(1-\alpha)(\kappa_0^* - \kappa_0) - (T_{kk} + T_{kK})\kappa_0^* = -V_K(0,0,0). \end{aligned}$$

Finally, for T to balance the budget (state by state) it must be that $T(K, K, 0, \bar{\theta}) = 0$ for all $(K, \bar{\theta})$ and that $T_{kk} + T_{\sigma\sigma} = 0$. Along with the other properties identified above, it is then easy to verify that this is equivalent to imposing the following: $T(0,0,0,0) = T_{\bar{\theta}}(0,0,0) = T_{\bar{\theta}\bar{\theta}} = 0$; $T_K(0,0,0) = -T_k(0,0,0) = V_K(0,0,0)$; $T_{KK} = -2T_{kK} - T_{kk} = -V_{\sigma\sigma} + 2V_{kK} + 2V_{KK}$; $T_{K\bar{\theta}} = -T_{k\bar{\theta}} = V_{K\theta} + V_{K\bar{\theta}}$; and finally $T_{\sigma\sigma} = -T_{kk}$. This also implies that the policy T is unique.

Part (i). Consider the case that ϕ is common knowledge. Now $\tilde{\alpha}$ and α^* are irrelevant and the policy implements the efficient allocation if and only if it induces $\tilde{\kappa}_0 = \kappa_0^*$, $\tilde{\kappa}_1 = \kappa_1^*$, and $\tilde{\kappa}_2 = \kappa_2^*$; whether $\tilde{\alpha}$ is equal to α^* is no longer relevant. Once again, there is a unique T_{kk} that induces $\tilde{\kappa}_1 = \kappa_1^*$. However, because there is no need to induce $\tilde{\alpha} = \alpha^*$, T_{kK} is free. As a result, $\tilde{\kappa}_2 = \kappa_2^*$ can now be induced by appropriately setting either $T_{k\bar{\theta}}$ or T_{kK}. It is then without any loss of optimality to set $T_{kK} = 0$ (or to any other arbitrary value) and then set $T_{k\bar{\theta}}$ as in (50). The rest of the parameters of the policy are then determined as in the proof of part (ii) above. ∎

Proof of Proposition 7. In the case of additive measurement error, $\mathbb{E}_i T(\tilde{k}_i, \tilde{K}, \tilde{\sigma}_k, \tilde{\theta}) = \mathbb{E}_i T(k_i, K, \sigma_k, \bar{\theta}) + SOT$, with

$$SOT = \frac{1}{2}(T_{kk} + 2T_{kK} + T_{KK})\sigma_\eta^2 + \frac{1}{2}(T_{kk} + T_{\sigma\sigma})\sigma_v^2 + T_{\bar{\theta}\bar{\theta}}\sigma_\varsigma^2.$$

The result then follows from noting that SOT is independent of k_i and therefore does not affect individual incentives. In the case of multiplicative measurement error,

$$SOT = SOT(k_i, K) = \tfrac{1}{2}T_{kk}\left(\sigma_\eta^2 + \sigma_v^2\right)k_i^2 + T_{kK}\sigma_\eta^2 k_i K + \tfrac{1}{2}T_{KK}\sigma_\eta^2 K^2 + \tfrac{1}{2}T_{\sigma\sigma}\sigma_\eta^2\sigma_k^2.$$

In this case, the measurement error does affect incentives, but this does not complicate the results.

Indeed, all the steps in Propositions 4, 5 and 6 hold with $\tilde{\kappa}$ and $\tilde{\alpha}$ redefined as follows:

$$\tilde{\kappa}_0 = \frac{V_k(0,0,0,0) - T_k(0,0,0)}{-V_{kk} - V_{kK} + T_{kk}\left(1+\sigma_\eta^2+\sigma_v^2\right) + T_{kK}(1+\sigma_\eta^2)} \qquad \tilde{\kappa}_1 = \frac{V_{k\theta}}{-V_{kk} + T_{kk}\left(1+\sigma_\eta^2+\sigma_v^2\right)}$$

$$\tilde{\kappa}_2 = \frac{V_{k\theta} + V_{k\bar{\theta}} - T_{k\bar{\theta}}}{-V_{kk} - V_{kK} + T_{kk}\left(1+\sigma_\eta^2+\sigma_v^2\right) + T_{kK}(1+\sigma_\eta^2)} - \tilde{\kappa}_1 \qquad \tilde{\alpha} = \frac{V_{kK} - T_{kK}(1+\sigma_\eta^2)}{-V_{kk} + T_{kk}\left(1+\sigma_\eta^2+\sigma_v^2\right)}$$

The only difference is that the optimal tax now depends on σ_η^2 and σ_v^2. ∎

Proof of Proposition 8. First, consider $t = 1$. Because information is exogenous in this period, that the equilibrium strategy at $t = 1$ is unique and solves (11) follows directly from the same argument as in Steps 1 and 2 of the proof of Proposition 2, with $\kappa(\theta) = \kappa_0 + (\kappa_1 + \kappa_2)\theta$ and with the coefficients $(\kappa_0, \kappa_1, \kappa_2)$ determined as in (35). Next consider $t = 2$. The information structure is now endogenous but uniquely determined by the unique equilibrium strategy for $t = 1$. That the equilibrium strategy at $t = 2$ is unique and solves (11) then follows again from Proposition 2. Repeating the same argument for all $t > 2$ establishes the result.

Proof of Proposition 9. First, consider $t = 1$. In this period, information is exogenous, with $\omega_{i,1} = (x_{i,1}, y_1, A_1)$. Standard Gaussian updating then gives

$$\mathbb{E}[\theta|\omega_{i,1}] = \frac{\pi_1^x}{\pi_1^x + \pi_1^y} X_{i,1} + \frac{\pi_1^y}{\pi_1^x + \pi_1^y} Y_1, \tag{51}$$

where $X_{i,1} = x_{i,1}$, $\pi_1^x = \sigma_{x,1}^{-2}$, $Y_1 = \frac{\sigma_\theta^{-2}}{\pi_1^y}\mu + \frac{\sigma_{y,1}^{-2}}{\pi_1^y}y_1 + \frac{\sigma_{a,1}^{-2}}{\pi_1^y}\tilde{\theta}_1$ and $\pi_1^y = \sigma_\theta^{-2} + \sigma_{y,1}^{-2} + \sigma_{a,1}^{-2}$. We then have that the unique solution to (11) is given by

$$k_1(\omega_{i,1}) = \kappa(\gamma_1 X_{i,1} + (1-\gamma_1)Y_1), \tag{52}$$

with $\gamma_1 \equiv [(1-\alpha)\pi_1^x] / [(1-\alpha)\pi_1^x + \pi_1^y]$. To see this, start by guessing that the equilibrium strategy satisfies (52) for *some* coefficient γ_1. Next, use this guess to compute aggregate activity as $K_1 = \kappa(\gamma_1\theta + (1-\gamma_1)Y_1)$. Finally, use the latter along with (11) and (51) to derive the equilibrium γ_1.

Next, consider $t = 2$. In the second period, $\omega_{i,2} = \omega_{i,1} \cup (x_{i,2}, y_2, \tilde{\theta}_2, \tilde{K}_1, \tilde{\sigma}_1)$. The endogenous signal is given by

$$\tilde{K}_1 = \kappa(\gamma_1\theta + (1-\gamma_1)Y_1) + \eta_2$$

The information about θ contained in $\tilde{K}_1$ is thus the same as that contained in

$$\tilde{y}_2 \equiv \frac{\tilde{K}_1 - \kappa((1-\gamma_1)Y_1)}{\kappa'\gamma_1} = \theta + \tilde{\eta}_2,$$

where $\tilde{\eta}_2 = \eta_2/[\kappa'\gamma_1]$ is Gaussian noise with variance $\sigma_{\tilde{\eta},2}^2 = \sigma_{\eta,2}^2/(\kappa')^2\gamma_1^2$. The signal $\tilde{\sigma}_1$, on the other hand, conveys no information about θ, because (52) implies that $\sigma_1 = (\kappa')^2\gamma_1^2\sigma_{x,1}^2$, which is

common knowledge. It follows that the period-2 public information about θ can be summarized in a sufficient statistic Y_2 such that the posterior about θ conditional on $(y_1, \tilde{\theta}_1, \tilde{K}_1, \tilde{\sigma}_1, y_2, \tilde{\theta}_2)$ is Gaussian with mean

$$Y_2 = \frac{\pi_1^y}{\pi_2^y} Y_1 + \frac{\sigma_{y,2}^{-2}}{\pi_2^y} y_2 + \frac{\sigma_{a,2}^{-2}}{\pi_2^y} \tilde{\theta}_2 + \frac{\gamma_1^2 (\kappa')^2 \sigma_{\eta,2}^{-2}}{\pi_2^y} \tilde{y}_2$$

and precision $\pi_2^y = \pi_1^y + \sigma_{y,2}^{-2} + \sigma_{a,2}^{-2} + \gamma_1^2 (\kappa')^2 \sigma_{\eta,2}^{-2}$. Similarly, the private information can be summarized in the sufficient statistic $X_{i,2}$ such that the posterior about θ conditional on $(x_{i,1}, x_{i,2})$ is Gaussian with mean

$$X_{i,2} = \frac{\pi_1^x}{\pi_2^x} X_{i,1} + \frac{\sigma_{x,2}^{-2}}{\pi_2^x} x_{i,2}$$

and precision $\pi_2^x = \pi_1^x + \sigma_{x,2}^{-2}$. The unique solution to (11) is then $k_2(\omega_{i,2}) = \kappa(\gamma_2 X_{i,2} + (1-\gamma_2) Y_2)$, with $\gamma_2 \equiv [(1-\alpha) \pi_2^x] / [(1-\alpha) \pi_2^x + \pi_2^y]$.

A similar argument applied also to $t \geq 3$, establishing that the unique equilibrium strategy is $k_{i,t}(\omega_{i,t}) = \kappa(\gamma_t X_{i,t} + (1-\gamma_t) Y_t)$, with $\gamma_t \equiv [(1-\alpha) \pi_t^x] / [(1-\alpha) \pi_t^x + \pi_t^y]$ and with the statistics $X_{i,t}$ an Y_t defined recursively as in the proposition. ∎

Proof of Proposition 10. We prove the result in three steps. Step 1 establishes that the optimal strategy is linear in the sufficient statistics for the case where there are no payoff interactions; this step echoes a similar result by Vives (1993). Step 2 extends this property to the case with payoff interactions. Step 3 completes the result by characterizing the optimal weight on the two statistics.

Step 1. We momentarily rule out payoff interactions by assuming

$$V(k, K, \sigma, \theta) = v(k, \theta) \equiv -(k - \theta)^2.$$

Let $h_t = \{y_1, \tilde{\theta}_1, \tilde{K}_1, ..., y_{t-1}, \tilde{\theta}_{t-1}, \tilde{K}_{t-1}, y_t, \tilde{\theta}_t\}$ denote the public history in period t and suppose agents follow a strategy $\mathbf{k} = \{k_t\}_{t=1}^N$ such that

$$k_t(\omega_{i,t}) = P_t(h_t) + \sum_{\tau=1}^t Q_{t,\tau} x_{i,\tau},$$

where $P_t(h_t)$ is a deterministic function of h_t and $Q_{t,\tau}$ are deterministic coefficients. It follows that $k_{i,t} = P_t + \gamma_t \theta + \sum_{\tau=1}^t Q_{t,\tau} \xi_{i,\tau}$, and hence $\tilde{K}_t = P_t + \gamma_t \theta + \eta_{t+1}$, where P_t is a shortcut for $P_t(h_t)$ and $\gamma_t \equiv \sum_{\tau=1}^t Q_{t,\tau}$. Welfare (ex-ante utility) is then $\mathbb{E}u = \sum_{t=1}^N w_t$, where

$$w_t \equiv \mathbb{E}[v(k_{i,t}, A_{t+1})] = \mathbb{E}[v(k_{i,t}, \theta)] - \sigma_{a,t+1}^2$$

and where

$$\mathbb{E}\left[v\left(k_{i,t},\theta\right)\right] = -\mathbb{E}\left[\mathbb{E}\left[\left\{\left(P_t + \gamma_t\theta + \sum_{\tau=1}^{t} Q_{t,\tau}\xi_{i,\tau}\right) - \theta\right\}^2 \middle| \theta, h_t\right]\right]$$

$$= -\mathbb{E}\left[(P_t + \gamma_t\theta - \theta)^2 + \sum_{\tau=1}^{t} Q_{t,\tau}^2 \sigma_{\xi,\tau}^2\right]$$

Now consider a strategy $\hat{\mathbf{k}} = \{\hat{k}_t\}_{t=1}^N$ that is a variation of the initial strategy $\mathbf{k} = \{k_t\}_{t=1}^N$ constructed as follows. First, pick an arbitrary t and let $\hat{k}_{i,s}(\omega_{i,s}) = k_{i,s}(\omega_{i,s})$ for all $s < t$. Next, in period t, pick an arbitrary function $\hat{P}_t$ and any coefficients $\hat{Q}_{t,\tau}$ such that $\sum_{\tau=1}^{t} \hat{Q}_{t,\tau} = \gamma_t$, and let $\hat{k}_t(\omega_{i,t}) = \hat{P}_t(h_t) + \sum_{\tau=1}^{t} \hat{Q}_{t,\tau}x_{i,\tau}$. Finally, for all $s > t$, let $\hat{k}_s(\omega_{i,s}) = \hat{P}_s(h_s) + \sum_{\tau=1}^{s} Q_{s,\tau}x_{i,\tau}$, where the functions $\hat{P}_s$ are such that $\hat{P}_s(..., \tilde{K}_t, ...) = P_s(..., \tilde{K}_t - \hat{P}_t(h_t) + P_t(h_t), ...)$.

By construction, at any period $s \neq t$, the strategy $\hat{\mathbf{k}}$ induces the same outcomes, and by implication the same per-period welfare level w_t, as the initial strategy $\mathbf{k}$. It follows that a necessary condition for the strategy $\mathbf{k}$ to be efficient is that, for all t and all h_t,

$$\left(P_t, (Q_{t,\tau})_{\tau=1}^t\right) \in \arg\min_{\hat{P}_t, \hat{Q}_{t,\tau}} \mathbb{E}\left[\left(\hat{P}_t + \gamma_t\theta - \theta\right)^2 + \sum_{\tau=1}^{t} \hat{Q}_{t,\tau}^2 \sigma_{\xi,\tau}^2 \middle| h_t\right]$$

$$\text{s.t.} \qquad \sum_{\tau=1}^{t} \hat{Q}_{t,\tau} = \gamma_t$$

This in turn is the case if and only if, for all t and all h_t

$$P_t(h_t) = (1 - \gamma_t)\mathbb{E}[\theta|h_t] \quad \text{and} \quad Q_{t,\tau} = \gamma_t \frac{\sigma_{\xi,\tau}^{-2}}{\sum_{j=1}^{t} \sigma_{\xi,j}^{-2}} \ \forall\tau. \tag{53}$$

Next note that, because P_t is public information, the observation in period $t+1$ of $\tilde{K}_t = K_t + \eta_t = P_t + \gamma_t\theta + \eta_t$ is informationally equivalent to the observation of a signal

$$\tilde{y}_{t+1} \equiv \frac{\tilde{K}_t - P_t}{\gamma_t} = \theta + \tilde{\eta}_{t+1} \tag{54}$$

where $\eta_{t+1} = \eta_{t+1}/\gamma_t$ is Gaussian noise with precision $\sigma_{\tilde{\eta},t+1}^{-2} = \gamma_t^2 \sigma_{\eta,t+1}^{-2}$. It follows that, given any linear strategy, the common posterior about θ in period t is Gaussian with mean $\mathbb{E}[\theta|h_t] = Y_t$ and precision π_t^y, where Y_t and π_t^y are defined recursively by

$$Y_t = \frac{\pi_{t-1}^y}{\pi_t^y}Y_{t-1} + \frac{\sigma_{y,t}^{-2}}{\pi_t^y}y_t + \frac{\sigma_a^{-2}}{\pi_t^y}\tilde{\theta}_t + \frac{\gamma_{t-1}^2 \sigma_{\eta,t}^{-2}}{\pi_t^y}\tilde{y}_t$$

$$\pi_t^y = \pi_{t-1}^y + \sigma_{y,t}^{-2} + \sigma_{a,t}^{-2} + \gamma_{t-1}^2 \sigma_{\eta,t}^{-2},$$

with initial conditions $Y_1 = \mu_0$ and $\pi_1^y = \sigma_0^{-2}$. Similarly, the private posteriors are Gaussian with

mean $\mathbb{E}\left[\theta|\omega_{i,t}\right] = \frac{\pi_t^x}{\pi_t^x+\pi_t^y}X_{i,t} + \frac{\pi_t^y}{\pi_t^x+\pi_t^y}Y_t$, and precision $\pi_t = \pi_t^x + \pi_t^y$, where

$$X_{i,t} = \frac{\pi_{t-1}^x}{\pi_t^x}X_{i,t-1} + \frac{\sigma_{x,t}^{-2}}{\pi_t^x}x_{i,t} \quad \text{and} \quad \pi_t^x = \pi_{t-1}^x + \sigma_{\xi,t}^{-2},$$

with initial conditions $X_{i,1} = x_{i,1}$ and $\pi_1^x = \sigma_{\xi,1}^{-2}$. Now note that $X_{i,t} = \sum_{\tau=1}^s \frac{\sigma_{x,\tau}^{-2}}{\sum_{j=1}^t \sigma_{x,j}^{-2}}x_{i,t}$, which together with (53) gives $\sum_{\tau=1}^s Q_{t,\tau}x_{i\tau} = \gamma_t X_{i,t}$. We conclude that a linear strategy $\mathbf{k}$ maximizes ex-ante utility only if, for all t and all $\omega_{i,t}$, $k_{i,t}\left(\omega_{i,t}\right) = (1-\gamma_t)Y_t + \gamma_t X_{i,t}$, for some γ_t.

Step 2. For more general payoffs V, let $\kappa^*\left(\theta\right) \equiv \arg\max_\kappa V\left(\kappa,\kappa,0,\theta\right)$. A similar argument as in Step 1 ensures that the efficient linear strategy must satisfy

$$k_t\left(\omega_{it}\right) = \kappa^*\left(\gamma_t X_{it} + (1-\gamma_t)Y_t\right)$$

for some γ_t. What then remains is to characterize the optimal $\{\gamma_t\}_{t=1}^N$, which is what we do next.

Step 3. Let $W_{vol} \equiv V_{kk} + 2V_{kK} + V_{KK}$ and $W_{dis} \equiv V_{kk} + V_{\sigma\sigma}$. Ex-ante utility is given by

$$\mathbb{E}u = \mathbb{E}W_{FB}(\theta) + \sum_{t=1}^N \beta^{t-1}\left(\kappa_1^* + \kappa_2^*\right)\left\{\frac{W_{dis}}{2}\left(\gamma_t\right)^2\left(\pi_t^x\right)^{-1} + \frac{W_{vol}}{2}\left(1-\gamma_t\right)^2\left(\pi_t^y\right)^{-1}\right\},$$

where $W_{FB}(\theta) \equiv \sum_{t=1}^N \beta^{t-1}W\left(\kappa^*\left(\theta\right),0,\theta\right)$ is the first-best level of welfare. Noting that the γ's impact only the evolution of public information, and using $W_{vol} < 0$, $W_{dis} < 0$, and $W_{vol}/W_{dis} = 1 - \alpha^*$, we infer that the optimal γ's solve the following problem:

$$\min_{\{\gamma_t\}_{t=1}^N} \sum_{t=1}^N \beta^{t-1}\left\{\gamma_t^2\left(\pi_t^x\right)^{-1} + (1-\alpha^*)\left(1-\gamma_t\right)^2\left(\pi_t^y\right)^{-1}\right\}$$

$$\text{s.t.} \quad \pi_{t+1}^y = \pi_t^y + \Xi_t + \left(\kappa^{*\prime}\right)^2 \sigma_{\eta,t}^{-2}\gamma_t^2 \quad \forall t$$

where $\Xi_t \equiv \sigma_{\varepsilon,t}^{-2} + \sigma_{a,t}^{-2}$ is the exogenous change in the precision of public information. Let $L_t\left(\pi_t^x,\pi_t^y\right)$ denote the associated value function in period t. (The existence of these values functions, and hence of the optimal strategy, follows by construction.) We then have

$$L_t\left(\pi_t^x,\pi_t^y\right) = \min_{\gamma_t}\left\{\gamma_t^2\left(\pi_t^x\right)^{-1} + (1-\alpha^*)\left(1-\gamma_t\right)^2\left(\pi_t^y\right)^{-1} + \beta L_{t+1}\left(\pi_{t+1}^x,\pi_{t+1}^y\right)\right\}$$

$$\text{s.t.} \quad \pi_{t+1}^y = \pi_t^y + \Xi_t + \left(\kappa^{*\prime}\right)^2\sigma_{\eta,t}^{-2}\gamma_t^2$$

The FOC for γ_t gives

$$\gamma_t\left(\pi_t^x\right)^{-1} - (1-\alpha^*)\left(1-\gamma_t\right)\left(\pi_t^y\right)^{-1} + \frac{1}{2}\beta\frac{\partial L_{t+1}}{\partial\pi_{t+1}^y}\frac{\partial\pi_{t+1}^y}{\partial\gamma_t} = 0.$$

The envelope condition for π_{t+1}^y and the law of motion for π_{t+1}^y give

$$\frac{\partial L_t}{\partial \pi_{t+1}^y} = -\left(1 - \alpha^*\right)\left(1 - \gamma_{t+1}\right)^2 \left(\pi_{t+1}^y\right)^{-2} \qquad \text{and} \qquad \frac{\partial \pi_{t+1}^y}{\partial \gamma_t} = 2\left(\kappa^{*\prime}\right)^2 \sigma_{\eta,t+1}^{-2} \gamma_t.$$

It follows that the optimal $\{\gamma_t\}_{t=1}^N$ satisfy

$$\gamma_t^{**} = \frac{\left(1 - \alpha^*\right)\pi_t^x}{\pi_t^y + \left(1 - \alpha^*\right)\pi_t^x - \beta\left(1 - \alpha^*\right)\left(1 - \gamma_{t+1}^{**}\right)^2 \pi_t^x \pi_t^y \left(\pi_{t+1}^y\right)^{-2}\left(\kappa^{*\prime}\right)^2 \sigma_{\eta,t+1}^{-2}}.$$

The SOC guarantees that the denominator is positive, and hence that $\gamma_t > 0$. (However, note that $1 - \gamma_t < 1$ is possible, which means that the sensitivity to public information can change sign.)

Finally, note that the preceding analysis presumes that the value functions L and the optimal strategy exist; this can be shown recursively. ∎

Proof of Proposition 11. Let $\{\gamma_t^{**}\}_{t=1}^N$ be the coefficients that characterize the efficient linear strategy as in Proposition 10 and let $\{\pi_t^x, \pi_t^y\}_{t=1}^N$ be the corresponding precisions of private and public information generated by the efficient linear strategy. The result then follows from letting α_t^{**} be the unique solution to

$$\frac{\left(1 - \alpha_t^{**}\right)\pi_t^x}{\left(1 - \alpha_t^{**}\right)\pi_t^x + \pi_t^y} = \gamma_t^{**}.$$

In fact, it is then and only then that the unique solution to (17) coincides with the strategy obtained in Proposition 10. ∎

Proof of Proposition 12. The existence of a policy that implements the efficient allocation follows from the same argument as in the proof of Proposition 7 for the case of additive measurement error. That the optimal contingency T_{kK} is necessarily higher than in the absence of informational externalities then follows directly from two facts: that $\alpha_t^{**} < \alpha^*$; that the measurement error per se does not affect the optimal T_{kK}. ∎

Proof of Proposition 13. Consider the environments with both exogenous and endogenous Gaussian signals studied in Section 6. The result follows directly from the proof of Proposition 11, where it is shown that, for all periods t, the present-value welfare losses L_t obtained along the efficient linear strategy are decreasing functions of π_t^x and π_t^y, the precisions of private and public information available in the beginning of period t. Putting aside informational externalities, the result can also be established for non-Gaussian signals using a Blackwell-like argument for the planner's problem that characterizes the efficient strategy. ∎

Appendix B: An interpretation of the coefficient α^*

Consider any arbitrary strategy $k : \Omega \to \mathbb{R}$. Given this strategy, define the function $\hat{k} : \Theta \times H \to \mathbb{R}$ by the following rule:

$$\hat{k}(\theta_i, h) \equiv \mathbb{E}[k(\omega_i)|\theta_i, h].$$

Accordingly, let $\hat{K}(h) \equiv \int \hat{k}(\theta, h)\, dh(\theta)$ and $\hat{\sigma}_k^2(h) \equiv \int [\hat{k}(\theta, h) - \hat{K}(h)]^2 dh(\theta)$. The action of any agent i can then be decomposed in three components:

$$k_i = \hat{k}_i + \epsilon + v_i$$

The term $\hat{k}_i \equiv \hat{k}(\theta_i, h)$ captures the variation in individual activity that is "explained" by variation in the underlying fundamentals. The term $\epsilon \equiv (K - \hat{K})$ captures the non-fundamental variation in individual activity that is common across agents; that is, ϵ captures the impact of common noise in information. Finally, the term $v_i \equiv (k - K) - (\hat{k} - \hat{K})$ captures the non-fundamental variation in individual activity that is idiosyncratic to the agent; that is, v_i captures the impact of idiosyncratic noise. (Note that, by construction, $\hat{k}_i$, ϵ and v_i are orthogonal one to the other.) The following result shows that a similar decomposition applies to ex-ante welfare; it then uses this decomposition to relate the coefficient α^* to social aversion over non-fundamental volatility.

Proposition 14. *(i) Given any strategy $k : \Omega \to \mathbb{R}$, ex-ante utility (welfare) is given by*

$$\mathbb{E}u = \mathbb{E}[V(\hat{k}, \hat{K}, \hat{\sigma}_k, \theta)] + \frac{1}{2}W_{vol} \cdot vol + \frac{1}{2}W_{dis} \cdot dis. \tag{55}$$

where $\mathbb{E}[V(\hat{k}, \hat{K}, \hat{\sigma}_k, \theta)]$ measures the welfare contribution of the fundamental component of activity; $vol \equiv \mathrm{Var}(\epsilon) = \mathrm{Var}(K) - \mathrm{Var}(\hat{K})$ and $dis \equiv \mathrm{Var}(v_i) = \mathrm{Var}(k - K) - \mathrm{Var}(\hat{k} - \hat{K})$ measure the non-fundamental volatility and the non-fundamental dispersion of activity; and $W_{vol} \equiv V_{kk} + 2V_{2kK} + V_{KK} < 0$ and $W_{dis} \equiv V_{kk} + V_{\sigma\sigma} < 0$ parameterize the social aversion to these two types of noise.

(ii) The efficient degree of complementarity is negatively related to social aversion to non-fundamental volatility relative to social aversion to non-fundamental dispersion:

$$\alpha^* = 1 - \frac{W_{vol}}{W_{dis}}.$$

Proof. Part (i) follows from taking a second-order Taylor expansion of $V(k_i, K, \sigma_k, \theta)$ around the point $(\hat{k}_i, \hat{K}, \hat{\sigma}_k, \theta)$, aggregating across all states to obtain ex-ante utility, and using the fact that, by construction, the random variables $\hat{k}_i$, ϵ and v_i are orthogonal to one another, with $\mathbb{E}[\epsilon] = \mathbb{E}[v_i] = 0$; a detailed derivation is available upon request. Part (ii) then follows from combining the definition of the coefficients W_{vol} and W_{dis} with the characterization of the coefficient α^* in the proof of Proposition 2. ∎

References

[1] Amador, Manuel, and Pierre-Olivier Weill (2007), "Learning from Private and Public observations of Others' Actions," Stanford University/UCLA mimeo.

[2] Amador, Manuel, and Pierre-Olivier Weill (2008), "Learning from Prices: Public Communication and Welfare," Stanford University/UCLA mimeo.

[3] Angeletos, George-Marios, and Alessandro Pavan (2007), "Efficient Use of Information and Social Value of Information," *Econometrica* 75, 1103-1142.

[4] Angeletos, George-Marios, Guido Lorenzoni, and Alessandro Pavan (2007), "Wall Street and Silicon Valley: A Delicate Interaction," MIT and Northwestern mimeo.

[5] Angeletos, George-Marios, and Jennifer La'O (2008), "Dispersed Information over the Business Cycle: Optimal Fiscal and Monetary Policy," MIT mimeo.

[6] Bacchetta, Philippe, and Eric van Wincoop (2005), "Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?," *American Economic Review* 95.

[7] Barro, Robert (1979), "On the Determination of the Public Debt," *Journal of Political Economy* 87, 940-71.

[8] Banerjee, Abhijit (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics* 107, 797-817.

[9] Chamley, Christophe (2004), *Rational Herds*, Cambridge University Press, Cambridge.

[10] Chari, V.V., and Patrick Kehoe (2003), "Hot Money," *Journal of Political Economy* 111.

[11] Chari, V.V., Lawrence Christiano, and Patrick Kehoe (1994), "Optimal Fiscal Policy in a Business Cycle Model," *Journal of Political Economy* 102, 617-652.

[12] Cremer, Jacques and Richard McLean (1985): "Optimal Selling Strategies under Uncertainty for a Monopolist when Demands are Interdependent," *Econometrica*, 53, 345–361.

[13] Hayek, Friedreich A. (1945), "The Use of Knowledge in Society," *American Economic Review* 35.

[14] Hellwig, Christian (2005), "Heterogeneous Information and the Welfare Effects of Public Information Disclosures," UCLA mimeo.

[15] Jehiel, Philippe and Benny Moldovanu (2001), "Efficient Design with Interdependent Valuations," *Econometrica*, 69, 1237–1259.

[16] Kocherlakota, Narayana (2005), "Advances in Dynamic Optimal Taxation," 9th World Congress of the Econometric Society.

[17] Laffont, Jean-Jacques (1985), "On the Welfare Analysis of Rational Expectations Equilibria with Asymmetric Information," *Econometrica* 53, 1-30.

[18] Lorenzoni, Guido (2008), "News Shocks and Optimal Monetary Policy," MIT mimeo.

[19] Lucas, Robert E., Jr. (1972), "Expectations and the Neutrality of Money," *Journal of Economic Theory* 4, 103-124.

[20] Lucas, Robert E., Jr., and Nancy L. Stokey (1983), "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics* 12, 55-93.

[21] McAfee, Preston, and Philip Reny (1992), "Correlated Information and Mechanism Design," *Econometrica*, 60, 395-421.

[22] McLean, Richard, and Andrew Postlewaite (2002), "Informational Size and Incentive Compatibility", *Econometrica*, 70, 2421-2454.

[23] McLean, Richard, and Andrew Postlewaite (2003), "Informational Size and Incentive Compatibility with Aggregate Uncertainty", *Games and Economic Behavior*, 45, 410-433.

[24] McLean, Richard, and Andrew Postlewaite (2004), "Informational Size and Efficient Auctions", *Review of Economic Studies* 71, 809-827.

[25] Messner, Simon, and Xavier Vives, (2005), "Informational and Economic Efficiency in REE with Asymmetric Information," mimeo, IESE.

[26] Morris, Stephen, and Hyun Song Shin (2002), "The Social Value of Public Information," *American Economic Review* 92, 1521-1534.

[27] Phelps, Edmund (1970), "Introduction: The New Microeconomics in Employment and Inflation Theory," in A.A. Alchian, C.C. Holt et al., *Microeconomic Foundations of Employment and Inflation Theory,* New York: Norton.

[28] Radner, Roy, (1962), "Team Decision Problems," *Annals of Mathematical Statistics* 33, 857-881.

[29] Townsend, Robert M. (1983), "Forecasting the Forecasts of Others," *Journal of Political Economy* 91, 546-588.

[30] Vives, Xavier (1988), "Aggregation of Information in Large Cournot Markets," *Econometrica* 56, 851-876.

[31] Vives, Xavier (1993), "How Fast Do Rational Agents Learn?," *Review of Economic Studies* 60, 329-347.

[32] Vives, Xavier (1997), "Learning from Others: a Welfare Analysis," *Games and Economic Behavior* 20, 177-200.

[33] Vives, Xavier (2008), *Information and Learning in Markets*, Princeton University Press.

[34] Woodford, Michael (2002), "Imperfect Common Knowledge and the Effects of Monetary Policy," in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton University Press.