Endogenous Information Acquisition in Coordination Games

David P. Myatt
Nuffield College, University of Oxford and London Business School
david.myatt@economics.ox.ac.uk

Chris Wallace
Department of Economics and Trinity College, University of Oxford
christopher.wallace@economics.ox.ac.uk

9 March 2011. (Previously: 3 August 2009 and 15 August 2010.)[1]

Abstract. In the context of a "beauty contest" coordination game (in which payoffs depend on the quadratic distance of actions from an unobserved state variable and from the average action) players choose how much costly attention to pay to various informative signals. Each signal has an underlying accuracy (how precisely it identifies the state) and a clarity (how easy it is to understand). The unique linear equilibrium has interesting properties: the signals which receive attention are the clearest available, even if they have poor underlying accuracy; the number of signals observed falls as the complementarity of players' actions rises; and, if actions are more complementary, the information endogenously acquired in equilibrium is more public in nature. The consequences of "rational inattention" constraints on information transmission and processing are also studied. JEL codes: C72, D83.

1. Coordination and Information Acquisition

In many scenarios of social-scientific interest, decision makers seek actions which are matched to both some unknown underlying feature of the world (a "fundamental" motive) and also matched to the actions taken by others (a "coordination" motive). Put somewhat more crudely, when modelled as a game the players wish to do the right thing (match the action to the fundamental) and do it together (coordinate with others' actions). In such scenarios, the participants may welcome any information which helps them to resolve uncertainty about the state of the world and the likely actions of others.

When information is costly an actor must balance the cost of information against its benefit; that benefit depends on the likely action choices of others, and so on the information which others acquire. If others pay close attention to an information source, then their actions will respond strongly to it; if their actions are to be predicted then knowledge of the information source is useful; hence, the coordination motive prompts a player to seek to know what other players know. In a two-stage listening-then-acting environment, this paper asks: to which information sources do players listen, and how do their information acquisition decisions respond to the properties of their environment?

[1]The authors thank colleagues, especially Torun Dewan, seminar participants, the editor Andrea Prat, and particularly two anonymous referees for their helpful comments and constructive suggestions.

As a leading economic example of the applications which motivate this paper, consider an industry in which the demand for a supplier's product depends on the (uncertain) state of the marketplace (perhaps the size of the customer base, or some aggregate-demand factor), on the supplier's own price, and also on the average price amongst the supplier's competitors. In this setting the profit-maximising price is typically increasing in the supplier's expectation of the underlying state of demand (this is the fundamental motive) and the expected industry-wide price level (the coordination motive). To improve decision making, a supplier may wish to engage in a (presumably costly) survey of market conditions. In the course of such market research, a supplier might examine different market segments, where a market segment could correspond to a geographic region, a point in time, or a particular product characteristic. Naturally, if competitors are devoting considerable time to research a particular market segment, then a supplier might wish to research this segment too; doing so can help predict those competitors' prices.

Motivating examples also emerge from political science and sociology. For instance, Dewan and Myatt (2008, 2011) viewed a political party as a group of activist members, where each member chooses a policy to advocate. Activists wish to do the right thing, by backing the right policy for the party, but they also value party unity. Such party members may pay costly attention to (possibly competing) party leaders who act as information sources; such leaders (if heeded) may help activists to develop a shared understanding of the party's direction. Similar logic can apply to a religious organisation: each member of a congregation may wish to live in accordance with some underlying (but uncertain) ideal spiritual values, while living compatibly with others; a case, perhaps, of literally singing from the same hymn sheet. Again, the motives are to do the right thing and to do it together. Learning about these things may involve dividing attention between information sources, such as the sermons of preachers or designated scriptures. Finally, turning back to economics another application of interest is the "shared knowledge" notion of corporate culture (Crémer, 1990, 1993). Using a team-theoretical (Marschak and Radner, 1972) framework, Crémer (1990, p. 55) noted that organisation members "will make observations relevant to the decisions that they have to take" and he went on to "study the trade-off faced by a firm between accumulating a diversified knowledge about the environment and providing common ground for decisions." This hints that different properties of information sources may be useful for the twin motives faced by agents who wish to coordinate their actions effectively in an uncertain world.

Throughout a recently developed literature the key features of the applications described here have been nicely captured by a tractable class of quadratic-payoff "beauty contest" games. In such a game the payoffs depend on the proximity of players' actions to an underlying state variable and to an aggregate measure of all actions. Players may have different information about the state variable, and so differences of opinion may frustrate coordination. The "beauty contest" terminology is drawn from a well-known parable told by Keynes (1936, Chapter 12); he described newspaper-based competitions whose entrants were invited to choose the prettiest faces from a set of photographs, but where it was optimal to nominate the most popular faces.

Beauty-contest models have received close attention following the contribution of Morris and Shin (2002). Such games have been applied to investment games (Angeletos and Pavan, 2004), to monopolistic competition (Hellwig, 2005), to financial markets (Allen, Morris, and Shin, 2006), to a range of other economic problems (Angeletos and Pavan, 2007), and to political leadership (Dewan and Myatt, 2008, 2011); many other papers report variants of the beauty-contest specification. Such games are also closely related to the macroeconomic island-economy parable (Lucas, 1973; Phelps, 1970) so long as players are interpreted as the island sectors and their actions are market-clearing prices (Amato, Morris, and Shin, 2002; Morris and Shin, 2005; Myatt and Wallace, 2008).

In most beauty-contest models information is exogenous. In the model of Morris and Shin (2002) players have two information sources: one is private (an independent signal realisation for each player) whereas the other is public (a common signal realisation). This paper moves away from the public-private distinction, allows multiple information sources, and considers endogenous information acquisition. Natural questions arise. Given the availability of multiple informative but costly signals, how carefully do players choose to listen to each? Do all information sources receive positive attention? How do equilibrium information-acquisition strategies (and, indeed, action choices) respond to the exogenous parameters of the model? These questions entail a step outside the established public-private taxonomy of signals, because the (endogenous) attention devoted to an information source determines how "public" it is.

More specifically, here players are granted costly access to a collection of information sources. Each source provides an informative signal with some source-specific "sender" noise; this sender noise determines the signal's underlying accuracy. A player then observes this signal with some additional player-specific "receiver" noise. The receiver noise, which determines the signal's clarity, is endogenous: if a player listens with greater care (and so at greater cost) then the receiver noise is reduced. The aforementioned industry-supply example helps to illustrate the components of signal noise. In a market-research context, perfect observation of a market segment (a region, time period, product characteristic, and so on) does not reveal completely the overall state of the market; each segment is (presumably) subject to its own idiosyncrasies. Thus, "sender noise" is the difference between demand in a segment and overall demand. An investigating supplier can conduct a costly survey of consumers within a market segment. The "receiver noise" is then the sampling error; if the survey is a random sample of a market segment then the variance is inversely proportional to survey size. More generally, sender noise is present at the origin of an information source, whereas receiver noise is error either in observation or understanding as players attempt to acquire and assimilate the data.

The players' information-acquisition decisions endogenously determine the correlation of their observations. These observations become highly correlated (a signal becomes very "public") if and only if all players pay very careful attention to the corresponding information source. In the market-research setting, if all suppliers saturate the same market segment with intensive surveys then they will obtain a common picture of that segment. More generally, the "publicity" of a signal depends on the mix of sender noise and receiver noise, with the latter endogenously determined.

Allowing players to choose how carefully to observe the information sources (instead of choosing whether or not to acquire a signal) has implications for the information-acquisition equilibrium: it is unique, and so comparative-static exercises are permitted. Robust messages emerge: only some signals receive attention; these are the clearest signals available, even if they have poor underlying accuracy; the number of such signals shrinks as the complementarity of actions rises; and, if actions become more complementary then the information endogenously acquired becomes more public in nature.

Turning back to the literature, most recent related research (amongst contributions that focus on beauty-contest games) has not considered endogenous information acquisition. One exception within political science is the model of leadership by Dewan and Myatt (2008), in which followers divide their attention between different leaders; leaders' speeches help their followers to learn about the world and to coordinate with each other. A notable exception within economics is a recent article by Hellwig and Veldkamp (2009). Their intuition that complementarity of action choice imposes complementarity upon information choice ("if an agent wants to do what others do, they want to know what others know") applies here. It suggests that there is scope for multiple equilibria; indeed, Hellwig and Veldkamp (2009, p. 224) argued that "[...] information choice imposes an additional requirement for equilibrium uniqueness: the information agents choose to acquire must also be private." The idea is that the acquisition of a public signal does two things: it informs a player about the underlying state directly and also about the likely actions of others. This second effect is present if and only if others acquire the signal too, and this naturally leads to multiple equilibria. When a signal is private (so that, conditional on the underlying state, realisations are independent) then it does not directly inform a player about others' likely moves. This removes a key ingredient of multiple equilibria.

This paper shows that the full privacy of signals is not a requirement for uniqueness. The results of Hellwig and Veldkamp (2009) depend upon the way in which players obtain their first bit of a signal. In this paper a player does not simply choose whether to obtain a particular (perhaps small) signal with a pre-determined publicity (or correlation); instead, a player chooses how much costly attention to pay to an information source. The first bit of a signal acquired (a situation in which a player pays relatively scant attention) is dominated by receiver noise. This ensures that the signal realisation is relatively uncorrelated with the signals received by others, and so is relatively private. Roughly speaking, this smooths out the first step of the information acquisition process and eliminates multiple equilibria, even though the informative signals actually acquired in equilibrium may be relatively public in nature.

Other research without a direct beauty-contest focus has allowed for endogenous information acquisition. The "rational inattention" literature associated with Sims (1998, 2003, 2005, 2006) has considered a world in which agents are free to construct informative signals, but face a constraint: there is a limit to the quantity of information which can be transmitted to them and absorbed by them; devoting attention to learning about one variable precludes paying attention to another. For example, in a recent paper Maćkowiak and Wiederholt (2009) considered the balance of attention between aggregate and idiosyncratic shocks. As Sims (2010) explained, such models "introduce the idea that people's abilities to translate external data into action are constrained by a finite Shannon 'capacity' to process information." This notion of capacity comes from information theory (Cover and Thomas, 2006; MacKay, 2003); when messages are appropriately coded, it is related to the minimal bandwidth required for successful communication.

With the rational-inattention approach in mind, and returning to the market-research setting, two stages of research can be envisaged. Firstly, a supplier must acquire data; the associated cost might be proportional to the sample size of a survey. Secondly, this data must be transmitted to and absorbed by the supplier's management. The limits to this second step correspond to the aforementioned Shannon capacity constraint.

The second information-transmission step is readily incorporated; it yields a particular cost function. However, this function is not convex. This is because the information content (and so the necessary bandwidth) arising from additional data is decreasing in the stock of existing information. So, whereas the cost of acquiring survey data may linearly (and so convexly) increase with the sample size, the cost of passing on the results rises only concavely. The uniqueness result of this paper uses the convexity of the cost function; when the costs of information acquisition stem from the constraints which feature in the rational-inattention literature then there can be multiple equilibria; an example is readily found. Nevertheless in some cases (the industry-supply example is one) uniqueness results can be maintained. Furthermore, the pattern of attention is predictable: players listen to the signals with the best accuracy, rather than those with the best clarity.

Comparing different approaches to information acquisition, three cases can be identified: firstly, players choose whether or not to pay to receive a signal (e.g., Hellwig and Veldkamp, 2009); secondly, they divide their time continuously between sampling different information sources (e.g., Dewan and Myatt, 2008); and, thirdly, they face information-processing constraints (e.g., Sims, 2003). This paper links these three different approaches by showing how they correspond to different cost-function specifications.

Turning to the structure of the paper, Sections 2–4 describe the model and the unique equilibrium in which actions respond linearly to signals. Sections 5–7 show how information acquisition, actions, and the publicity of informative signals respond to the coordination motive and to other parameters. Sections 8–9 relate the model to the rational-inattention literature, and consider the impact of imposing a constraint upon information transmission. Finally, Section 10 relates the results to those of the existing literature.

2. A Beauty-Contest Coordination Game

The model considered here is a quadratic-payoff "beauty contest" game in which players' payoffs depend upon the proximity of their actions to an unobserved underlying state variable and to the average action taken by all players. The twist is that the information sources upon which players condition their actions are both costly and endogenous.

More formally, a simultaneous-move game is played by a unit mass of players indexed by $\ell \in [0, 1]$. An individual player's move consists of the following three steps.

(1) A player chooses an information-acquisition policy $z_\ell \in \mathcal{R}^n_+$. The interpretation is that there are n information sources, and the element $z_{i\ell}$ of the vector z_ℓ is the amount of costly attention which player ℓ pays to the ith informative source.
(2) After this information-acquisition choice, the player observes a vector of n signals $x_\ell \in \mathcal{R}^n$ which inform the player about some unobserved state variable θ, where the precisions of these signals depend upon the earlier choice of z_ℓ.
(3) Finally, a player takes a real-valued signal-contingent action $a_\ell \in \mathcal{R}$.

A player ℓ's pure strategy is a pair $\{z_\ell, A_\ell(\cdot)\}$ where z_ℓ is the information-acquisition component and the function $A_\ell(\cdot) : \mathcal{R}^n \mapsto \mathcal{R}$ specifies the action $a_\ell = A_\ell(x_\ell)$ which is to be taken following the observation of the n signal realisations $x_\ell \in \mathcal{R}^n$.

A player's payoff depends on the proximity of the player's action a_ℓ to the underlying state variable θ, the action's proximity to the average action $\bar{a} \equiv \int_0^1 a_l dl$, and the player's information acquisition z_ℓ. Assembling these three elements, a player's payoff is

$$u_\ell = \bar{u} - (1-\gamma)(a_\ell - \theta)^2 - \gamma(a_\ell - \bar{a})^2 - C(z_\ell). \tag{1}$$

The parameter $\gamma \in (-1, 1)$ determines a player's concern for aligning with others (the coordination motive) relative to matching the state variable (the player's fundamental motive). If $\gamma = 0$ then coordination is irrelevant. The model allows for $\gamma < 0$, in which case a player wishes to differ from others. Nevertheless, the restriction $|\gamma| < 1$ is imposed; if $|\gamma| > 1$ then a strategy-revision process driven by best replies is explosive, and some of the analysis reported throughout the paper fails.[2] The final component of (1) is the cost of acquiring, transmitting, and processing information. Throughout most of the paper the cost function $C(z_\ell)$ is assumed to be increasing, convex, and differentiable. However, when information-processing constraints are explicitly incorporated (in Section 8) a different formulation for information-acquisition costs is considered.

Before moving on to describe the information sources available to players, the general specification of (1) is related to the motivating example from the introduction to the paper:

[2] This is easy to see in a complete-information model. If θ is known then a player's unique best reply to an average action $\bar{a}$ taken by others is $a_\ell = (1-\gamma)\theta + \gamma\bar{a}$, and the unique Nash equilibrium is for all players to choose $a_\ell = \theta$. However, consider a strategy-revision process comprising myopic best replies. Specifically, begin with a strategy profile in which the average action is $a^{(0)} \neq \theta$. If all players adopt a myopic best reply to this then they will all take the action $a^{(1)}$ satisfying $a^{(1)} - \theta = \gamma(a^{(0)} - \theta)$. Repeating this step k times readily yields $a^{(k)} - \theta = \gamma^k(a^{(0)} - \theta)$. This process explodes if $|\gamma| > 1$.

an application in which a supplier's demand depends on the state of the marketplace and the average price amongst competitors. For this application, a_ℓ is the price set by supplier ℓ, $\bar{a}$ is the industry-wide average price amongst others, and θ is a demand-shock parameter. A simple linear specification for the demand q_ℓ for ℓ's product is

$$q_\ell = (2-\beta)\theta - a_\ell + \beta\bar{a}, \tag{2}$$

for some positive parameter $\beta < 1$, where the coefficient $(2-\beta)$ on θ is a convenient (for algebraic purposes) rescaling of the demand-shift parameter. Setting costs to zero (with no loss of insight) it is straightforward to confirm that a supplier's profit satisfies

$$a_\ell((2-\beta)\theta - a_\ell + \beta\bar{a}) = -\left(1-\frac{\beta}{2}\right)(a_\ell-\theta)^2 - \frac{\beta(a_\ell-\bar{a})^2}{2} + \left(1-\frac{\beta}{2}\right)\theta^2 + \frac{\beta\bar{a}^2}{2}. \tag{3}$$

Notice that the final two terms are independent of supplier ℓ's price a_ℓ and so are strategically irrelevant; they may be safely neglected, leaving only the first two quadratic-loss terms. It is easy to see that the remaining components of a supplier's profit in (3) combine to take the form of the payoffs in (1); to do this, simply define $\gamma = \beta/2$. Thus, in this context γ indexes the importance of others' prices relative to its own price on a supplier's demand; if others' prices are irrelevant, so that $\beta = \gamma = 0$, then the coordination motive absent. Note also that a restriction is endogenously imposed upon the parameter γ. As competitors' prices have less impact upon demand than a supplier's own price (so $\beta < 1$), it must be that $\gamma < \frac{1}{2}$, a point returned to in later sections.

Before moving on, two technical issues are briefly discussed. Firstly, the player set is a unit mass and so each individual is negligible. In the context of the example above, the continuum-of-players specification implies that each price-setting supplier is best thought of as a monopolistic competitor rather than an oligopolist. The unit-mass assumption serves mainly to simplify exposition, but is not crucial to the results. Appropriately modified, many messages emerging from the paper carry over to a world with a finite number of players.[3]

Secondly, a player's payoff depends on the average action $\bar{a}$ taken across all players. (Equivalently, given the unit-mass-of-players assumption, this is the average taken across all other players apart from player ℓ.) Of course, this average is not always well-defined.[4] However, for the class of equilibria considered later in the paper (specifically, those in which the action chosen by a player is a linear function of the informative signals observed) the average remains well defined both in equilibrium and following a single-player deviation. Furthermore, the specification of the game may be completed by placing payoffs on the extended real line and setting $u_\ell = -\infty$ whenever $\bar{a}$ does not exist.

[3]The appendix to Myatt and Wallace (2008) demonstrates the changes needed to consider an L-player version of beauty-contest games of the kind considered here. That paper does not include endogenous information acquisition, but otherwise uses the same informational environment and structure studied here.

[4]For example, consider a strategy profile in which players choose actions which form a Cauchy distribution across the player set. The mean of the Cauchy does not exist, and so $\bar{a}$ is not well-defined.

3. Information Sources

Players begin with no knowledge of the underlying state; they share an improper prior over θ. Eliminating the prior serves solely to simplify the statement of the results; and no insight is lost by doing so. Indeed, a common prior can be accommodated easily by using one of the n informative signals to reflect prior beliefs, and making it costless to observe perfectly that signal. Furthermore, an explicit prior is specified in Section 8, when "rational inattention" constraints to players' information processing are considered.

Turning to the n information sources, the ith signal observed by player ℓ satisfies

$$x_{i\ell} = \theta + \eta_i + \varepsilon_{i\ell}, \quad \text{where} \quad \eta_i \sim N(0, \kappa_i^2) \quad \text{and} \quad \varepsilon_{i\ell} \sim N\left(0, \frac{\xi_i^2}{z_{i\ell}}\right), \tag{4}$$

and where the various noise terms are all independently distributed.

The general interpretation of (4) is that each information source has associated with it some "sender" noise η_i which reflects the quality or accuracy of an underlying signal $\bar{x}_i \equiv \theta + \eta_i$; the accuracy is indexed by the precision $1/\kappa_i^2$. A player ℓ who chooses to pay attention to the information source i does so imperfectly, owing to "receiver" noise, by observing $x_i = \bar{x}_i + \varepsilon_{i\ell}$. The receiver noise reflects the clarity with which the information is imparted, indexed by $1/\xi_i^2$, and the attention $z_{i\ell}$ that player ℓ pays to source i, so that the overall clarity of the observation is determined by the precision $z_{i\ell}/\xi_i^2$. The observation precision (or clarity) linearly increases with the choice variable $z_{i\ell}$, and so a player's information acquisition can be interpreted conveniently as a sample size. Furthermore, the choice $z_{i\ell} = 0$ is straightforwardly interpreted as the decision to ignore the ith information source completely (equivalently, the realisation $x_{i\ell}$ is pure noise in this case).

The illustrative example discussed in previous sections yields a specific interpretation of (4). Consider again a supplier maximising the profit (3) associated with the demand function (2) specified in Section 2. Naturally, the supplier may investigate demand conditions. Imagine, then, that it conducts market research in n different segments of the marketplace. A market segment $i \in \{1, \ldots, n\}$ could be thought of as a geographic region, as a particular class of consumers, as a period of time, or even, more broadly, as the opinions of consumers about a particular product characteristic. The underlying signal that can be obtained from a segment is then equal to the market-wide demand state θ plus some segment-specific shock η_i; this shock (the "sender noise" in this scenario) may be related, for instance, to the idiosyncrasies of a geographic region. The best that a survey can do is to identify perfectly the segment-specific demand conditions $\bar{x}_i = \theta + \eta_i$. However, any survey is subject to sampling error; this is the "receiver noise" $\varepsilon_{i\ell}$. If a supplier obtains a random sample of a market segment then the variance of the sampling error is inversely proportional to the sample size. Thus, the information-acquisition decision $z_{i\ell}$ can be thought of as the number of consumers in market segment i interviewed by a market researcher from supplier ℓ. Furthermore, if the supplier faces a price-per-interview then a natural specification for costs is the linear form $C(z_\ell) = \text{constant} \times \sum_{i=1}^{n} z_{i\ell}$.

Conditional on θ, information sources are independent, but players' observations of each source are correlated: for two players ℓ and ℓ', $\mathrm{cov}[x_{i\ell}, x_{i\ell'} \,|\, \theta] = \kappa_i^2$, and so observations move together unless the underlying signal $\bar{x}_i$ has perfect precision. Furthermore, the correlation of players' observations depends straightforwardly on the mix of sender noise and receiver noise. More formally, the model specification is equivalent to one in which

$$x_{i\ell} \,|\, \theta \sim N(\theta, \sigma_{i\ell}^2) \quad \text{and} \quad \mathrm{cov}[\, x_{i\ell}, x_{i\ell'} \,|\, \theta \,] = \rho_{i\ell\ell'} \sigma_{i\ell} \sigma_{i\ell'}, \tag{5}$$

for all $\ell' \neq \ell$ and for all i. This emerges from the specification (4) via the transformations

$$\sigma_{i\ell}^2 = \kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}} \quad \text{and} \quad \rho_{i\ell\ell'} = \kappa_i^2 \left[\left(\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}} \right) \left(\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell'}} \right) \right]^{-\frac{1}{2}}. \tag{6}$$

Paying more attention to an information source i (by increasing $z_{i\ell}$) not only reduces the overall variance $\sigma_{i\ell}^2$ of that signal (or, equivalently, increases the precision), but also makes it more correlated with others' observations of i (the correlation coefficient $\rho_{i\ell\ell'}$ increases).

The specification (4) and transformations (6) can be related to established models in the literature. Setting $z_{i\ell} = z_i$ for all ℓ for expositional simplicity, the correlation of players' observations of an information source is $\rho_i = \kappa_i^2/[\kappa_i^2 + (\xi_i^2/z_i)]$. The case $\rho_i = 0$, so that observations are conditionally uncorrelated, is obtained when $\kappa_i^2 = 0$, and corresponds to the "private" signal from the two-source world of Morris and Shin (2002). In contrast, the case $\rho_i = 1$, obtained in the limit as $z_i \to \infty$ or by setting $\xi_i^2 = 0$, so that players' observations coincide, corresponds to the "public" signal of Morris and Shin (2002).

For general values of κ_i^2, ξ_i^2, and z_i a signal's correlation satisfies $0 < \rho_i < 1$ so the signal is neither purely private nor purely public. As noted above, the correlation coefficient (and hence publicity of a signal) is both endogenous and also directly linked to the precision of a signal. In particular, the correlation coefficient vanishes as the attention paid to an information source shrinks to zero. What this means is that as a player begins to acquire information from a source, so that z_i moves up from zero, the signal is initially private in nature, and only becomes more public as increasing attention is devoted to it.[5]

Two further technical issues are mentioned before concluding this section. Firstly, a signal's distribution is not properly specified when a player chooses $z_{i\ell} = 0$. However, this does not cause any particular problems since, as noted above, choosing $z_{i\ell} = 0$ is equivalent to ignoring an information source. Secondly, for $\xi_i^2 > 0$ obtaining a perfectly public signal is impossible. However, this can be resolved by extending the choice of information acquisition to include $z_{i\ell} = \infty$, so long as the cost $\lim_{z_{i\ell} \to \infty} C(z_\ell)$ is well-defined.

[5]This contrasts with the specifications used by Hellwig and Veldkamp (2009). Their players either acquire a signal or do not. This is equivalent to restricting a player's choice of $z_{i\ell}$ to take only two values. They also considered a specification in which a player's information-acquisition decision is continuous. However, that specification insists that the correlation coefficient does not change with the information acquired. In the model proposed here, this is equivalent to assuming that a signal's correlation coefficient remains bounded away from zero even when hardly any attention is paid to it. As Section 10 explains, it is this feature which is responsible for the presence of multiple linear equilibria in their model.

4. EQUILIBRIUM

A player's strategy $\{z_\ell, A_\ell(\cdot)\}$ specifies the action $A_\ell(x_\ell)$ taken in response to each possible signal realisation x_ℓ. There are good reasons to follow the established literature by focusing on strategies in which a player's action $A_\ell(x_\ell)$ is a linear function of the signal realisations. To see why, suppose that all others use a strategy $\{z, A(\cdot)\}$. Differentiating the quadratic objective function confirms that player ℓ's best-reply action is

$$A_\ell(x_\ell) = (1-\gamma)\,\mathrm{E}[\theta \,|\, x_\ell] + \gamma\,\mathrm{E}[A(x_{\ell'}) \,|\, x_\ell], \tag{7}$$

which is a weighted average of the player's expectations of the state variable and of the average action.[6] Given the normality assumptions, the first expectation is linear in x_ℓ. If $A(\cdot)$ is linear, then the second expectation is also linear in x_ℓ. Hence, if other players use a linear strategy then the unique best reply is linear. Furthermore, relatively mild restrictions on the class of strategies used by players ensure that equilibrium strategies are linear. One such restriction is to consider non-linear strategies that are nonetheless bounded by linear strategies. A strategy $A(\cdot)$ satisfies this restriction if there is a linear function $\bar{A}(\cdot)$ such that $|A(x_\ell) - \bar{A}(x_\ell)|$ remains bounded for all x_ℓ. If an equilibrium strategy satisfies this restriction, then it must itself be linear (Dewan and Myatt, 2008).[7]

A strategy is linear if there are weights $w_\ell \in \mathcal{R}^n$ such that $A_\ell(x_\ell) = \sum_{i=1}^n w_{i\ell} x_{i\ell}$. Given linearity, a player's strategy takes the form $\{z_\ell, w_\ell\}$, and it is straightforward to confirm that in the context of an equilibrium strategy $\sum_{i=1}^n w_{i\ell} = 1$, so that a player's action is a weighted average of the signals received, and $w_{i\ell}$ is the influence of the ith information source. (This claim is verified formally in Appendix A.[8]) Given that all other players employ a strategy $\{z, w\}$ then the expected payoff of a player ℓ choosing $\{z_\ell, w_\ell\}$ is

$$\mathrm{E}[u_\ell] = \bar{u} - \underbrace{\sum\nolimits_{i=1}^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right]}_{L^\star(w_\ell, z_\ell)} - \underbrace{\gamma \sum\nolimits_{i=1}^n (w_{i\ell} - w_i)^2 \kappa_i^2}_{L^\dagger(w_\ell, w)} - C(z_\ell). \tag{8}$$

---

[6] Note that a player's forecast of the average action is equivalent to the forecast of the action of an arbitrary player $\ell' \neq \ell$. The average action is $\bar{a} = \int_0^1 a_{\ell'}\, d\ell'$, so taking expectations $\mathrm{E}[\bar{a} \,|\, x_\ell] = \int_0^1 \mathrm{E}[a_{\ell'} \,|\, x_\ell]\, d\ell'$. The expectation in the integrand does not depend on the particular label ℓ', and so $\mathrm{E}[\bar{a} \,|\, x_\ell] = \mathrm{E}[a_{\ell'} \,|\, x_\ell]$. Of course, $a_{\ell'} = A(x_{\ell'})$, which upon substitution yields the final term of (7).

[7] Morris and Shin (2002) claimed that the linear equilibrium of a beauty-contest game is unique. Angeletos and Pavan (2007, fn. 5) observed that their logic is not watertight. Dewan and Myatt (2008) proved uniqueness within the class of strategies which (as described here) do not stray too far from linearity; their approach could be extended to strategies which do not diverge from a finite-term polynomial strategy. A second approach is to consider a related game in which state, signal, and action spaces are bounded, and show that the unique equilibrium converges to the unique linear equilibrium of an unbounded game as the various bounds are removed (Calvó-Armengol, de Martí Beltran, and Prat, 2009). Finally, arguments from the classic study of team-decision problems (Radner, 1962) can be exploited: for an appropriately specified finite-player version of the game considered here, and given the introduction of an appropriate proper and normal prior, the unique symmetric strategy profile which maximises the ex ante expected payoff of a randomly chosen player is the unique linear equilibrium. Contrary to some claims within the literature, it seems that a full uniqueness proof is unavailable. This is because there are some strategy profiles for which payoffs are not defined; footnote 4 mentions Cauchy-distributed actions as an example.

[8] Appendix A also contains various calculations, such as the derivation of (8), omitted from the main text.

Given that others play linearly (and, following the discussion in footnote 7, there is little if any loss of generality by supposing that they do) a player's best reply is to choose a pair of vectors $\{z_\ell, w_\ell\}$ to maximise (8) subject to the constraint $\sum_{i=1}^n w_{i\ell} = 1$. An inspection confirms that (8) is strictly concave, and a player's best reply is unique.

Before characterising a player's best reply and the unique symmetric linear equilibrium to the beauty-contest game, the components of (8) are discussed.

Consider each element of $L^\star(w_\ell, z_\ell)$. This summation is the quadratic loss experienced by a player when all players use the same weights on their signals. By placing weight on the ith information source a player is exposed to both the sender noise $\eta_{i\ell}$ (with variance κ_i^2) and receiver noise ε_i (with variance $\xi_i^2/z_{i\ell}$). The receiver noise, which is idiosyncratic to player ℓ, pushes the player's action away from both the state variable θ and also the average action $\bar{a}$. Given that all players use the same weights, the sender noise pushes the player's action away from the state variable θ but does not push it away from the actions of others; the reason is that η_i is a common shock to all players, and so (as long as they use a common linear strategy) it has no bearing on the coordination-motive component of a player's payoff. For this reason, the variance term κ_i^2 is a multiple of the coefficient $(1-\gamma)$.

Next, consider each element of $L^\dagger(w_\ell, w)$. This second summation is the quadratic loss experienced by a player owing to the use of a different strategy from other players. Each loss here arises because of the various sender noise terms η_i. If $w_{i\ell} = w_i$ then player ℓ's action reacts to the shock η_i in the same way as other players, and so a_ℓ and $\bar{a}$ do not move apart. However, if $w_{i\ell} \neq w_i$ then owing to the different reactions the receiver noise can move a player away from others. Since this reflects the desire to coordinate (or, indeed, the desire to differ if $\gamma < 0$) then $L^\dagger(w_\ell, w)$ attracts the coefficient γ.

Notice that $L^\dagger(w_\ell, w)$ disappears when players use the same strategy. Furthermore, beginning from a symmetric strategy profile, changes in a player's strategy have no first-order effect on $L^\dagger(w_\ell, w)$, and so when considering a local deviation a player needs only to consider the effect of that deviation on $L^\star(w_\ell, z_\ell)$ and $C(z_\ell)$. $\mathrm{E}[u_\ell]$ is concave in w_ℓ and z_ℓ, and so consideration of local deviations is all that is needed. This means that in a symmetric equilibrium each player acts as though minimising $L^\star(w_\ell, z_\ell) + C(z_\ell)$. These observations form a useful lemma.

Lemma 1. *A strategy $\{z, w\}$ forms a symmetric equilibrium if and only if it solves*

$$\min \sum_{i=1}^n w_i^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_i}\right] + C(z_\ell) \quad \textit{subject to} \quad \sum_{i=1}^n w_i = 1. \tag{9}$$

This lemma relies upon the maintained assumption that $C(\cdot)$ is convex. Such convexity ensures that the first-order conditions from maximisation of $\mathrm{E}[u_\ell]$ in the context of a symmetric equilibrium successfully solve (9). However, if $C(\cdot)$ is not convex then an equilibrium strategy $\{w, z\}$ can only be guaranteed to generate a local minimum of

$L^\star(w_\ell, z_\ell) + C(z_\ell)$. Any global minimiser of $L^\star(w_\ell, z_\ell) + C(z_\ell)$ will necessarily generate an equilibrium (a moment's inspection confirms that the addition of $L^\dagger(w_\ell, w)$ helps to dissuade a player from deviating from a symmetric profile). However, a local (but not global) minimiser of $L^\star(w_\ell, z_\ell) + C(z_\ell)$ might also generate an equilibrium.

The solution to this minimisation problem (9) generates Proposition 1. (The proofs of this result and all the other propositions are collected together in Appendix A.)

Proposition 1 (Basic Equilibrium Characterisation)**.** *In the unique linear symmetric equilibrium, the influence w_i of the ith signal and the attention z_i paid to it satisfy*

$$w_i = \frac{\hat{\psi}_i}{\sum_{j=1}^n \hat{\psi}_j} \quad \textit{and} \quad z_i = \frac{\xi_i w_i}{\sqrt{C_i'(z)}}, \quad \textit{with} \quad \hat{\psi}_i = \frac{1}{(1-\gamma)\kappa_i^2 + \xi_i^2/z_i}, \tag{10}$$

and where $\hat{\psi}_i = 0$ for any information source which is ignored (so that $z_i = w_i = 0$).

The weight attached to a particular signal is large when that signal is listened to carefully: w_i moves together with z_i. Moreover, signals have more weight attached to them whenever they are clearer or more accurate; that is, when ξ_i^2/z_i and κ_i^2 fall.

Putting aside the information-acquisition decisions for a moment, the equilibrium influence of an information source (this is determined by $\hat{\psi}_i$) depends less strongly on a signal's underlying accuracy than on its clarity whenever players value coordination (so that $\gamma > 0$). Indeed, if only coordination matters (so that γ is close to one) then a signal's influence is proportional to its clarity. This is natural: changing a signal's underlying accuracy affects only the ability of players to hit the truth (which matters to the extent that hitting the truth matters; that is, $1-\gamma$) whereas enhancing a signal's clarity helps players both to coordinate and also to hit the true value of θ.

Another perspective is provided by considering the informativeness of each source of information and the degree to which different signal realisations coincide. Drawing upon the discussion in Section 3, the variance of the ith signal and the correlation coefficient between two realisations $x_{i\ell}$ and $x_{i\ell'}$, both conditioned on the underlying state θ, are

$$\sigma_i^2 = \kappa_i^2 + \frac{\xi_i^2}{z_i} \quad \text{and} \quad \rho_i = \frac{\kappa_i^2}{\kappa_i^2 + (\xi_i^2/z_i)}. \tag{11}$$

The precision $\psi_i \equiv 1/\sigma_i^2$ measures how the ith signal informs a player about the fundamental. The correlation coefficient ρ_i determines how public that signal is. Comparing two signals that receive positive attention in equilibrium, the influence on players' action choices of signal i relative to signal j is given by

$$\frac{w_i}{w_j} = \frac{\sigma_j^2}{\sigma_i^2}\frac{1-\gamma\rho_j}{1-\gamma\rho_i}. \tag{12}$$

Thus the relative influence is the product of two terms. The first ratio is the precision of the ith signal relative to the jth. Notice that this is all that matters when $\gamma = 0$. The second ratio measures the relative publicity of the signals; when $\gamma > 0$, so that coordination is desirable, this drives influence toward the signal with the higher correlation coefficient. Signals that are more public (more highly correlated) are more useful for the players' coordination motive.[9] When $\gamma < 0$ (coordination is undesirable) the reverse is true.

The next three sections of the paper examine the properties of the equilibrium described in Proposition 1 via comparative-static exercises. Firstly, Section 5 analyses how information acquisition varies with the exogenous parameters, in particular the coordination preferences of the players (γ). Secondly, Section 6 relates the (endogenous) publicity of information sources to the nature of comparative-static predictions. Thirdly, Section 7 examines how equilibrium actions and beliefs vary with the coordination motive.

5. Information Acquisition

The main focus of this paper is on the introduction of endogenous information acquisition to an otherwise-standard beauty contest, and so the determinants of z (the information-acquisition policy) are now considered. Taking (10) and substituting yields, for $z_i > 0$,

$$z_i = \frac{\xi_i(K_i - \xi_i)}{(1-\gamma)\kappa_i^2} \quad \text{where} \quad K_i \equiv \frac{1}{\sqrt{\partial C(z)/\partial z_i} \sum_{j=1}^n \hat{\psi}_j}. \tag{13}$$

Treating K_i as a constant for the moment, (13) suggests that the attention paid to an information source is increasing in the accuracy (that is, the precision $1/\kappa_i^2$) of the underlying signal $\bar{x}_i = \theta + \eta_i$. Put more crudely, players listen more carefully to an information source whenever its provider has more to say. However, notice (again treating K_i as a constant for the moment) that z_i is potentially non-monotonic in the clarity (determined by $1/\xi_i^2$) with which the information is communicated. This is rather natural: ξ_i^2 is effectively the price of obtaining a noisy observation of $\bar{x}_i$ with precision z_i/ξ_i^2, and so z_i is a player's expenditure on that information source. This expenditure is increasing and then decreasing in the price charged. A final observation is that (13) applies only so long as $\xi_i < K_i$. When ξ_i exceeds K_i then $z_i = 0$. This indicates that an information source is likely to receive attention only if it is communicated with sufficient clarity.

This discussion of (13) treats K_i as a constant; but of course it is not. Nevertheless, with a little more structure the suggested comparative-static properties do hold. To proceed further it proves useful to examine a particular form for the cost of information acquisition. Consider a world in which z_i is the time spent listening to signal i; picking up on the market-research story from the industry-supply example featuring in the introduction,

[9] In related work Myatt and Wallace (2008) called the term $\beta_i \equiv 1/(1-\gamma\rho_i)$ "publicity". Thus $w_i \propto \psi_i\beta_i$. The focus there is on the macroeconomic island-economy parable and follows closely in the spirit of Morris and Shin (2002). As a result the restriction $\gamma \geq 0$ holds and so β_i is increasing in ρ_i. Thus the notion of publicity conveniently captures the correlation of signals across 'islands'. The emphasis is on macroeconomic performance in the presence of informative announcements about θ by a social planner (for instance, a central bank), treated as an additional signal. Since there are no players—the beauty-contest game is only a useful isomorphism—it does not make sense to speak of objective functions, and so endogenous information acquisition cannot be incorporated into that framework immediately. Nevertheless, the informational structure there can be recovered in the current paper by (for example) setting $z_{i\ell} = 1$ for all i and ℓ.

this fits well with the interpretation of z_i as a sample size, so that the precision of the observation increases linearly with z_i. In such a world a natural specification is $C(z) = c(Z)$ where $Z \equiv \sum_{i=1}^n z_i$, and where $c(\cdot)$ is an increasing, convex, and differentiable cost function which reflects the opportunity cost of spending a total period of time Z gathering information; if a market researcher's time can be purchased on the open market then it would be natural to suppose that $c(\cdot)$ is linear. Of course, the various information sources continue to vary in their clarity, so that listening to a given signal i for some period of time longer need not reveal the same quantity of information that listening to j for the same extra time would yield. It proves convenient to label the information sources in decreasing order of clarity, so that $\xi_1 < \xi_2 < \ldots < \xi_n$.[10] Equivalently, higher-indexed information sources are more expensive to acquire. Note, however, that this labelling has no implications for the underlying accuracy of the information; the clearest signal may well be subject to high-variance sender noise.

Given this specific form for the cost function, the marginal cost of information acquisition is independent of i; a little more formally, $\partial C(z)/\partial z_i = c'(Z)$ for all i. Inspecting (13), this implies that K_i is equal to some constant K across all i. A direct implication is that $z_i > 0$ if and only if $\xi_i < K$: the clearest signals receive attention and consequently influence players' actions, whilst the remaining signals are ignored.

Proposition 2 (Signal Acquisition: Additive Attention)**.** *Suppose that $C(z) = c(Z)$ where $Z \equiv \sum_{j=1}^n z_j$ and $c(\cdot)$ is increasing, convex, and differentiable. There is a unique K such that*

$$z_i = \frac{\xi_i \max\{(K - \xi_i), 0\}}{(1-\gamma)\kappa_i^2}. \tag{14}$$

Only the clearest signals (those that satisfy $\xi_i < K$) receive attention. Other things equal, signals with better accuracy receive more attention; raising the marginal-cost schedule $c'(\cdot)$ reduces the attention paid to all signals; and the attention paid to a signal is non-monotonic in its clarity.

The number of signals which attract attention falls as the marginal-cost schedule rises, as the accuracy of information sources improves, and as coordination becomes more important. When γ is sufficiently close to one then only one signal (the clearest) receives attention.

It is striking that not all signals necessarily receive attention: sufficient clarity is necessary (and, indeed, sufficient). Whilst clarity determines which information sources receive positive attention, accuracy determines—for those signals in use—how much attention each receives. Other things equal, more accurate (higher quality) signals receive more attention. Note, however, that a signal with appalling underlying accuracy (κ_i^2 is very high) is nevertheless both acquired and has influence (albeit receiving very little attention, and exerting very little influence) so long as its clarity is sufficient.

[10] Ties are excluded for convenience only. The propositions and proofs could be extended to accommodate ties (in a straightforward but cumbersome manner) but no fresh insight would be gained.

This feature is usefully understood by considering the marginal benefit to increased attention. Differentiating the quadratic-loss term from (9) it is readily verified that

$$-\frac{\partial}{\partial z_i}\left[\sum_{j=1}^{n} w_j^2\left((1-\gamma)\kappa_j^2+\frac{\xi_j^2}{z_j}\right)\right]=\frac{w_i^2\xi_i^2}{z_i^2}\propto\frac{1}{\xi_i^2}\left[\frac{\xi_i^2/z_i}{(1-\gamma)\kappa_i^2+\xi_i^2/z_i}\right]^2. \qquad (15)$$

(Note that while marginal calculations are used in this discussion, the claims of the results apply globally and not just locally.) This marginal benefit of increased attention depends on both κ_i^2 and ξ_i^2. However, an inspection of (15) confirms that as z_i shrinks to zero this marginal benefit depends only on the clarity of the information source. Intuitively, when z_i is small the total amount of noise in an information source is dominated by the receiver noise. So, when thinking about which information source to acquire a player begins with the clearest. However, as z_i increases away from zero the marginal benefit of further attention is no longer dominated by receiver noise, and so the accuracy of the underlying signal $\bar{x}_i = \theta + \eta_i$ becomes important. This means that information sources which are clear but inaccurate are acquired, but receive only a limited attention span.

Another notable feature of Proposition 2 is that the attention paid to a signal is non-monotonic in its clarity. Directly this is because the marginal benefit from increased attention paid to a signal is small whenever the signal is opaque (ξ_i^2 is large) or very clear (ξ_i^2 is small). Fixing z_i, when a signal is opaque then it attracts relatively little influence (w_i is low) and so the marginal benefit to paying further attention to it is low. On the other hand, when the signal is very clear (that is, ξ_i^2 is small) then there is very little receiver noise remaining in it; formally ξ_i^2/z_i is very small. This reduces the marginal benefit of increased attention, as an inspection of (15) confirms. A further interpretation is that z_i is the expenditure of a player on the ith information source. The product received from this expenditure is the precision z_i/ξ_i^2. Thus, in effect, ξ_i^2 is the price of acquiring the ith signal. The fact that attention is non-monotonic in clarity reflects the fact that optimised expenditure on a product is non-monotonic in its price.

This discussion suggests that it is the properties of the first bit of a signal, as z_i rises away from zero, that determine whether an information source is used. This is also true for a second natural specification in which the cost function is additively separable, so that $C(z) = \sum_{i=1}^{n} c_i(z_i)$. It is immediate that if $c_i'(0) = 0$ then $z_i > 0$; if listening to a signal for a very short period of time adds nearly nothing to costs, it will always be worth doing so. A more interesting situation is one in which information acquisition is always costly at the margin. For this case, without loss of generality set $\xi_i^2 = \xi^2$ for all i (because the ξ_i^2 parameter could be incorporated into the ith element of the cost function c_i), and label the information sources so that $c_1'(0) < c_2'(0) < \cdots < c_n'(0)$.[11] Thus the lower-indexed information sources are less costly at the margin when a player begins to bring a source into limited use. In this setting, attention is again focused on lower-indexed signals; it is useful to refer to these as the signals that are cheapest to acquire.

[11]Once again, no new insight is gained by considering the case of ties.

Proposition 3 (Signal Acquisition: Additive Costs). *Suppose that $C(z) = \sum_{j=1}^{n} c_j(z_j)$ where each $c_j(\cdot)$ is an increasing, convex, and differentiable function. Only the set of signals that are cheapest to acquire receive attention. The number of such signals falls as the accuracy of information sources improves and as coordination becomes more important to the players.*

The first claim does not imply that only a strict subset of signals are acquired; it is possible that all n information sources receive attention. However, those that receive no attention are the ones that are (perhaps unsurprisingly) the most expensive at the initial margin.

Related results also hold. For instance, it is natural to say that information source i is cheaper at the margin than j if the marginal-cost schedule for i lies everywhere below that for j. If this is the case, and if signal i has better underlying accuracy than j, then of course signal i attracts more attention (and gains more influence) than signal j.

The second claim of Proposition 3 echoes a result of Proposition 2: attention focuses on fewer signals as actions become complementary; as the coordination motive dominates players select a single focal point (the clearest or cheapest signal) and match their actions to it. If the coordination motive disappears then the focal-point motive is absent and a player cares only about identifying θ. In this case, players divide their attention across a wide range of information sources simply because there are decreasing returns to each individual signal; ignoring $C(z_\ell)$ for a moment, from (8) notice that $\mathrm{E}[u_\ell]$ is concave in $z_{i\ell}$.

Another result of interest is the effect of changing clarity or, equivalently, the cost of attention. Reducing the marginal cost of information acquisition (shifting down $c'(\cdot)$) is equivalent to increasing simultaneously the clarity of all signals. Whereas this has a predictable monotonic effect on the number of signals which are acquired, this is not the case when the clarities are changed individually: fixing ξ_j^2 for $j \neq i$, the size of the attention-receiving set is generally non-monotonic in ξ_i^2.[12] This is perhaps unsurprising, given that the relationship between z_i and ξ_i^2 is also non-monotonic. This non-monotonicity can be interpreted as a simple income effect: as ξ_i^2 falls it becomes cheaper to maintain a particular precision of observation of the ith signal. This frees the observer to divert attention elsewhere. Similarly, the non-monotonicity arises because a lower ξ_i^2 allows a player to obtain the same observation but with a lower (and hence cheaper) value of z_i.

Propositions 2 and 3 establish some properties of information-acquisition strategies. Players may restrict attention to a subset of signals; either the clearest (as in Proposition 2) or the cheapest to acquire (Proposition 3). Furthermore, these results record how the size of the attention-grabbing set changes with the players' environment. These results do not,

[12] Consider a world in which $n = 2$ and where $m = 1$; given that $\xi_1^2 < \xi_2^2$ it is always possible to construct such a scenario by choosing $1 - \gamma$ sufficiently small. Increasing ξ_1^2 up to ξ_2^2 will raise m, as certainly both signals are acquired whenever their clarities are equal. Also, when ξ_1^2 is lowered toward zero then m also rises. (Technically, some other conditions need to be imposed for this to be true; it is sufficient to impose an Inada condition on $c'(\cdot)$ by supposing that $c'(0) = 0$.) The reason is that the first signal becomes almost free to listen to: this reduces z_1 and so lowers the marginal cost of paying attention to the second information source. Drawing these observations together, there is no monotonic relationship between m and ξ_i^2.

however, reveal fully the amount of attention paid to each source as parameters change. Although more signals are acquired as the coordination motive weakens and as the accuracy of signals falls, it is not the case that each signal receives more individual attention. Indeed, for many specifications (including those in this section) any change in accuracy or the coordination motive which raises the attention given to one signal must necessarily reduce the attention paid to another.[13] Before describing how the pattern of attention changes, however, it is useful to consider the notion of a signal's publicity.

6. Publicity and Information Acquisition

Many contributions to the "beauty contest" literature have specified signals that are either public (perfectly correlated noise in signals) or private (uncorrelated noise). Here, and as already suggested in Sections 3-4, the correlation coefficient can index the general "publicity" of a signal. In equilibrium, the correlation between two players' observations of an information source is given by ρ_i in (11). (It is convenient to set $\rho_i = 0$ for a source with $z_i = 0$, which is the limit as $z_i \to 0$.) Two features distinguish the modelling framework here from existing work: firstly, the publicity of a signal can and does take intermediate values; and secondly, that publicity is endogenous.[14]

Once the (endogenous) publicities of signals (via their correlation coefficients) have been established, it is straightforward to explain how the pattern of attention paid to information sources changes with the players' desire for coordination. Intuitively, relatively public signals act as effective focal points for players coordination. As the desire for coordination weakens (γ falls) such signals become less influential and so the attention paid to them falls. In tandem, the attention paid to relatively private signals grows. This intuition is confirmed (at least for the leading cost specifications of interest) by the next proposition, which also describes the effect of changing signal accuracy.

Proposition 4 (Comparative-Static Exercises (i))**.** *Suppose that either $C(z) = c(\sum_{j=1}^{n} z_j)$ or $C(z) = \sum_{j=1}^{n} c_j(z_j)$, where $c(\cdot)$ and the various $c_j(\cdot)$ functions are increasing, convex, and differentiable. As the desire for coordination rises attention moves away from more private signals and toward more public signals: that is, there is a $\hat{\rho}$ such that the attention paid to signal i is locally increasing in γ if and only if $\rho_i > \hat{\rho}$. An increase in the underlying accuracy of a signal (a fall in κ_i^2) increases the attention paid to it, while reducing the attention paid to all other signals.*

The final comparative-static prediction is natural: attention falls away from poorer quality information sources. The effect of the coordination motive is more interesting, however: the change in the attention paid to an information source depends upon the associated signal's publicity, but this publicity is itself endogenous. In particular, as γ rises attention moves away from relatively private (uncorrelated) signals and so, as an inspection of (11)

[13] This statement holds, for instance, whenever the cost function satisfies $\partial^2 C(z)/\partial z_i \partial z_j \geq 0$.

[14] The first of these features is also present in Myatt and Wallace (2008); however, the endogenous information acquisition which is the central theme of this paper is absent from their model.

confirms, those signals become less correlated and so even more private; at the same time, the greater attention paid to the relatively public signals (that is, the highly correlated signals) makes them even more public by increasing their correlation coefficients. In essence, the heightened coordination motive spreads out the pattern of signals' publicities.

Since the correlation coefficients of signals (their publicities) are endogenous, it is interesting to consider how the exogenous properties of an information source, namely its underlying accuracy and its clarity, determine its equilibrium publicity. Given the use of the additive-attention specification for costs, so that $C(z) = c(\sum_{j=1}^{n} z_j)$, this is readily determined. To see this, notice that the correlation coefficient of the ith signal satisfies

$$\frac{\rho_i}{1-\rho_i} = \frac{\kappa_i^2}{\xi_i^2/z_i} = \frac{\max\{(K-\xi_i),0\}}{(1-\gamma)\xi_i}, \tag{16}$$

where the second equality is obtained by substituting in the solution for z_i from (14). Notice that the effect of the signal-accuracy term κ_i^2 cancels out; hence a signal is more correlated if it is clearer, in the sense that ξ_i^2 is lower.

Proposition 5 (Comparative-Static Exercises (ii))**.** *Suppose $C(z) = c(Z)$ where $Z \equiv \sum_{j=1}^{n} z_j$ and where $c(\cdot)$ is increasing, convex, and differentiable. In equilibrium the clearest signals are also the most public: if $\xi_i^2 < \xi_j^2$ then $\rho_i \geq \rho_j$. So, as the coordination motive strengthens, attention moves toward the clearest signals from the less clear signals. Total attention Z is decreasing in γ, and so players spend less on information acquisition as their desire to coordinate strengthens.*

The final claim is obtained by straightforward algebraic manipulations. The total attention Z becomes constant once γ is large enough for only one signal to receive attention.

The comparative-static results relating to γ may be recast in terms of the accuracy of the underlying signals. From (10), scaling up all of the κ_i^2s proportionately is equivalent to increasing $(1-\gamma)$ (κ_i^2 and $1-\gamma$ enter as a product and only in the expression for $\hat{\psi}_i$). Using Proposition 5, increasing all n of the κ_i^2s proportionately (equivalently reducing their accuracy) will (i) move attention away from the clearest signals and toward the less clear, (ii) increase the total attention paid, and so (iii) increase players' expenditure on information acquisition. Hence a general decrease in signal accuracy results in higher expenditure: the reduced accuracy increases the marginal benefits generated by any (endogenous) increase in clarity and hence induces players to pay more heed overall.

7. Equilibrium Actions and Beliefs

Having established some properties of players' information acquisition, consideration is now given to the beliefs which are induced and the actions which are taken. The statistical characteristics of actions and beliefs have been highlighted in the literature: Angeletos and Pavan (2007) referred to the "non-fundamental volatility" $\mathrm{var}[\bar{a} \mid \theta]$ and "dispersion" $\mathrm{var}[a_\ell \mid \theta, \bar{a}]$ of actions; these indicators appear also in Angeletos and Pavan (2004), where actions are interpreted as investment decisions. Angeletos and Pavan (2007) reported that these terms rise and fall, respectively, as the coordination motive strengthens. Here the information structure is a little richer; there are more than the familiar two public-and-private information sources, and the nature of informative signals is endogenous. In this broader setting it is useful to check that the properties of volatility, dispersion, and other indicators are retained. One purpose of this section is to do just that.

On average each action matches the underlying state: $\mathrm{E}[a_\ell \,|\, \theta] = \theta$. However, actions vary, and the extent to which players hit θ is measured by the variance $\mathrm{var}[a_\ell \,|\, \theta]$. Further measures of players' performance include the pairwise covariance $\mathrm{cov}[a_\ell, a_{\ell'} \,|\, \theta]$ (which is equal to the variance of the average action, $\mathrm{var}[\bar{a} \,|\, \theta]$, or what has been called non-fundamental volatility); the variance of actions across the player set $\mathrm{var}[a_\ell \,|\, \bar{a}, \theta]$ (the dispersion of actions); and the pairwise correlation coefficient $\mathrm{cov}[a_\ell, a_{\ell'} \,|\, \theta] / \mathrm{var}[a_\ell \,|\, \theta]$ of actions. If more structure is imposed on the cost function then three measures (variance, covariance, and correlation) all move together as the players' desire for coordination is changed, whereas the dispersion (the variance conditional on the average action) moves in the opposite direction. The specification imposed here is linear: $C(z) \propto \sum_{j=1}^{n} z_j$, which is equivalent to imposing a constant marginal cost of a player's time in a world where z_i is interpreted as the time spent listening to an information source. This functional form greatly simplifies the solution for K used in (14) and generates the following proposition.

Proposition 6 (Properties of Equilibrium Actions)**.** *Suppose that* $C(z) =$ *constant* $\times \sum_{j=1}^{n} z_j$*. The variance* $\mathrm{var}[a_\ell \,|\, \theta]$*, the covariance* $\mathrm{cov}[a_\ell, a_{\ell'} \,|\, \theta] = \mathrm{var}[\bar{a} \,|\, \theta]$*, and the correlation coefficient* $\mathrm{cov}[a_\ell, a_{\ell'} \,|\, \theta] / \mathrm{var}[a_\ell \,|\, \theta]$ *of players' actions all rise with the players' concern for coordination, whereas the conditional variance* $\mathrm{var}[a_\ell \,|\, \bar{a}, \theta]$ *falls.*

As the truth becomes less important and coordination more so, the correlation between players' actions rises, but they take actions that vary more around θ. Moreover, this result continues to apply as γ falls below zero. That is, if players are interested in doing what others do not, they will take increasingly uncorrelated actions (but based on the same information sources). This is despite the fact that, for small γ, the very strong preference to hit θ drives the variability of actions around θ down.

The properties of players' posterior beliefs also change with the coordination motive. Previous results have shown that as γ falls, players listen to more signals, listen for longer, and shift their attention away from the clearer information sources. However, it remains to establish what this means for posterior beliefs. It is natural to examine the conditional expectation of θ given the information acquired: that is, $\mathrm{E}[\theta \,|\, x_\ell]$. The following proposition begins with the variance of this expectation.

Proposition 7 (Properties of Equilibrium Expectations)**.** *Suppose that* $C(z) =$ *constant* $\times \sum_{j=1}^{n} z_j$*. The variance of conditional expectations about* θ*,* $\mathrm{var}[\mathrm{E}[\theta \,|\, x_\ell] \,|\, \theta]$*, increases with* γ*: as the coordination motive strengthens, beliefs about* θ *become more variable. If coordination is less important* ($\gamma < \frac{1}{2}$) *or there are few signals* ($n \leq 3$)*, the covariance of conditional expectations increases with* γ*: as the coordination motive strengthens, the coincidence of beliefs increases.*

Put rather more crudely, when players become more concerned with coordination then they tend to believe the wrong thing about θ, but at least they believe it together; in essence, their beliefs become more public (correlated) in nature.

Notice that the linear form of the cost function used in Propositions 6 and 7 fits with the market-research story; it corresponds to the case where there is a constant marginal cost of interviewing each additional surveyed consumer. Furthermore, the condition $\gamma < \frac{1}{2}$ used in Proposition 7 is automatically satisfied in the industry-supply setting; as Section 2 noted, this inequality corresponds to the assumption that the demand for a product is more sensitive to its own price than to the industry-wide average price.

8. Information Transmission and Rational Inattention

In the industry-supply scenario from the introduction and elsewhere in the paper, the cost of information acquisition is interpreted as a supplier paying for market researchers to survey various market segments. As noted in the previous section, if there were some fixed price per interview then a linear specification for $C(\cdot)$ might be natural. However, another view of the costs of information acquisition is suggested by the "rational inattention" literature. Here, the costs can be associated with the transmission, evaluation, and incorporation of the information into the decision-making process.

The rational inattention literature (Sims, 2010, provides a recent survey) supposes that there is a constraint on the information that may be processed (transmitted, evaluated, and so forth). It uses ideas from information theory (Cover and Thomas, 2006; MacKay, 2003) to model this. For data with a finite support the relevant concept is Shannon capacity (or Shannon entropy), which is in turn related to coding theory. Given a probability distribution over messages that could be sent, a coding system may be constructed (the Huffman algorithm) that optimally allocates bandwidth—shorter codes are used for common messages. Roughly speaking, entropy measures the average length of an optimally coded message.[15] When there are M different possible messages and message m occurs with probability p_m then the entropy is $-\operatorname{E}[\log p_m] = -\sum_{m=1}^{M} p_m \log p_m$, where the logarithm base determines the units of measurement. Entropy is minimised when a message always takes on a single value (no bandwidth is required, since it is known what the message will say) and is maximised by a uniform distribution over possible messages. It is a measure of the amount of uncertainty over a random variable.

The entropy definition may be extended to a continuous variable via the notion of differential entropy. This is defined as $H(x) \equiv -\operatorname{E}[\log f(x)]$ where x is a random variable with density $f(x)$. Similarly, the conditional differential entropy $H(x \,|\, y)$ measures the uncertainty about x after another random variable y has been observed. The change in entropy following such an observation is the mutual information between x and y, labelled $\mathcal{I}(x, y)$,

[15] Somewhat more precisely, entropy provides a lower bound to this length, and the use of an optimal coding algorithm achieves an average message length within one "bit" of the entropy.

and has the property $\mathcal{I}(x,y) = H(x) - H(x\,|\,y) = H(y) - H(y\,|\,x)$. The mutual information is a measure of how much bandwidth is needed to transmit the data required to update beliefs from (in an obvious notation) $F(x)$ to $F(x\,|\,y)$.

The differential entropy takes a convenient form when a variable is normally distributed. If x is an n-dimensional multivariate normal distribution then

$$H(x) = \tfrac{1}{2}\log\left[(2\pi e)^n \det[\Omega_x]\right], \tag{17}$$

where $\det[\Omega_x]$ is the determinant of the covariance matrix Ω_x. If y is another n-dimensional random variable and x and y are joint normally distributed then

$$\mathcal{I}(x,y) = H(x) - H(x\,|\,y) = \tfrac{1}{2}\log\left(\det[\Omega_x]/\det[\Omega_{x\,|\,y}]\right), \tag{18}$$

where $\Omega_{x\,|\,y}$ is the covariance matrix for the conditional distribution. The formula (18) may be applied to the model considered in this paper. Player ℓ observes a vector x_ℓ of noisy observations which are informative about the vector of true underlying signals $\bar{x}$. So, in this case a measure of the information transmitted to a player during the information-acquisition process is the mutual information $\mathcal{I}(x_\ell, \bar{x})$. The evaluation of this simply requires the calculation of the covariance matrices $\mathrm{var}[\bar{x}]$ and $\mathrm{var}[\bar{x}\,|\,x_\ell]$.

Note that with a diffuse prior, the prior entropy is undefined, and so a proper prior must be incorporated at this juncture. In particular, suppose that $\theta \sim N(\bar{\theta}, \varpi^2)$.[16] This is equivalent to introducing a signal (call it signal zero) with $\kappa_0^2 \equiv \varpi^2$ and $\xi_0^2 \equiv 0$. In the previous sections this could be interpreted as a costless signal.

Lemma 2. *The mutual information between x_ℓ and $\bar{x}$ satisfies*

$$2\,\mathcal{I}(x_\ell,\bar{x}) = \log\left(1 + \varpi^2 \sum\nolimits_{i=1}^{n} \frac{1}{\kappa_i^2 + (\xi_i^2/z_{i\ell})}\right) + \sum\nolimits_{i=1}^{n} \log \frac{1/\kappa_i^2 + z_{i\ell}/\xi_i^2}{1/\kappa_i^2}, \tag{19}$$

and also, when expressed in terms of the variance σ_i^2 and correlation ρ_i, satisfies

$$2\,\mathcal{I}(x_\ell,\bar{x}) = \log\left(\frac{1}{\varpi^2} + \sum\nolimits_{i=1}^{n} \frac{1}{\sigma_i^2}\right) - \log\frac{1}{\varpi^2} - \sum\nolimits_{i=1}^{n} \log(1-\rho_i). \tag{20}$$

The mutual information is increasing and concave in the information-acquisition choice z.

The first two terms in (20) represent the proportional change in the precision of a player's beliefs about θ. A third term reveals that the mutual information is higher for more public (higher correlation) signals. In essence, this is because such signals (holding σ_i^2 constant) contain more sender noise, and so there is more prior uncertainty about the underlying signal $\bar{x}_i$. By listening to such an information source, a player learns (and so absorbs information) about both θ and $\bar{x}_i$. In contrast, for a very private signal (κ_i^2 is low, and so $\bar{x}_i$ is close to θ) the listener is, in effect, learning only about θ.

Contributors to the rational-attention literature have modelled decision makers who face a capacity-constrained information channel. For instance, Maćkowiak and Wiederholt

[16] A proper prior over θ implies a proper prior over the underlying signal realisations.

(2009) studied a model in which firms divide their attention between idiosyncratic and aggregate conditions, and face a limit to the information they receive. Such a constraint can take the form $\mathcal{I}(x_\ell, \bar{x}) \leq \bar{\mathcal{I}}$ for some capacity term $\bar{\mathcal{I}}$. A related approach is for a player to incur a cost $C(z) = c(\mathcal{I}(x_\ell, \bar{x}))$, where $c(\cdot)$ is an increasing function.

Adopting this particular kind of entropy-based approach involves a different perspective on the information-acquisition process. For the market-research story which has featured throughout the paper, a specification such as $C(z) = c\sum_{i=1}^n z_i$ makes sense when the major source of costs is the deployment of researchers to conduct surveys. However, if such research is not so costly then the major bottleneck could be the transmission of the market-research data to a management team, and the subsequent assimilation of the information by that team. If this second aspect of the information-acquisition process is more important, then it may be more natural to employ an entropy-derived cost function.

9. Transmission Costs and Multiple Equilibria

Using an entropy-based cost function, however, generates a problem. Mutual information is strictly concave in z (Lemma 2) and (for $n \geq 2$) a cost function based on it cannot be convex. The discussion following Lemma 1 indicates that finding an equilibrium no longer corresponds to solving the minimisation problem of (9) when $C(z)$ is not convex.

Recall that for $\{z, w\}$ to form an equilibrium then given its play by others it should solve

$$\min_{w_\ell, z_\ell} \underbrace{\sum\nolimits_{i=0}^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}}\right]}_{L^\star(w_\ell, z_\ell)} + \gamma \underbrace{\sum\nolimits_{i=0}^n (w_{i\ell} - w_i)^2 \kappa_i^2}_{L^\dagger(w_\ell, w)} + C(z_\ell). \tag{21}$$

(The summations include a 0th term for the prior, where $\xi_0^2/z_{0\ell} = 0$ and $\kappa_0^2 = \varpi^2$.)

It has been observed that w_ℓ has no first-order effect on $L^\dagger(w_\ell, w)$ local to w, and so an equilibrium $\{z, w\}$ needs to be a local minimiser of $L^\star(w_\ell, z_\ell) + C(z_\ell)$, or more generally a stationary point.[17] If $C(\cdot)$ is convex then there is only one candidate for this, and so only one equilibrium. If convexity fails, however, then there may be multiple local minimisers of $L^\star(w_\ell, z_\ell) + C(z_\ell)$. A global minimiser is always an equilibrium. However, a local minimiser can also form an equilibrium: whereas a player can choose a non-local deviation which can strictly lower $L^\star(w_\ell, z_\ell) + C(z_\ell)$, the second-order effect of $L^\dagger(w_\ell, w)$ kicks in and can be strong enough to prevent that deviation. Assembling these observations, the next result develops Lemma 1 and offers a partial characterisation of symmetric linear equilibria; here, "payoff maximising" refers to a player's ex ante expected payoff.

Lemma 3. *If costs are entropy-based, so that $C(z) = c(\mathcal{I}(x_\ell, \bar{x}))$, then there may be multiple equilibria. A strategy which minimises $L^\star(w, z) + C(z)$ is a payoff-maximising equilibrium. Any other equilibrium is either a local minimiser or a stationary point of $L^\star(w, z) + C(z)$.*

[17] If an equilibrium is a local maximum then the convexity of $L^\dagger(w_\ell, w)$ in w must be strong enough to ensure that $L^\star(w_\ell, z_\ell) + L^\dagger(w_\ell, w) + C(z_\ell)$ achieves a local minimum.

Lemma 3 reveals the possibility of multiple equilibria, and so it is useful to find an example that fulfills this possibility. In the presence of a proper prior it is possible to do this by considering a world with only one information source ($n = 1$). Abusing (but, of course, simplifying) notation slightly, subscripts are dropped here so that the sender and receiver noise variances for this single signal are κ^2 and ξ^2/z respectively, and the weight placed on this signal in the linear equilibrium strategy is w, so that the remaining weight $1 - w$ is placed on the prior. Using Proposition 1, these weights satisfy

$$w = \frac{(1-\gamma)\varpi^2}{(1-\gamma)(\varpi^2+\kappa^2)+(\xi^2/z)} \quad \text{and} \quad 1-w = \frac{(1-\gamma)\kappa^2+(\xi^2/z)}{(1-\gamma)(\varpi^2+\kappa^2)+(\xi^2/z)}. \tag{22}$$

Adopting the cost function $C(z) = 2c\mathcal{I}(x_\ell, \bar{x})$, so that costs are linearly increasing in the bandwidth required for the transmission of information, when $n = 1$ the entropy-based cost function takes the particularly simple form

$$C(z) = c\log\left(\frac{\kappa^2+\varpi^2+(\xi^2/z)}{(\xi^2/z)}\right), \tag{23}$$

while the expected quadratic loss from the beauty-contest components is

$$L^\star(w,z) = w^2((1-\gamma)\kappa^2+(\xi^2/z)) + (1-w)^2(1-\gamma)\varpi^2. \tag{24}$$

In these expressions ξ^2 and z only enter as a ratio (this is true more generally when costs are entropy based) and so there is nothing lost by setting $\xi^2 = 1$ (this is a change in the units of z). Doing so, and substituting the solutions for w and $1-w$, a player's loss as a function of z is $L(z) = L^\star(z) + C(z)$ where $L^\star(z) \equiv \min_{w\in[0,1]} L^\star(w,z)$. This satisfies

$$L(z) = \frac{(1-\gamma)\varpi^2(1+(1-\gamma)\kappa^2 z)}{1+(1-\gamma)(\varpi^2+\kappa^2)z} + c\log\left(1+(\kappa^2+\varpi^2)z\right). \tag{25}$$

An examination of $L(z)$ permits the identification of candidate equilibria. For instance, a z which successfully minimises this expression subject to $z \geq 0$ will yield a payoff-maximising equilibrium (Lemma 3). This approach yields the next result.

Proposition 8 (Equilibria with Entropy-Derived Information-Acquisition Costs)**.** *Suppose that there is a single information source and that the cost of paying attention to it is linearly increasing in the mutual information, so that $C(z) = 2c\mathcal{I}(x_\ell, \bar{x})$. Define:*

$$\bar{c} = \frac{((1-\gamma)\varpi^2)^2}{\kappa^2+\varpi^2}. \tag{26}$$

If $\gamma < \frac{1}{2}$ (so that the coordination motive is relatively weak) then there is a unique equilibrium. Players acquire no new information ($z = w = 0$) if and only if $c \geq \bar{c}$.

If $\gamma > \frac{1}{2}$ (the coordination motive is stronger) then there may be multiple equilibria. If $c > \bar{c}/(2(1-\gamma))$ then there is a unique equilibrium in which players acquire no new information ($z = w = 0$). If $c < \bar{c}$ then there is a unique equilibrium in which players pay attention to the signal. However, if $\bar{c} \leq c \leq \bar{c}/(2(1-\gamma))$ then $L(z)$ is locally minimised at $z = 0$, but for c close enough to $\bar{c}$ then $L(z)$ has a local maximum and a local minimum for two positive values of z.

Interestingly, the uniqueness result is retained if $\gamma < \frac{1}{2}$. This is natural: one force in favour of multiple equilibria is the presence of the coordination motive, and so when this motive is weakened there is only one equilibrium. If γ is smaller then the convexity of $L^\star(z)$ (the expected quadratic loss from the play of the beauty-cost game) overcomes the concavity of the entropy-based cost function $C(z)$. Recall also that the inequality $\gamma < \frac{1}{2}$ is automatically satisfied in the context of the industry-supply example which has been discussed throughout the paper.

Nevertheless, Proposition 8 also confirms that multiple equilibria may be present when the coordination motive is strong. This is readily illustrated using the parameters

$$\gamma = \tfrac{4}{5}, \quad \varpi^2 = 5, \quad \kappa^2 = 0, \quad \text{and} \quad c = \tfrac{1}{4}. \tag{27}$$

It is straightforward to evaluate the loss $L^\star(z_\ell, w_\ell) + L^\dagger(w_\ell, w) + C(z_\ell)$ for a player ℓ choosing a strategy $\{z_\ell, w_\ell\}$ when others choose w. (The information acquisition decision of others is of no direct relevance to player ℓ.) Figure 1 illustrates various losses as a function of the weight w_ℓ placed on the informative signal by the player. For each choice of w_ℓ the optimal information acquisition choice is used. It is readily verified that this satisfies

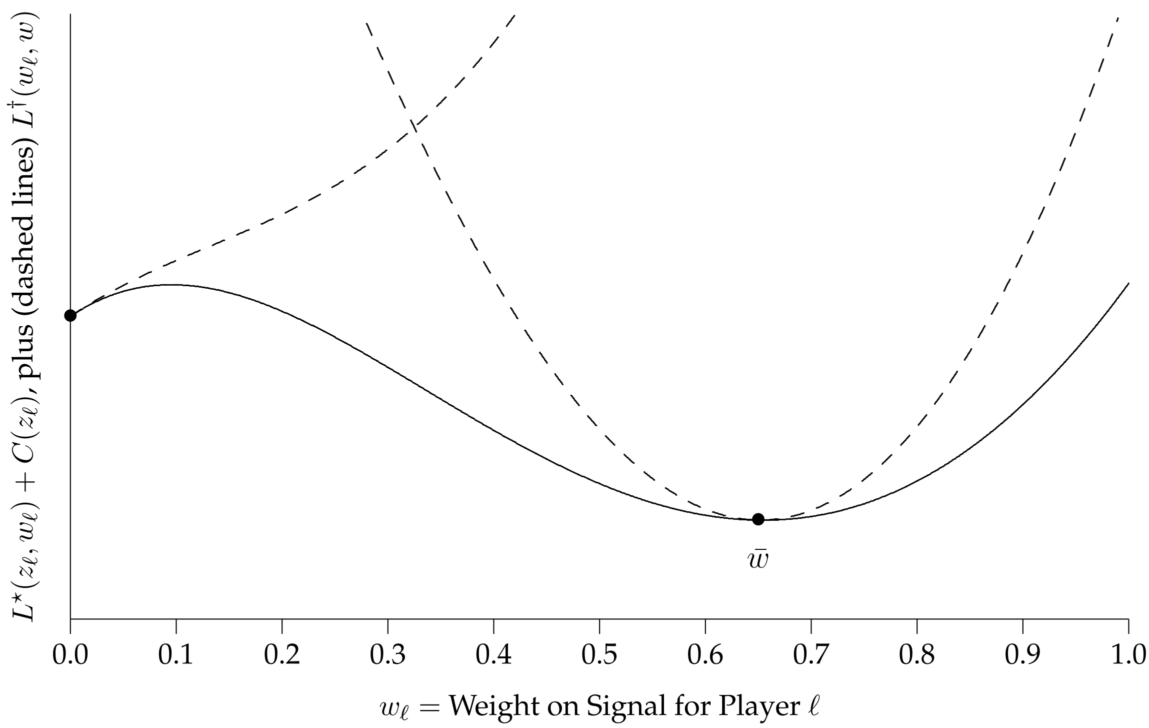
$$z_\ell = \frac{w_\ell^2}{2c}\left[1 + \sqrt{1 + \frac{4c}{w_\ell^2(\kappa^2 + \varpi^2)}}\right]. \tag{28}$$

The parameter choices made here yield thresholds $\bar{c} = \frac{1}{5}$ and $\bar{c}/(2(1-\gamma)) = \frac{1}{2}$ which enclose the cost parameter c, and so (from Proposition 8) multiple equilibria may arise. The solid line in Figure 1 illustrates that there are multiple local minima to $L^\star(z_\ell, w_\ell) + C(z_\ell)$. The inclusion of the deviate-from-others term $L^\dagger(w_\ell, w)$ to generate the dashed lines demonstrates that this example exhibits multiple equilibria: one in which no weight is put on the signal, and another in which it attracts significant weight.

The presence of multiple equilibria and other aspects of the entropy-based cost structure make it difficult to characterise fully the equilibrium set. Nevertheless, some progress can be made. Returning to the general case of n information sources, it is natural to ask: which information sources do players choose to use?

The first (and easy) result is that the clarity of an information source, determined by ξ_i^2, no longer matters. The parameter ξ_i^2 changes the cost of acquiring data, but in the context of entropy-based information-transmission constraints, data is not directly costly. Instead, the cost arises from the information content of the data and this depends on ξ_i^2/z_i. The only substantive characteristic of an information source is its underlying accuracy, determined by the sender-noise variance κ_i^2. A second natural result that might be expected is that better accuracy helps players, and that they choose (in equilibrium) to acquire the signals with better accuracy. This is true, but nonetheless requires a little work.

The extra complication arises because increased accuracy raises costs as well as benefits; this contrasts with the earlier specifications in which the cost of the ith signal is independent of κ_i^2. What this means is that increased signal accuracy (a fall in κ_i^2) can sometimes



For $n = 1$ this figure illustrates the expected loss to a player ℓ as a function of the weight w_ℓ placed on the signal. The cost function $C(z) = 2c\mathcal{I}(x_\ell, \bar{x})$ is based on the mutual information from Lemma 2. The parameter choices are $\gamma = \frac{4}{5}$, $\varpi^2 = 5$, $\kappa^2 = 0$, and $c = \frac{1}{4}$. The solid line illustrates $L^\star(z_\ell, w_\ell) + C(z_\ell)$ as a function of w_ℓ, where for each w_ℓ the information acquisition z_ℓ is chosen optimally. There are two local minima, at $w = 0$ and $w = \bar{w} > 0$ where $\bar{w} \approx 0.65$. The latter minimum generates a payoff-maximising equilibrium. The dashed lines illustrate $L^\star(z_\ell, w_\ell) + L^\dagger(w_\ell, w) + C(z_\ell)$ for $w \in \{0, \bar{w}\}$, and so include the term $L^\dagger(w_\ell, w)$ which punishes player ℓ for deviating from the choice w by others. Including this extra term for $w = 0$ ensures that $w_\ell = 0$ is a unique best reply from player ℓ, and so $w = z = 0$ is an equilibrium, even though it is not payoff maximising.

FIGURE 1. Multiple Equilibria with Entropy-Based Information Costs

hurt rather than harm a player. However, when evaluated in the context of an equilibrium a player's attention choice already takes into account the conflicting costs and benefits of each signal, and at this point increased accuracy is always welcome. The derivations which confirm this (relegated to Appendix A) generate Proposition 9.

Proposition 9 (Properties of Acquired Signals with Entropy-Derived Costly Attention). *Suppose that the players face an information-transmission constraint: the cost of information is an increasing function of the mutual information arising from the signal observations. In the payoff-maximising equilibrium, only the most accurate signals receive attention and exert influence; if the accuracy of a signal is sufficiently poor then it is ignored. A player's equilibrium payoff is strictly increasing in the underlying accuracy of signals that are used.*

The first claim of this proposition is a corollary of the fact that signal accuracy is payoff-improving in equilibrium. It means that a player would always find it optimal to swap a lower-accuracy signal for a higher-accuracy alternative.

Emerging from this section, then, are two messages which contrast with earlier results. Firstly, if rational-inattention information-transmission constraints are present then multiple equilibria may arise. This multiplicity arises because of non-convexities in players' cost functions; the rational-attention approach generates increasing returns on the cost side. Nevertheless, in a leading case of interest (when $n = 1$, so that there is a prior plus a single informative signal) there is a unique equilibrium so long as the coordination motive is not too strong; this is satisfied in the leading industry-supply scenario. Secondly, the move to entropy-derived costs generates a different pattern of attention. When the cost of information is based on the raw data obtained then players use the subset of signals with the best clarity, but not necessarily with the best underlying accuracy. In contrast, when data is cheap but there are limitations to transmission and data processing then the pattern switches to one in which the signals acquired are those with better accuracy.

10. Related Literature and Concluding Remarks

Researchers including Morris and Shin (2002, 2005), Hellwig (2005), and Angeletos and Pavan (2004, 2007) have studied models in which the players of beauty-contest games have exogenous access to information sources; for most papers (although not all) such informative signals are either "public" or "private" in nature.[18] This paper contributes in two ways: firstly, it allows for endogenous information acquisition; and secondly, it allows that acquisition to change the nature (in particular, the publicity as well as the precision) of the signals. Other recent papers have also considered endogenous information acquisition, and so this concluding section relates this paper to work by Dewan and Myatt (2008), by Hellwig and Veldkamp (2009), and by contributors to the rational-attention literature, such as Maćkowiak and Wiederholt (2009).[19,20]

[18] There are some recent exceptions. Papers which admit a more general signal structure include Myatt and Wallace (2008), Baeriswyl and Cornand (2006, 2007), Baeriswyl (2007), Angeletos and Pavan (2009), as well as those mentioned below: Dewan and Myatt (2008) and Hellwig and Veldkamp (2009).

[19] The small and most directly related literature discussed here is distinct from the contemporaneous literature on dynamic coordination games with endogenous information (Angeletos and Pavan, 2009; Angeletos and La'O, 2009, for instance). There the endogeneity arises from the fact that agents observe noisy signals of past behaviour which aggregate the dispersed (and exogenous) information available to agents up until that point. However, agents do not choose what to observe nor how carefully to observe it. Related to this literature, various recent contributions use a similar approach to study, for example, asset pricing and informational feedback effects (Ozdenoren and Yuan, 2008) or how the aggregate trading of currency speculators endogenously generates information for a policy maker (Goldstein, Ozdenoren, and Yuan, 2011).

[20] Another related and interesting strand of literature is the work of Calvó-Armengol and de Martí Beltran (2007, 2009) and particularly Calvó-Armengol, de Martí Beltran, and Prat (2009). In these papers a set of players arranged on a network share information they hold concerning the state of the world with others they are linked to on the network, before playing a beauty contest of the sort studied above. The papers study the impact that the network structure has upon the spread of actions in the game where (in the first two papers) that network structure is exogenous and (in the third paper) the players themselves decide

The model of Dewan and Myatt (2008) is closely related to this one; many of their results are special cases of those presented here. They used a beauty-contest game as a metaphor for a political party. Party members must advocate a policy, and in so doing want to do the right thing for the party (a policy close to θ) while preserving party unity (a policy close to the "party line"). Before making their decisions they listen to leaders. These leaders personify information sources. Party members can divide a fixed unit of time between listening to different leaders. Their model is equivalent to a special case of the one presented here: their cost function takes the form $C(z) = c(Z)$ for $Z = \sum_{j=1}^{n} z_j$ where costs are zero if $Z \leq 1$ and infinite (or, at least, sufficiently large) otherwise.[21]

This paper offers a complementary perspective to the messages of Hellwig and Veldkamp (2009). Their first main result shows that the incentive for players to acquire information is enhanced when others acquire information and when actions are complementary; as their title suggests, players want to know what others know whenever they want to do what others do. They then considered information-acquisition equilibria, and the main message from that work is that there can be many (linear) equilibria. For instance, when players are faced with a choice between acquiring a signal or not, there may be an equilibrium in which everyone acquires the signal and another equilibrium in which everyone ignores it; this seems natural given the complementarity inherent in information acquisition. This result survives even when information acquisition is "near continuous" so that the precision of the signal in question and its cost are both small. This contrasts noticeably with the findings of this paper, in which the equilibrium is unique. So what explains the difference in the messages of these two papers?

In the model specified by Hellwig and Veldkamp (2009), the publicity of a signal (its correlation coefficient) is exogenous.[22] Even if only a small amount is spent on information acquisition, so that the extra precision obtained and the extra cost incurred are both small, the correlation coefficient is bounded away from zero. For instance, a player can acquire a very small bit of a public signal. Heuristically, at least, such a public signal is more valuable if others are acquiring it too; this is the source of the multiplicity. As Hellwig and Veldkamp (2009) correctly observed, this does not happen with private signals: when signal realisations are uncorrelated, there is a unique equilibrium.

---

whether to form pairwise links with other players at some cost, thereby endogenising the network structure and hence the information acquired. Given that extant information is passed between players, the focus of this work is elsewhere, however it is related to the current paper to the extent that information acquisition is endogenous and co-determined with the actions of the underlying beauty contest.

[21] Dewan and Myatt (2008) also allowed the properties of information sources to be endogenous by considering the rhetorical strategies of leaders: such leaders vary their clarities (ξ_i^2) in order to attract attention.

[22] In the world of Hellwig and Veldkamp (2009) a player either acquires a signal or not. For an acquired signal the variance of the receiver noise is fixed. Within the context of this paper, this is equivalent to specifying, for some $\bar{z}_i$, a cost function where $c_i(z_i) = 0$ for $z_i = 0$, $c_i(z_i) = \bar{c}_i$ for $0 < z_i \leq \bar{z}_i$, and $c_i(z_i) = \infty$ otherwise. Obviously, this is non-convex and so an ingredient of the uniqueness result in this paper is missing. Although this is a technical reason for the present of multiple equilibria in their model, it is not a useful explanation; they key issue is the exogeneity of the correlation coefficient.

Here there is a more nuanced view. As a player pays more attention to an information source (z_i grows) then the correlation of the signal realisations rises too; hence the publicity of a signal, as well as its precision, is under the control of the acquiring player. Thus, implicitly at least, this model endogenises the nature of acquisition as well as the decision to acquire. Crucially, the first bit of a signal acquired is private in nature: the correlation coefficient falls to zero as z_i vanishes. This smoothes things out sufficiently to ensure that there is a unique equilibrium. Thus, when Hellwig and Veldkamp (2009, p. 224) stated that a requirement for uniqueness is that "the information agents choose to acquire must also be private" they were correct only when the decision is to acquire or not; if players choose how carefully to listen then the important feature is that the signal is almost perfectly private when a player pays almost no attention to it. The move from a model with multiple equilibria to one with a unique equilibrium is of interest because it allows for rich comparative-static exercises. Whereas Hellwig and Veldkamp (2009) offered the knowing-what-others-know and the multiple-equilibria messages, here the uniqueness of the equilibrium allows specific predictions about what kind of information is acquired and how acquisition decisions change with both the nature of the information and the nature of the coordination problem faced by the players.

Multiple equilibria can re-appear with a very different cost specification. The rational-inattention literature (Sims, 2003; Maćkowiak and Wiederholt, 2009) has promoted the study of information-transmission constraints derived from information theory and coding theory. This leads naturally to a cost function which exhibits increasing returns: doubling the attention paid to an information source does not double the costs. The marginal cost of increased attention is decreasing, simply because the marginal datum acquired contains less new information than an infra-marginal datum, and so there is less to transmit. This all fits within the general model of this paper, but the non-convexities in the cost function again permit multiple equilibria.

The take-home message from this paper, then, is that the nature of equilibrium endogenous information acquisition in coordination games turns upon the nature of the cost function which players face. One possibility is that a player decides simply whether to acquire a signal or not, where the clarity of the acquired signal is exogenous. For instance, if a player acquires some specific economic data (stock prices, perhaps) then the content might be unambiguous. This is the world of Hellwig and Veldkamp (2009). There are multiple equilibria and so limited opportunities for comparative-static exercises. A second possibility is that players choose how much attention to pay to each information source, so that the clarity of the acquired signal is endogenous. For instance, if a supplier learns about market conditions by conducting market research then better information can be acquired by a larger (and so more costly) survey. This the world of Section 3–7. Under natural cost specifications there is a unique equilibrium. Players pay attention to the clearest signals (even if their underlying accuracies are poor) and the subset of attention-grabbing signals shrinks as the coordination motive grows. The third possibility involves constraints to information transmission and comprehension. For instance, a supplier may find it easy to acquire data (consider, for instance, the wealth of scanner-based data cheaply acquired by a supermarket chain) but costly to assimilate and process it. This is a world in which the entropy-based information constraints suggested by Sims (1998, 2003, 2005, 2006) become relevant. There are increasing returns on the cost side, and so multiple equilibria can return. Nevertheless, there can still be a unique equilibrium when the coordination motive is not too strong (this is true in the industry-supply example). In contrast to the costly data case, players choose to acquire the most accurate sources of information rather those with the best clarity.

Appendix A. Omitted Proofs

Proof of Lemma 1. In both the text and the lemma it is claimed that any linear equilibrium strategy satisfies $\sum_{i=1}^n w_i = 1$. To see why, consider a linear equilibrium strategy profile $A(x_\ell) = w' x_\ell$, where w' is the transpose of $w \in \mathcal{R}^n$. Given the linearity, $\mathrm{E}[A(x_{\ell'}) \,|\, x_\ell] = w' \,\mathrm{E}[x_{\ell'} \,|\, x_\ell]$. Given normality, the latter conditional expectation satisfies $\mathrm{E}[x_{\ell'} \,|\, x_\ell] = B x_\ell$ where B is a $n \times n$ inference matrix with the property that the rows of B sum to one. Similarly, $\mathrm{E}[\theta \,|\, x_\ell] = a' x_\ell$ where the elements of $a \in \mathcal{R}^n$ also sum to one. Using (7), $w' x_\ell = (1-\gamma)\,\mathrm{E}[\theta \,|\, x_\ell] + \gamma\,\mathrm{E}[A(x_{\ell'}) \,|\, x_\ell] = [(1-\gamma)a + \gamma B' w]' x_\ell$, and hence $w = (1-\gamma)a + \gamma B' w$. Given that the elements of a sum to one and each column of B' sums to one, this equality can only hold if the elements of w sum to one. So, when looking for linear equilibria it is sufficient to look for those satisfying $\sum_{i=1}^n w_i = 1$. Moreover, any best reply to such a strategy also satisfies this equality. Thus it is permissible to impose the constraint $\sum_{i=1}^n w_i = 1$ upon each player when seeking equilibria.

To obtain (8), note that $\sum_{i=1}^n w_{i\ell} = 1$ for player ℓ implies that $a_\ell - \theta = \sum_{i=1}^n w_{i\ell}(\eta_i + \varepsilon_{i\ell})$, and so

$$\mathrm{E}[(a_\ell - \theta)^2] = \sum\nolimits_{i=1}^n w_{i\ell}^2 \left(\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}} \right). \tag{29}$$

The average action is $\bar{a} = \theta + \sum_{i=1}^n \eta_i$, since the individual-specific errors disappear via the law of large numbers and so $a_\ell - \bar{a} = \sum_{i=1}^n w_{i\ell}\varepsilon_{i\ell} + \sum_{i=1}^n (w_{i\ell} - w_i)\eta_i$. Hence:

$$\mathrm{E}[(a_\ell - \bar{a})^2] = \sum\nolimits_{i=1}^n \frac{w_{i\ell}^2 \xi_i^2}{z_{i\ell}} + \sum\nolimits_{i=1}^n (w_{i\ell} - w_i)^2 \kappa_i^2. \tag{30}$$

Substituting these two expressions yields the expression for $\mathrm{E}[u_\ell]$ given in (8). Given this solution, the pair $\{z, w\}$ yield a symmetric equilibrium if and only if

$$\{z, w\} \in \operatorname*{arg\,min}_{z_\ell \in \mathcal{R}_+^n, w_\ell \in \mathcal{R}^n} L^\star(w_\ell, z_\ell) + L^\dagger(w_\ell, w) + C(z_\ell) \quad \text{subject to} \quad \sum\nolimits_{i=1}^n w_{i\ell} = 1, \quad \text{and where}$$

$$L^\star(w_\ell, z_\ell) \equiv \sum\nolimits_{i=1}^n w_{i\ell}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{i\ell}} \right] \quad \text{and} \quad L^\dagger(w_\ell, w) \equiv \gamma \sum\nolimits_{i=1}^n (w_{i\ell} - w_i)^2 \kappa_i^2. \tag{31}$$

This combined loss function is strictly convex in its arguments. Thus, the unique solution to the minimisation problem is determined by the relevant first-order conditions. Local to w, however, changes in w_ℓ have no first-order effect on $L^\dagger(w_\ell, w)$. Thus the component $L^\dagger(w_\ell, w)$ can be ignored when dealing with the relevant first-order conditions. This all implies that $\{z, w\}$ uniquely minimises $L^\star(w_\ell, z_\ell) + C(z_\ell)$, subject of course to the constraint $\sum_{i=1}^n w_{i\ell} = 1$. $\square$

Proof of Proposition 1. The expression for z_i can be obtained from the first-order condition with respect to z_i. To obtain the solutions for the influence weights w, fix z and note that the optimisation problem is to minimise $L^\star \equiv \sum_{i=1}^n (w_i^2/\hat{\psi}_i)$ subject to $\sum_{i=1}^n w_i = 1$. A solution must satisfy $\partial L^\star/\partial w_i = \partial L^\star/\partial w_j$ for all $i \neq j$, which holds if and only if $w_i \propto \hat{\psi}_i$.

It is useful at this point to derive (13) Once the equilibrium weights w have been substituted into the objective function, the solution for z emerges by minimising $L^\star(z) + C(z)$ where

$$L^\star(z) \equiv \frac{1}{\sum_{i=1}^n \hat{\psi}_i} \quad \text{and where} \quad \hat{\psi}_i \equiv \frac{1}{(1-\gamma)\kappa_i^2 + \xi_i^2/z_i}. \tag{32}$$

For $z_i > 0$ the first-order condition with respect to z_i takes the form

$$-\frac{\partial L^*(z)}{\partial z_i} = \frac{\partial C(z)}{\partial z_i} \quad \Leftrightarrow \quad \frac{\xi_i^2}{((1-\gamma)\kappa_i^2 z_i + \xi_i^2)^2} = \left(\sum_{j=1}^n \hat{\psi}_j\right)^2 \frac{\partial C(z)}{\partial z_i} \equiv \frac{1}{K_i^2}, \tag{33}$$

which can be re-arranged to yield (13). (Note that the first-order condition can hold only if $\xi_i < K_i$. Furthermore, a solution to the minimisation problem also requires $\xi_i \geq K_i$ whenever $z_i = 0$.) $\square$

Proof of Proposition 2. $\partial C(z)/\partial z_i = c'(Z)$ for all i and so $K_i = K$ for all i. The calculation of (13) noted that $\xi_i < K_i$ when $z_i > 0$ and $\xi_i \geq K_i$ when $z_i = 0$. Given that that $K_i = K$ for all i, this implies that the information sources attracting attention are those with the lowest ξ_i. This yields the first claim. Substituting the expression for z_i from (13) into $\hat{\psi}_i$ yields

$$\sum_{j=1}^n \hat{\psi}_j = \frac{1}{(1-\gamma)K} \sum_{j=1}^m \frac{(K-\xi_j)}{\kappa_j^2} = \frac{1}{(1-\gamma)K} \sum_{j=1}^n \frac{\max\{(K-\xi_j), 0\}}{\kappa_j^2}. \tag{34}$$

The second part of (13) yields $1/K = \sqrt{c'(Z)} \sum_{j=1}^n \hat{\psi}_j$. Combining this with (34):

$$c'\left(\sum_{j=1}^n \frac{\xi_j \max\{(K-\xi_j), 0\}}{(1-\gamma)\kappa_j^2}\right) \left(\sum_{j=1}^n \frac{\max\{(K-\xi_j), 0\}}{(1-\gamma)\kappa_j^2}\right)^2 = 1. \tag{35}$$

The left-hand side of (35) is increasing in K, and so (35) yields a unique solution for K. This can be used to obtain the solution for the individual attention levels paid to each information source.

Turning to the properties of the equilibrium, the first claim follows by inspection. The second claim is obtained by observing that anything which increases the left-hand side of (35) must reduce K and so the attention paid to any signal. The third claim is by inspection. Regarding the number of attention-receiving signals, the left-hand side of (35) is increasing in γ and κ_i^2 for each i, and also falls as $c'(\cdot)$ falls. Hence the solution K (and so the number of attention-grabbing signals) decreases with γ and κ_i for each i but falls as $c'(\cdot)$ rises. Finally, as γ approaches one from below, K converges to a lower bound $\bar{K}$. If $\bar{K} > \xi_1$ then the left-hand side of (35) diverges, and so the equality cannot hold. Hence it must be the case that $\bar{K} = \xi_1$, which means that K must fall below ξ_2 for $1-\gamma$ sufficiently small, and so all signals $i > 1$ are ignored for γ close enough to one. $\square$

Proof of Proposition 3. Contrary to the proposition, suppose that players ignore the ith information source (so that $z_i = 0$) while listening to source $i+1$ (so that $z_{i+1} > 0$). Now

$$\xi < K_{i+1} = \frac{1}{\sqrt{c'_{i+1}(z_{i+1})} \sum_{j=1}^n \hat{\psi}_j} \leq \frac{1}{\sqrt{c'_{i+1}(0)} \sum_{j=1}^n \hat{\psi}_j} < \frac{1}{\sqrt{c'_i(0)} \sum_{j=1}^n \hat{\psi}_j} = K_i. \tag{36}$$

The first inequality holds because $z_{i+1} > 0$; the second is from the convexity of $c_{i+1}(\cdot)$; and the third inequality holds by assumption. This implies $\xi < K_i$, which contradicts $z_i = 0$.

Combine the equalities from (13) to obtain

$$z_i = \frac{\xi_i}{(1-\gamma)\kappa_i^2} \max\left\{\left(\frac{1}{\Psi\sqrt{c_i'(z_i)}} - \xi_i\right), 0\right\} \quad \text{where} \quad \Psi \equiv \sum\nolimits_{j=1}^{n} \hat{\psi}_j. \tag{37}$$

This also holds for $z_i = 0$. Treating Ψ as a constant, the right-hand side of the first equation in (37) is decreasing in z_i and so (37) yields a unique solution $z_i = f_i(\gamma, \kappa_i^2, \xi_i, \Psi)$ for some function $f_i(\cdot)$. That solution is increasing in γ, but decreasing in κ_i^2, and Ψ. Given this, the second equation in (37) can be written

$$\Psi = \sum\nolimits_{j=1}^{n} \frac{1}{(1-\gamma)\kappa_j^2 + [\xi_j^2 / f_j(\gamma, \kappa_j^2, \xi_j, \Psi)]}. \tag{38}$$

Given the observations made so far, the right-hand side of this equation is decreasing in Ψ, and so (38) yields a unique solution for Ψ. The right-hand side is also increasing in γ and decreasing in κ_j^2 for each j, and so the solution Ψ is respectively increasing and decreasing in these parameters. This property of Ψ is enough to establish the proposition's remaining claims. To see why, inspect (37) and notice that an information source i is ignored if and only if $\xi_i \Psi \sqrt{c_i'(0)} > 1$. If γ is increased or κ_j^2 is reduced, then the consequent increase in Ψ strengthens this inequality and so information source i continues to be ignored. □

Proof of Proposition 4. Consider the cost specification $C(z) = c(\sum_{j=1}^{n} z_j)$ and an information source satisfying $z_i > 0$. Differentiate the solution for z_i stated in Proposition 2 to obtain

$$\frac{dz_i}{d\gamma} = \frac{\xi_i(K-\xi_i)}{(1-\gamma)^2\kappa_i^2} - \frac{\xi_i}{(1-\gamma)\kappa_i^2}\frac{dK}{d\gamma} < 0 \quad \Leftrightarrow \quad \xi_i > K + (1-\gamma)\frac{dK}{d\gamma}, \tag{39}$$

and so attention falls with γ if and only if the clarity of an information source is sufficiently poor. However, in equilibrium the correlation coefficient ρ_i of a signal is monotonic in its clarity:

$$\rho_i = \frac{\kappa_i^2}{\kappa_i^2 + \xi_i^2/z_i} = \frac{K-\xi_i}{K-\gamma\xi_i}. \tag{40}$$

Turning to the specification $C(z) = \sum_{j=1}^{n} c_j(z_j)$, use (37) for $z_i > 0$ to obtain

$$z_i = \frac{\xi_i}{(1-\gamma)\kappa_i^2}\left(\frac{1}{\Psi\sqrt{c_i'(z_i)}} - \xi_i\right). \tag{41}$$

Now, z_i is decreasing in γ if and only the right-hand side is decreasing in γ when z_i is fixed. Differentiating the right-hand side yields

$$\begin{aligned}\frac{\partial}{\partial\gamma}\left[\frac{\xi_i}{(1-\gamma)\kappa_i^2}\left(\frac{1}{\Psi\sqrt{c_i'(z_i)}} - \xi_i\right)\right] &= \frac{\xi_i}{(1-\gamma)^2\kappa_i^2}\left(\frac{1}{\Psi\sqrt{c_i'(z_i)}} - \xi_i\right) - \frac{\xi_i}{(1-\gamma)\kappa_i^2}\left(\frac{1}{\Psi^2\sqrt{c_i'(z_i)}}\right)\frac{\partial\Psi}{\partial\gamma} \\ &= \frac{z_i}{1-\gamma} + \left(z_i + \frac{\xi_i^2}{(1-\gamma)\kappa_i^2}\right)\frac{\partial\log\Psi}{\partial\gamma} > 0 \quad \Leftrightarrow \quad 1 - \left(1-\gamma+\frac{\xi_i^2/z_i}{\kappa_i^2}\right)\frac{\partial\log\Psi}{\partial\gamma} < 0.\end{aligned} \tag{42}$$

The term specific to i is monotonic in the correlation coefficient $\rho_i = \kappa_i^2/(\kappa_i^2 + \xi_i^2/z_i)$. Specifically:

$$\frac{dz_i}{d\gamma} < 0 \quad \Leftrightarrow \quad \left(1-\gamma+\frac{1-\rho_i}{\rho_i}\right)\frac{\partial\log\Psi}{\partial\gamma} > 1, \tag{43}$$

where the last step uses the fact that Ψ is increasing in γ. This final equality holds if and only if ρ_i is sufficiently small; that is, if and only if the information source is relatively private. $\square$

For the next three proofs, the notation m indicates the number of active information sources; hence $z_i > 0$ for $i \leq m$ but $z_i = 0$ for all $i > m$, where signals have been ordered appropriately.

Proof of Proposition 5. The first claims follow from arguments given in the proof of Proposition 4. To establish that $Z = \sum_{j=1}^n z_j$ is decreasing in γ, suppose (for the purpose of contradiction) that it is not. Summing the expression in (39) for $dz_i/d\gamma$ across the m active information sources and re-arranging, total attention Z is increasing in γ if and only if

$$\sum_{j=1}^m \frac{\xi_j(K-\xi_j)}{\kappa_j^2} > -(1-\gamma)\frac{dK}{d\gamma}\sum_{j=1}^m \frac{\xi_j}{\kappa_j^2}. \tag{44}$$

Inspecting (35), note that Z is the argument of the $c'(\cdot)$ term. Hence if Z is increasing in γ then the squared term must be decreasing in γ. This is so if and only if

$$\sum_{j=1}^m \frac{K-\xi_j}{\kappa_j^2} < -(1-\gamma)\frac{dK}{d\gamma}\sum_{j=1}^m \frac{1}{\kappa_j^2}. \tag{45}$$

Combining the two inequalities of (44) and (45) gives the single inequality

$$\begin{aligned}
\sum_{i=1}^m\sum_{j=1}^m \frac{\xi_j(K-\xi_j)}{\kappa_i^2\kappa_j^2} &> \sum_{i=1}^m\sum_{j=1}^m \frac{\xi_j(K-\xi_i)}{\kappa_i^2\kappa_j^2} \\
\Leftrightarrow \quad \sum_{i\neq j} \frac{\xi_j(K-\xi_j)+\xi_i(K-\xi_i)}{\kappa_i^2\kappa_j^2} &> \sum_{i\neq j} \frac{\xi_j(K-\xi_i)+\xi_i(K-\xi_j)}{\kappa_i^2\kappa_j^2} \\
\Leftrightarrow \quad 0 &> \sum_{i\neq j} \frac{\xi_i^2+\xi_j^2-2\xi_i\xi_j}{\kappa_i^2\kappa_j^2} = \sum_{i\neq j}\frac{(\xi_i-\xi_j)^2}{\kappa_i^2\kappa_j^2}.
\end{aligned} \tag{46}$$

The final expression is positive, which generates the desired contradiction. $\square$

Proof of Proposition 6. Setting $C(z) = \zeta\sum_{i=1}^n z_i$, $K_i = K$ where $1/K = \sqrt{\zeta}\sum_{i=1}^n \hat{\psi}_i$. Algebra yields

$$w_i = \frac{\sqrt{\zeta}}{1-\gamma}\frac{\max\{(K-\xi_i),0\}}{\kappa_i^2} \quad \text{and} \quad \sigma_i^2 \equiv \kappa_i^2 + \frac{\xi_i^2}{z_i} = \kappa_i^2\left[\frac{K-\gamma\xi_i}{K-\xi_i}\right], \tag{47}$$

where the solution for σ_i^2 applies and is needed only for $i \leq m$. Hence

$$\mathrm{var}[a_\ell \,|\, \theta] = \sum_{i=1}^m w_i^2\sigma_i^2 = \frac{\zeta}{(1-\gamma)^2}\sum_{i=1}^m \frac{(K-\xi_i)(K-\gamma\xi_i)}{\kappa_i^2}. \tag{48}$$

Given the cost assumptions, the equation (35) determining K becomes

$$\zeta\left(\sum_{j=1}^m \frac{K-\xi_j}{(1-\gamma)\kappa_j^2}\right)^2 = 1 \quad \Rightarrow \quad K = \bar{\xi} + \frac{1-\gamma}{\sqrt{\zeta}\sum_{j=1}^m(1/\kappa_j^2)} \quad \text{where} \quad \bar{\xi} \equiv \frac{\sum_{j=1}^m(\xi_j/\kappa_j^2)}{\sum_{j=1}^m(1/\kappa_j^2)}. \tag{49}$$

Substituting K back into $\mathrm{var}[a_\ell \,|\, \theta]$ yields, after some algebraic simplification,

$$\begin{aligned}
\mathrm{var}[a_\ell \,|\, \theta] &= \zeta\sum_{i=1}^m \frac{1}{\kappa_i^2}\left(\frac{\bar{\xi}-\xi_i}{1-\gamma} + \frac{1}{\sqrt{\zeta}\sum_{j=1}^m(1/\kappa_j^2)}\right)\left(\frac{\bar{\xi}-\xi_i}{1-\gamma} + \frac{1}{\sqrt{\zeta}\sum_{j=1}^m(1/\kappa_j^2)} + \xi_i\right) \\
&= \bar{\xi}\sqrt{\zeta} + \frac{1}{\sum_{j=1}^m(1/\kappa_j^2)} + \zeta\left(\frac{\gamma}{(1-\gamma)^2}\right)\sum_{i=1}^m \frac{(\xi_i-\bar{\xi})^2}{\kappa_i^2}.
\end{aligned} \tag{50}$$

This is increasing in γ if $\gamma > -1$, which is a maintained parameter restriction of the model. Turning to the pairwise covariance between players' actions,

$$\operatorname{cov}[a_\ell, a_{\ell'} \,|\, \theta] = \sum\nolimits_{i=1}^{n} w_i^2 \kappa_i^2 = \frac{\zeta}{(1-\gamma)^2} \sum\nolimits_{i=1}^{m} \frac{(K-\xi_i)^2}{\kappa_i^2}$$
$$= \frac{\zeta}{(1-\gamma)^2} \sum\nolimits_{i=1}^{m} \frac{1}{\kappa_i^2} \left(\frac{1-\gamma}{\sqrt{\zeta} \sum_{j=1}^{m} (1/\kappa_j^2)} - (\xi_i - \bar{\xi}) \right)^2 = \frac{1}{\sum_{j=1}^{m} (1/\kappa_j^2)} + \frac{\zeta}{(1-\gamma)^2} \sum\nolimits_{i=1}^{m} \frac{(\xi_i - \bar{\xi})^2}{\kappa_i^2}, \tag{51}$$

where K has been substituted as before. By inspection, this covariance is increasing in γ. This covariance is the variance of the average action, conditional on the state θ: $\operatorname{cov}[a_\ell, a_{\ell'} \,|\, \theta] = \operatorname{var}[\bar{a} \,|\, \theta]$. The variance of a player's action conditional on this average is also readily calculated:

$$\operatorname{var}[a_\ell \,|\, \bar{a}, \theta] = \operatorname{var}[a_\ell \,|\, \theta] - \operatorname{var}[\bar{a} \,|\, \theta] = \bar{\xi}\sqrt{\zeta} - \frac{\zeta}{1-\gamma} \sum\nolimits_{i=1}^{m} \frac{(\xi_i - \bar{\xi})^2}{\kappa_i^2}, \tag{52}$$

and this is decreasing in γ. The correlation coefficient of action across players (conditional on θ) is

$$\hat{\rho}_{\ell\ell'} \equiv \frac{\operatorname{cov}[a_\ell, a_{\ell'} \,|\, \theta]}{\operatorname{var}[a_\ell \,|\, \theta]} = \frac{\frac{1}{\sum_{j=1}^{m}(1/\kappa_j^2)} + \frac{\zeta}{(1-\gamma)^2} \sum_{i=1}^{n} \frac{(\xi_i - \bar{\xi})^2}{\kappa_i^2}}{\bar{\xi}\sqrt{\zeta} + \frac{1}{\sum_{j=1}^{m}(1/\kappa_j^2)} + \zeta\left(\frac{\gamma}{(1-\gamma)^2}\right) \sum_{i=1}^{m} \frac{(\xi_i - \bar{\xi})^2}{\kappa_i^2}} = \frac{(1-\gamma)^2 + B}{(A+1)(1-\gamma)^2 + \gamma B}$$
$$\text{where} \quad A \equiv \bar{\xi}\sqrt{\zeta} \sum\nolimits_{j=1}^{m} (1/\kappa_j^2) \quad \text{and} \quad B \equiv \zeta \sum\nolimits_{i=1}^{m} \frac{(\xi_i - \bar{\xi})^2}{\kappa_i^2} \sum\nolimits_{j=1}^{m} \frac{1}{\kappa_j^2}. \tag{53}$$

Differentiating with respect to γ:

$$\frac{d\hat{\rho}_{\ell\ell'}}{d\gamma} = -\frac{B^2 - (1-\gamma)^2 B - 2(1-\gamma)AB}{((A+1)(1-\gamma)^2 + \gamma B)^2} < 0 \quad \Leftrightarrow \quad B < (1-\gamma)^2 + 2(1-\gamma)A. \tag{54}$$

For $1-\gamma$ small enough, $m = 1$ and so $B = 0$ and so this inequality holds. Fixing A and B, the inequality strengthens as γ falls. The only remaining case is when m increases following a fall in γ, so that A and B both change. However, straightforward but long and tedious algebraic manipulations confirm that such an increase in m serves to strengthen the inequality. □

Proof of Proposition 7. Write $\psi_i \equiv 1/\sigma_i^2$ where $\sigma_i^2 = \kappa_i^2 + (\xi_i^2/z_i)$ for the precision of the ith signal. The precision of a player's posterior beliefs about θ is $\sum_{j=1}^{m} \psi_j$ and $\operatorname{var}[\operatorname{E}[\theta \,|\, x_\ell] \,|\, \theta] = 1/\sum_{j=1}^{m} \psi_j$. The proposition's claim, therefore, is that $\sum_{j=1}^{m} \psi_j$ is decreasing in γ. Taking σ_i^2 from (47), differentiating ψ_i with respect to γ, substituting in the derivative of K with respect to γ obtained from differentiating the expression for K stated in (49), and re-arranging yields

$$\frac{d\psi_i}{d\gamma} = -\frac{\xi_i}{\kappa_i^2} \frac{\xi_i - \bar{\xi}}{(K - \gamma\xi_i)^2}. \tag{55}$$

Hence $\sum_{j=1}^{m} \psi_j$ is decreasing in γ if and only if

$$\sum\nolimits_{j=1}^{m} \frac{\xi_j}{\kappa_j^2} \frac{\xi_j - \bar{\xi}}{(K-\gamma\xi_j)^2} > 0 \;\Leftrightarrow\; \sum\nolimits_{j=1}^{m} \frac{\xi_j^2}{\kappa_j^2} \frac{1}{(K-\gamma\xi_j)^2} > \bar{\xi} \times \sum\nolimits_{j=1}^{m} \frac{\xi_j}{\kappa_j^2} \frac{1}{(K-\gamma\xi_j)^2}$$
$$\Leftrightarrow \quad \sum\nolimits_{k=1}^{m} \frac{1}{\kappa_k^2} \sum\nolimits_{j=1}^{m} \frac{\xi_j^2}{\kappa_j^2} \frac{1}{(K-\gamma\xi_j)^2} > \sum\nolimits_{k=1}^{m} \frac{\xi_k}{\kappa_k^2} \sum\nolimits_{j=1}^{m} \frac{\xi_j}{\kappa_j^2} \frac{1}{(K-\gamma\xi_j)^2}. \tag{56}$$

This inequality involves two products, the jkth elements of which cancel from both sides whenever $j = k$. Consider $j \neq k$. Collecting together the terms on either side in a typical such jkth element, a sufficient condition for the above inequality is that, for all j and k,

$$\frac{1}{\kappa_j^2\kappa_k^2}\left[\frac{\xi_j^2}{(K-\gamma\xi_j)^2}+\frac{\xi_k^2}{(K-\gamma\xi_k)^2}\right] > \frac{1}{\kappa_j^2\kappa_k^2}\left[\frac{\xi_j\xi_k}{(K-\gamma\xi_j)^2}+\frac{\xi_j\xi_k}{(K-\gamma\xi_k)^2}\right]$$
$$\Leftrightarrow \quad \frac{\xi_j(\xi_j-\xi_k)}{(K-\gamma\xi_j)^2} > \frac{\xi_k(\xi_j-\xi_k)}{(K-\gamma\xi_k)^2}. \quad (57)$$

Suppose first that $\xi_j > \xi_k$, then dividing by the (positive) common element simplifies this inequality to $\xi_j(K-\gamma\xi_k)^2 > \xi_k(K-\gamma\xi_j)^2$. Multiplying out, cancelling the common component and collecting terms again simplifies further to $(\xi_j-\xi_k)K^2 > (\xi_j-\xi_k)\gamma^2\xi_j\xi_k$. Given that $\xi_j > \xi_k$ has been assumed, the first term on each side can be cancelled and the result is true if $K > \xi_j$ (for $\gamma > -1$, which is assumed throughout). But, since $z_j > 0$ for such j, K is certainly larger than ξ_j. Finally, when $\xi_j < \xi_k$, the penultimate two inequalities both reverse (returning exactly the same final inequality) and the result holds once more, since $K > \xi_k$.

Set $\zeta = 1$ without loss of generality. The covariance of interest is

$$\operatorname{cov}[\mathrm{E}[\theta\,|\,x_\ell],\mathrm{E}[\theta\,|\,x_{\ell'}]\,|\,\theta] = \frac{\sum_{j=1}^m\psi_j^2\,\mathrm{E}[(x_{j\ell}-\theta)(x_{j\ell'}-\theta)\,|\,\theta]}{(\sum_{j=1}^m\psi_j)^2} = \frac{\sum_{j=1}^m\psi_j^2\kappa_j^2}{(\sum_{j=1}^m\psi_j)^2} = \frac{\sum_{j=1}^m\psi_j\rho_j}{(\sum_{j=1}^m\psi_j)^2}, \quad (58)$$

where the second equality follows from independence across information sources, the third by definition, and where $\rho_i \equiv \psi_i\kappa_i^2$. Simplifying notation and differentiating with respect to γ gives

$$\frac{d\,\mathrm{cov}}{d\gamma} = \frac{1}{(\sum_{j=1}^m\psi_j)^2}\sum_{j=1}^m\left(\frac{d\psi_j}{d\gamma}\rho_j+\psi_j\frac{d\rho_j}{d\gamma}\right) - \frac{2}{(\sum_{j=1}^m\psi_j)^3}\sum_{j=1}^m\frac{d\psi_j}{d\gamma}\sum_{j=1}^m\psi_j\rho_j$$
$$= \frac{2}{(\sum_{j=1}^m\psi_j)^2}\sum_{j=1}^m\frac{d\psi_j}{d\gamma}\left(\rho_j-\rho\right), \quad (59)$$

where $\rho \equiv \sum_{j=1}^m\psi_j\rho_j/\sum_{j=1}^m\psi_j$. Therefore, the covariance increases with γ if and only if the final term above is positive. From (55), $d\psi_i/d\gamma < 0$ if and only if $\xi_i > \bar{\xi}$. Again from (55),

$$\rho_i = \psi_i\kappa_i^2 = \frac{K-\xi_i}{K-\gamma\xi_i} \quad \text{and so} \quad \rho_i > \rho_j \quad \Leftrightarrow \quad \xi_i < \xi_j \quad (60)$$

is confirmed by straightforward algebra. Now, the differential of the covariance can be written

$$\frac{d\,\mathrm{cov}}{d\gamma} = \sum_{\xi_j<\bar{\xi}}\underbrace{\frac{d\psi_j}{d\gamma}}_{+\mathrm{ve}}(\rho_j-\rho) + \sum_{\xi_j>\bar{\xi}}\underbrace{\frac{d\psi_j}{d\gamma}}_{-\mathrm{ve}}(\rho_j-\rho) > \sum_{\xi_j<\bar{\xi}}\frac{d\psi_j}{d\gamma}(\bar{\rho}-\rho) + \sum_{\xi_j>\bar{\xi}}\frac{d\psi_j}{d\gamma}(\bar{\rho}-\rho), \quad (61)$$

where $\bar{\rho} = (K-\bar{\xi})/(K-\gamma\bar{\xi})$. Thus, collecting together the terms in the summation again,

$$\frac{d\,\mathrm{cov}}{d\gamma} > 0 \quad \text{if} \quad (\bar{\rho}-\rho)\sum_{j=1}^m\frac{d\psi_j}{d\gamma} > 0 \quad \Leftrightarrow \quad \bar{\rho} < \rho, \quad (62)$$

where the latter statement follows from Proposition 7. Recall $\sqrt{c'} = 1$, and so, using (49) for K,

$$\bar{\rho} = \frac{K-\bar{\xi}}{K-\gamma\bar{\xi}} = \frac{1}{1+\sum_{j=1}^m\xi_j/\kappa_j^2}, \quad \text{and} \quad \rho = \frac{\sum_{j=1}^m\psi_j\rho_j}{\sum_{j=1}^m\psi_j} \quad (63)$$

by definition. So $\rho > \bar{\rho}$ if and only if $\sum_{j=1}^{m}\psi_j\rho_j(1+\sum_k \xi_k/\kappa_k^2) > \sum_{j=1}^{m}\psi_j$. Rearranging, this occurs if and only if $\sum_{j=1}^{m}\psi_j(\rho_j(1+\sum_k \xi_k/\kappa_k^2)-1) > 0$. Using the definitions for ρ_i and for K from (49),

$$\rho > \bar{\rho} \quad \Leftrightarrow \quad \sum\nolimits_{j=1}^{m}\psi_j\frac{\bar{\xi}-\xi_j}{K-\gamma\xi_j} > 0 \quad \Leftrightarrow \quad \sum\nolimits_{j=1}^{m}\frac{1}{\kappa_j^2}\frac{(K-\xi_j)(\bar{\xi}-\xi_j)}{(K-\gamma\xi_j)^2} > 0$$
$$\Leftrightarrow \quad \sum\nolimits_{k=1}^{m}\frac{\xi_k}{\kappa_k^2}\sum\nolimits_{j=1}^{m}\frac{1}{\kappa_j^2}\frac{K-\xi_j}{(K-\gamma\xi_j)^2} > \sum\nolimits_{k=1}^{m}\frac{1}{\kappa_k^2}\sum\nolimits_{j=1}^{m}\frac{\xi_j}{\kappa_j^2}\frac{K-\xi_j}{(K-\gamma\xi_j)^2}. \quad (64)$$

The jkth terms cancel when $j = k$. For $j \neq k$, collect together the relevant jkth terms, so that a sufficient condition for the above inequality to hold is that, for all $j \neq k$,

$$\frac{1}{\kappa_j^2\kappa_k^2}\left[\frac{\xi_k(K-\xi_j)}{(K-\gamma\xi_j)^2}+\frac{\xi_j(K-\xi_k)}{(K-\gamma\xi_k)^2}\right] > \frac{1}{\kappa_j^2\kappa_k^2}\left[\frac{\xi_j(K-\xi_j)}{(K-\gamma\xi_j)^2}+\frac{\xi_k(K-\xi_k)}{(K-\gamma\xi_k)^2}\right]$$
$$\Leftrightarrow \quad \frac{(\xi_k-\xi_j)(K-\xi_j)}{(K-\gamma\xi_j)^2} > \frac{(\xi_k-\xi_j)(K-\xi_k)}{(K-\gamma\xi_k)^2}. \quad (65)$$

Suppose initially that $\xi_k > \xi_j$, so that the first term of the numerator cancels, then this reduces to

$$\frac{K-\xi_j}{(K-\gamma\xi_j)^2} > \frac{K-\xi_k}{(K-\gamma\xi_k)^2} \quad (66)$$

whenever $\xi_k > \xi_j$. In other words, $(K-\xi)/(K-\gamma\xi)^2$ must be decreasing in ξ. Now

$$\frac{d}{d\xi}\frac{K-\xi}{(K-\gamma\xi)^2} = \frac{1}{(K-\gamma\xi)^2}\left(\frac{2\gamma(K-\xi)}{K-\gamma\xi}-1\right) < 0$$
$$\Leftrightarrow \quad K-\gamma\xi > 2\gamma(K-\xi) \quad \Leftrightarrow \quad \gamma < \frac{K}{2K-\xi}, \quad (67)$$

which is only satisfied for all ξ if $\gamma < \frac{1}{2}$ (because $\xi_i < K$ for all $z_i > 0$ as usual). It remains to be shown that the covariance is also increasing in γ when $n \leq 3$. From (59), and multiplying out ρ,

$$\frac{d\,\mathrm{cov}}{d\gamma} > 0 \quad \Leftrightarrow \quad \sum\nolimits_{k=1}^{m}\psi_k\sum\nolimits_{j=1}^{m}\frac{d\psi_j}{d\gamma}\rho_j > \sum\nolimits_{k=1}^{m}\psi_k\rho_k\sum\nolimits_{j=1}^{m}\frac{d\psi_j}{d\gamma}. \quad (68)$$

This inequality involves two products. The jkth terms cancel when $j = k$. For $j \neq k$, collect together the relevant jkth terms, so that a sufficient condition for the inequality is, for all $j \neq k$,

$$\psi_k\frac{d\psi_j}{d\gamma}\rho_j + \psi_j\frac{d\psi_k}{d\gamma}\rho_k > \psi_k\rho_k\frac{d\psi_j}{d\gamma} + \psi_j\rho_j\frac{d\psi_k}{d\gamma} \quad \Leftrightarrow \quad \psi_k(\rho_j-\rho_k)\frac{d\psi_j}{d\gamma} > \psi_j(\rho_j-\rho_k)\frac{d\psi_k}{d\gamma}. \quad (69)$$

Suppose initially that $\xi_k > \xi_j$ so that $\rho_k < \rho_j$, then this last inequality simplifies to $\psi_k \times d\psi_j/d\gamma > \psi_j \times d\psi_k/d\gamma$. Now recall that $d\psi_i/d\gamma < 0$ if and only if $\xi_i > \bar{\xi}$. If $n = 2$, then $\xi_k > \bar{\xi} > \xi_j \Rightarrow d\psi_k/d\gamma < 0 < d\psi_j/d\gamma$ (as $\bar{\xi}$ is a weighted sum of ξ_k and ξ_j). Therefore, since $\psi_i > 0$ for all i, the required inequality holds for sure (the case $\xi_k < \xi_j$ follows in exactly the same way).

For $n = 3$, note that there are two possibilities: $\xi_1 < \xi_2 < \bar{\xi} < \xi_3$ and $\xi_1 < \bar{\xi} < \xi_2 < \xi_3$. (Ties cause no problems, as may be readily verified.) The latter of these two cases may be dealt with by reference to (69) alone. Recall that $\rho_i = \psi_i\kappa_i^2$ and hence the inequality in (69) may be written

$$\frac{d\,\mathrm{cov}}{d\gamma} > 0 \quad \text{if} \quad \rho_k(\rho_j-\rho_k)\frac{d\rho_j}{d\gamma} > \rho_j(\rho_j-\rho_k)\frac{d\rho_k}{d\gamma} \quad \text{for all } j \neq k. \quad (70)$$

Now, $d\rho_j/d\gamma < d\rho_k/d\gamma$ if $\xi_j > \xi_k > \bar{\xi}$. To confirm this, differentiate an appropriate function $\rho(\xi)$, constructed from (55) in an obvious way, with respect to ξ, giving

$$\frac{d}{d\xi}\frac{d\rho(\xi)}{d\gamma} \equiv -\frac{d}{d\xi}\frac{\xi(\xi-\bar{\xi})}{(K-\gamma\xi)^2} = -\frac{\xi(K-\gamma\bar{\xi})+(\xi-\bar{\xi})K}{(K-\gamma\xi)^3}, \tag{71}$$

which is certainly negative if $\xi > \bar{\xi}$. Therefore each pair of comparisons required in (70) between jkth elements is satisfied: $\xi_3 > \xi_2 > \bar{\xi} > \xi_1 \Rightarrow \rho_3 < \rho_2 < \rho_1$ and $d\rho_3/d\gamma < d\rho_2/d\gamma < 0M d\rho_1/d\gamma$.

For the former case, $\xi_1 < \xi_2 < \bar{\xi} < \xi_3$, this comparison cannot be done. Instead, again note that $\rho_1 > \rho_2 > \rho_3$ and that $d\psi_1/d\gamma > 0$, $d\psi_2/d\gamma > 0$, and $d\psi_3/d\gamma < 0$. Now, writing out the full expression in (68), and eliminating the common jkth terms with $j = k$ from both sides,

$$\begin{aligned}\frac{d\,\mathrm{cov}}{d\gamma} > 0 \quad\Leftrightarrow\quad & \psi_1\underbrace{(\rho_2-\rho_1)}_{-\mathrm{ve}}\underbrace{\frac{d\psi_2}{d\gamma}}_{+\mathrm{ve}} + \psi_2\underbrace{(\rho_3-\rho_2)}_{-\mathrm{ve}}\underbrace{\frac{d\psi_3}{d\gamma}}_{-\mathrm{ve}} + \psi_3\underbrace{(\rho_1-\rho_3)}_{+\mathrm{ve}}\underbrace{\frac{d\psi_1}{d\gamma}}_{+\mathrm{ve}} \\ > \quad & \psi_2\underbrace{(\rho_2-\rho_1)}_{-\mathrm{ve}}\underbrace{\frac{d\psi_1}{d\gamma}}_{+\mathrm{ve}} + \psi_3\underbrace{(\rho_3-\rho_2)}_{-\mathrm{ve}}\underbrace{\frac{d\psi_2}{d\gamma}}_{+\mathrm{ve}} + \psi_1\underbrace{(\rho_1-\rho_3)}_{+\mathrm{ve}}\underbrace{\frac{d\psi_3}{d\gamma}}_{-\mathrm{ve}}. \end{aligned} \tag{72}$$

The problem lies in the very first term. The other left-hand-side terms are positive and the right-hand-side terms are negative. Hence it suffices to show that the absolute value of the first left-hand-side term is smaller than that of the last right-hand-side term. Note that ψ_1 is identical (and positive) in both terms. $|\rho_1-\rho_3| = \rho_1-\rho_3 > \rho_1-\rho_2 = |\rho_2-\rho_1|$, so the second element in the right-hand side term exceeds that in the left-hand side term. It remains to be shown that

$$\left|\frac{d\psi_3}{d\gamma}\right| = -\frac{d\psi_3}{d\gamma} > \frac{d\psi_2}{d\gamma} = \left|\frac{d\psi_2}{d\gamma}\right| \quad\Leftrightarrow\quad \frac{\xi_3}{\kappa_3^2}\frac{\xi_3-\bar{\xi}}{(K-\gamma\xi_3)^2} > \frac{\xi_2}{\kappa_2^2}\frac{\bar{\xi}-\xi_2}{(K-\gamma\xi_2)^2}. \tag{73}$$

Since $\xi_3 > \xi_2$, it is sufficient to show that this holds for $\gamma > 0$ (the case of $\gamma < \frac{1}{2}$ has already been proved for all n), this latter inequality will hold if

$$\frac{\xi_3-\bar{\xi}}{\kappa_3^2} > \frac{\bar{\xi}-\xi_2}{\kappa_2^2} \quad\Leftrightarrow\quad \frac{\xi_2}{\kappa_2^2}+\frac{\xi_3}{\kappa_3^2} > \bar{\xi}\left(\frac{1}{\kappa_2^2}+\frac{1}{\kappa_2^2}\right). \tag{74}$$

This inequality holds: $\bar{\xi} = \sum_{i=1}^{3}\xi_i/\kappa_i^2 \big/ \sum_{i=1}^{3} 1/\kappa_i^2$ by definition and any weighted average of the two higher ξ_is will always be larger than the smallest ξ_i, completing the proof for $n = 3$. $\square$

Proof of Lemma 2. There is now a proper prior $\theta \sim N(\bar{\theta}, \varpi^2)$, and so a proper prior about the $n \times 1$ vector $\bar{x}$. Abusing notation, so that $\bar{\theta}$ is also an $n \times 1$ vector with identical entries equal to the scalar $\bar{\theta}$, and ϖ^2 is an $n \times n$ matrix with every element equal to the scalar ϖ^2, $\bar{x} \sim N(\bar{\theta}, \varpi^2 + K)$ where $K \equiv \mathrm{diag}[\kappa^2]$ is an $n \times n$ diagonal matrix with ith diagonal element κ_i^2. Now, by definition, signal observations are distributed $x_\ell \,|\, \bar{x} \sim N(\bar{x}, \Xi)$ where $\Xi = \mathrm{diag}[\xi^2/z_\ell]$. Standard Bayesian updating yields a posterior $\bar{x} \,|\, x_\ell \sim N(\mathrm{E}[\bar{x} \,|\, x_\ell], \mathrm{var}[\bar{x} \,|\, x_\ell])$ where

$$\begin{aligned} \mathrm{E}[\bar{x} \,|\, x_\ell] &= ((\varpi^2+K)^{-1}+\Xi^{-1})^{-1}((\varpi^2+K)^{-1}\mu+\Xi^{-1}x_\ell) \\ \text{and}\quad \mathrm{var}[\bar{x} \,|\, x_\ell] &= ((\varpi^2+K)^{-1}+\Xi^{-1})^{-1}. \end{aligned} \tag{75}$$

Dealing with the individual components, $\Xi^{-1} = \mathrm{diag}[z_\ell/\xi^2]$. The Sherman-Morrison formula for updating rank-one updates of invertible matrices yields

$$(\varpi^2+K)^{-1} = K^{-1} - \frac{K^{-1}\varpi\varpi' K^{-1}}{1+\varpi' K^{-1}\varpi}, \tag{76}$$

where notation is again abused: ϖ is an $n \times 1$ vector (as well as the corresponding scalar). Let

$$\varrho \equiv \frac{K^{-1}\varpi}{\sqrt{1+\varpi' K^{-1}\varpi}} \quad \text{so that} \quad (\varpi^2 + K)^{-1} + \Xi^{-1} = K^{-1} + \Xi^{-1} - \varrho\varrho'. \tag{77}$$

The determinant of the posterior covariance matrix $\det[\mathrm{var}[\bar{x} \,|\, x_\ell]]$ is required. This is the reciprocal of $\det[(\varpi^2 + K)^{-1} + \Xi^{-1}]$. So $\det[\mathrm{var}[\bar{x} \,|\, x_\ell]] = (\det[K^{-1} + \Xi^{-1} - \varrho\varrho'])^{-1}$. Applying the matrix determinant lemma for rank-one updates yields

$$\begin{aligned}\det[K^{-1} + \Xi^{-1} - \varrho\varrho'] &= (1 - \varrho'(K^{-1} + \Xi^{-1})^{-1}\varrho)\det[K^{-1} + \Xi^{-1}] \\ &= \left(1 - \frac{\varpi' K^{-1}(K^{-1} + \Xi^{-1})^{-1}K^{-1}\varpi}{1 + \varpi' K^{-1}\varpi}\right)\det[K^{-1} + \Xi^{-1}].\end{aligned} \tag{78}$$

Considering each of these components in turn,

$$K^{-1} + \Xi^{-1} = \mathrm{diag}\left[\frac{1}{\kappa_i^2} + \frac{z_{i\ell}}{\xi_i^2}\right] \quad \Rightarrow \quad \det[K^{-1} + \Xi^{-1}] = \prod\nolimits_{i=1}^{n}\left(\frac{1}{\kappa_i^2} + \frac{z_{i\ell}}{\xi_i^2}\right). \tag{79}$$

Also,

$$\begin{aligned}(K^{-1} + \Xi^{-1})^{-1} &= \mathrm{diag}\left[\frac{1}{(1/\kappa_i^2) + (z_{i\ell}/\xi_i^2)}\right] \\ \text{and} \quad K^{-1}(K^{-1} + \Xi^{-1})^{-1}K^{-1} &= \mathrm{diag}\left[\frac{(1/\kappa_i^2)^2}{(1/\kappa_i^2) + (z_{i\ell}/\xi_i^2)}\right].\end{aligned} \tag{80}$$

Noting that pre- and post-multiplication by the vector ϖ essentially sums the elements of the quadratic-form matrix while scaling by ϖ^2,

$$\varpi' K^{-1}(K^{-1} + \Xi^{-1})^{-1}K^{-1}\varpi = \varpi^2 \sum\nolimits_{i=1}^{n} \frac{(1/\kappa_i^2)^2}{(1/\kappa_i^2) + (z_{i\ell}/\xi_i^2)}. \tag{81}$$

Similarly, $1 + \varpi' K^{-1}\varpi = 1 + \varpi^2 \sum_{i=1}^{n} \frac{1}{\kappa_i^2}$. Now consider $\det[\mathrm{var}[\bar{x}]]$. This can be obtained by eliminating Ξ^{-1} from (78), or equivalently ignoring $z_{i\ell}$ terms in (79), and hence in (80) and (81):

$$\frac{1}{\det[\mathrm{var}[\bar{x}]]} = \det[K^{-1} - \varrho\varrho'] = \frac{1}{1 + \varpi^2 \sum_{i=1}^{n}(1/\kappa_i^2)} \prod\nolimits_{i=1}^{n}\left(\frac{1}{\kappa_i^2}\right). \tag{82}$$

Comparing the posterior and prior determinants, and following some simplification,

$$\frac{\det[\mathrm{var}[\bar{x}]]}{\det[\mathrm{var}[\bar{x} \,|\, x_\ell]]} = \left(1 + \varpi^2 \sum_{i=1}^{n} \frac{1}{\kappa_i^2 + (\xi_i^2/z_{i\ell})}\right) \prod\nolimits_{i=1}^{n} \frac{1/\kappa_i^2 + z_{i\ell}/\xi_i^2}{1/\kappa_i^2}, \tag{83}$$

which yields (19), as required. □

Proof of Lemma 3. Follows from the arguments used to derive Lemma 1, and subsequent arguments and discussion given in the main text. □

Proof of Proposition 8. Differentiate $L(z)$ from (25) with respect to z to obtain:

$$\begin{aligned}L'(z) &= \frac{[(1-\gamma)(\varpi^2 + \kappa^2)z + 1](1-\gamma)^2\varpi^2\kappa^2 - (1-\gamma)^2\varpi^2((1-\gamma)\kappa^2 z + 1)(\varpi^2 + \kappa^2)}{[(1-\gamma)(\varpi^2 + \kappa^2)z + 1]^2} + \frac{c(\kappa^2 + \varpi^2)}{1 + (\kappa^2 + \varpi^2)z} \\ &= \frac{c(\kappa^2 + \varpi^2)}{1 + (\kappa^2 + \varpi^2)z} - \frac{((1-\gamma)\varpi^2)^2}{[(1-\gamma)(\varpi^2 + \kappa^2)z + 1]^2} = \frac{Q(z)}{[(1-\gamma)(\varpi^2 + \kappa^2)z + 1]^2[1 + (\kappa^2 + \varpi^2)z]} \\ &\text{where} \quad Q(z) \equiv c(\kappa^2 + \varpi^2)[(1-\gamma)(\varpi^2 + \kappa^2)z + 1]^2 - ((1-\gamma)\varpi^2)^2[1 + (\kappa^2 + \varpi^2)z].\end{aligned} \tag{84}$$

The sign of $L'(z)$ is determined by the $Q(z)$, which is a convex quadratic. Any interior minimiser of $L(z)$ must satisfy $Q(z) = 0$ where $Q(z)$ is increasing. The unique candidate for this is the largest root of $Q(z)$. (There is also an interior maximiser at the smaller root of $Q(z)$. If this is positive, then, given the discussion in the text, it could form part of an equilibrium.) There is also the possibility of a boundary solution at $z = 0$, which requires $Q(0) > 0$. Evaluating at $z = 0$,

$$Q(0) = (\kappa^2 + \varpi^2)\left[c - \frac{((1-\gamma)\varpi^2)^2}{\kappa^2 + \varpi^2}\right] \quad \text{and} \quad Q'(0) = (\kappa^2 + \varpi^2)^2\left[2(1-\gamma)c - \frac{((1-\gamma)\varpi^2)^2}{\kappa^2 + \varpi^2}\right]. \tag{85}$$

From (26), recall that $\bar{c} = ((1-\gamma)\varpi^2)^2/(\kappa^2 + \varpi^2)$.

Begin by supposing that $\gamma < \frac{1}{2}$. From (85), if $c > \bar{c}$ then $Q(0) > 0$ and so $L(z)$ is locally increasing at zero. Hence, $z = 0$ is a local minimiser. $2(1-\gamma)c > \bar{c}$, and so $Q'(0) > 0$, which means that the quadratic $Q(z)$ is increasing for all positive z. This means that there can be no positive solution to $Q(z)$, and so $z = 0$ is the unique minimiser, and so there is a unique equilibrium. If $c < \bar{c}$ then $Q(0) < 0$, and so $z = 0$ cannot be an equilibrium. Moreover, beginning from $Q(0) < 0$ there is only one positive solution to $Q(z) = 0$ and so only one local minimiser of $Q(z)$.

Next suppose that $\gamma > \frac{1}{2}$. If $c < \bar{c}$ then $Q(0) < 0$, and the argument in the previous paragraph applies: there is a unique and positive local minimiser of $Q(z)$, and so a unique equilibrium. Similarly, if $c > \bar{c}/(2(1-\gamma)) > \bar{c}$ then there is a solution at $z = 0$ (because $Q(0) > 0$) but no positive solution (because $Q'(0) > 0$). The remaining case is when $\bar{c} < c < \bar{c}/(2(1-\gamma))$. $Q(z)$ begins at $Q(0) > 0$ (so there is a local minimiser at $z = 0$) but is decreasing, and so there is the possibility of two roots of $Q(z)$ at positive values of z. The existence of such roots is guaranteed when $Q(0)$ is close enough to zero, which holds when c is close enough to $\bar{c}$. □

Proof of Proposition 9. Without loss, set $\xi_i^2 = 1$ for all i. It is useful to perform the change of variables $y_i = z_i\kappa_i^2$. There is no loss in doing so, since choosing y_i is equivalent to choosing z_i. With this change in hand, the mutual information between x_ℓ and $\bar{x}$ satisfies

$$2\,\mathcal{I}(x_\ell, \bar{x}) = \log(1 + \varpi^2 Y) + \sum\nolimits_{i=1}^{n} \log(1 + y_i) \quad \text{where} \quad Y \equiv \sum\nolimits_{i=1}^{n} \frac{1}{\kappa_i^2}\frac{y_i}{1 + y_i}. \tag{86}$$

Differentiating this with respect to both y_i and $(1/\kappa_i^2)$ yields

$$\frac{\partial C(y)}{\partial y_i} = \frac{c'(\mathcal{I}(x_\ell, \bar{x}))}{2}\left[\frac{\varpi^2}{\kappa_i^2(1 + \varpi^2 Y)(1 + y_i)^2} + \frac{1}{1 + y_i}\right] \tag{87}$$

$$\text{and} \quad \frac{\partial C(y)}{\partial(1/\kappa_i^2)} = \frac{c'(\mathcal{I}(x_\ell, \bar{x}))}{2}\left[\frac{\varpi^2 y_i}{(1 + \varpi^2 Y)(1 + y_i)}\right]. \tag{88}$$

The change of variable ensures that the beauty-contest loss function $L^\star(z)$ becomes

$$L^\star(y) = \frac{1}{\Psi} \quad \text{where} \quad \Psi \equiv \frac{1}{\varpi^2} + \sum\nolimits_{i=1}^{n} \frac{y_i}{\kappa_i^2(1 + (1-\gamma)y_i)}. \tag{89}$$

Differentiating this with respect to both y_i and $(1/\kappa_i^2)$ yields

$$\frac{\partial L^\star(y)}{\partial y_i} = -\frac{1}{\Psi^2}\frac{1}{\kappa_i^2(1 + (1-\gamma)y_i)^2} \quad \text{and} \quad \frac{\partial L^\star(y)}{\partial(1/\kappa_i^2)} = -\frac{1}{\Psi^2}\frac{y_i}{(1 + (1-\gamma)y_i)}. \tag{90}$$

For signals that are ignored, the underlying accuracy $(1/\kappa_i^2)$ has no first-order effect. For signals that are used, so that $y_i > 0$, the relevant first-order condition holds, and so

$$\frac{\partial L^\star(y)}{\partial y_i} + \frac{\partial C(y)}{\partial y_i} = 0 \quad \Leftrightarrow \quad \frac{c'(\mathcal{I}(x_\ell, \bar{x}))}{2} \left[\frac{\varpi^2}{(1+\varpi^2 Y)} + \kappa_i^2(1+y_i)\right] = \frac{1}{\Psi^2}\left(\frac{1+y_i}{1+(1-\gamma)y_i}\right)^2. \tag{91}$$

The proposition claims that an increase in underlying signal accuracy improves payoffs. It is sufficient to show this locally. That is

$$\frac{\partial L^\star(y)}{\partial(1/\kappa_i^2)} + \frac{\partial C(y)}{\partial(1/\kappa_i^2)} < 0 \quad \Leftrightarrow \quad \frac{c'(\mathcal{I}(x_\ell, \bar{x}))}{2} \left[\frac{\varpi^2}{(1+\varpi^2 Y)(1+y_i)}\right] < \frac{1}{\Psi^2(1+(1-\gamma)y_i)}. \tag{92}$$

Dividing each side of this inequality by the relevant terms from the first-order condition, cancelling terms, and re-arranging yields

$$\frac{\varpi^2}{1+\varpi^2 Y} < \frac{1+(1-\gamma)y_i}{1+y_i}\left[\frac{\varpi^2}{1+\varpi^2 Y} + \kappa_i^2(1+y_i)\right] \quad \Leftrightarrow \quad \frac{\gamma y_i}{1+y_i}\frac{\varpi^2}{1+\varpi^2 Y} < \kappa_i^2(1+(1-\gamma)y_i). \tag{93}$$

A sufficient condition for this inequality is it holds when $\gamma = 1$, or

$$\frac{y_i}{1+y_i}\frac{\varpi^2}{1+\varpi^2 Y} < \kappa_i^2 \quad \Leftrightarrow \quad \frac{1}{\kappa_i^2}\frac{y_i}{1+y_i} < \frac{1}{\varpi^2} + Y. \tag{94}$$

Given the definition of Y this is automatically satisfied. Hence, at any equilibrium players' payoffs are increased by increasing the precision of the underlying signals. This implies that the payoff-maximising equilibrium must use the signals with the best underlying accuracy. To see why, note that if it didn't (so a signal with lower accuracy was used, while one with higher accuracy was not) then switching the use of the two signals is equivalent to raising the accuracy of the in-use signal. This increases payoffs, and so demonstrates that there was a profitable deviation. □

REFERENCES

ALLEN, F., S. MORRIS, AND H. S. SHIN (2006): "Beauty Contests and Iterated Expectations in Asset Markets," *The Review of Financial Studies*, 19(3), 720–752.

AMATO, J. D., S. MORRIS, AND H. S. SHIN (2002): "Communication and Monetary Policy," *Oxford Review of Economic Policy*, 18(4), 495–503.

ANGELETOS, G.-M., AND J. LA'O (2009): "Incomplete Information, Higher-Order Beliefs and Price Inertia," *Journal of Monetary Economics*, 56(S1), S19–S37.

ANGELETOS, G.-M., AND A. PAVAN (2004): "Transparency of Information and Coordination in Economies with Investment Complementarities," *American Economic Review: AEA Papers and Proceedings*, 94(2), 91–98.

——— (2007): "Efficient Use of Information and Social Value of Information," *Econometrica*, 75(4), 1103–1142.

——— (2009): "Policy with Dispersed Information," *Journal of the European Economic Association*, 7(1), 11–60.

BAERISWYL, R. (2007): "Central Bank's Action and Communication," *University of Munich*, unpublished manuscript.

BAERISWYL, R., AND C. CORNAND (2006): "Can Opacity of a Credible Central Bank Explain Excessive Inflation?," *University of Munich*, unpublished manuscript.

——— (2007): "Monetary Policy and its Informative Value," *University of Munich*, unpublished manuscript.

CALVÓ-ARMENGOL, A., AND J. DE MARTÍ BELTRAN (2007): "Communication Networks: Knowledge and Decision," *American Economic Review: AEA Papers and Proceedings*, 97(2), 86–91.

——— (2009): "Information Gathering in Organizations: Equilibrium, Welfare, and Optimal Network Structure," *Journal of the European Economic Association*, 7(1), 116–161.

CALVÓ-ARMENGOL, A., J. DE MARTÍ BELTRAN, AND A. PRAT (2009): "Endogenous Communication in Complex Organizations," *unpublished manuscript*, Universitat Pompeu Fabra.

COVER, T. M., AND J. A. THOMAS (2006): *Elements of Information Theory*. Wiley, New York, second edn.

CRÉMER, J. (1990): "Common Knowledge and the Co-ordination of Economic Activities," in *The Firm as a Nexus of Treaties*, ed. by M. Aoki, B. Gustafsson, and O. E. Williamson, chap. 3, pp. 53–76. Sage, London.

——— (1993): "Corporate Culture and Shared Knowledge," *Industrial and Corporate Change*, 2(1), 351–386.

DEWAN, T., AND D. P. MYATT (2008): "The Qualities of Leadership: Direction, Communication, and Obfuscation," *American Political Science Review*, 102(3), 351–368.

——— (2011): "On the Rhetorical Strategies of Leaders: Speaking Clearly, Standing Back, and Stepping Down," forthcoming in the *Journal of Theoretical Politics*.

GOLDSTEIN, I., E. OZDENOREN, AND K. YUAN (2011): "Learning and Complementarities in Speculative Attacks," *Review of Economic Studies*, 78(1), 263–292.

HELLWIG, C. (2005): "Heterogeneous Information and the Welfare Effects of Public Information Disclosure," *UCLA*, unpublished manuscript.

HELLWIG, C., AND L. VELDKAMP (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies*, 76(1), 223–251.

KEYNES, J. M. (1936): *The General Theory of Employment, Interest and Money*. Macmillan, London.

LUCAS, R. E. (1973): "Some International Evidence on Output-Inflation Tradeoffs," *American Economic Review*, 63(3), 326–334.

MACKAY, D. J. C. (2003): *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, Cambridge.

MAĆKOWIAK, B., AND M. WIEDERHOLT (2009): "Optimal Sticky Prices under Rational Inattention," *American Economic Review*, 99(3), 769–803.

MARSCHAK, J., AND R. RADNER (1972): *Economic Theory of Teams*. Yale University Press, New Haven.

MORRIS, S., AND H. S. SHIN (2002): "Social Value of Public Information," *American Economic Review*, 92(5), 1521–1534.

——— (2005): "Central Bank Transparency and the Signal Value of Prices," *Brookings Papers on Economic Activity*, No. 2(2005), 1–43.

MYATT, D. P., AND C. WALLACE (2008): "On the Sources and Value of Information: Public Announcements and Macroeconomic Policy," *Department of Economics Working Paper*, 411, Oxford University.

OZDENOREN, E., AND K. YUAN (2008): "Feedback Effects and Asset Prices," *Journal of Finance*, LXIII(4), 1939–1975.

PHELPS, E. S. (1970): "Introduction: The New Microeconomics in Employment and Inflation Theory," in *Microeconomic Foundations of Employment and Inflation Theory*, ed. by E. S. Phelps et. al., pp. 1–23. Norton, New York.

RADNER, R. (1962): "Team Decision Problems," *The Annals of Mathematical Statistics*, 33(2), 857–881.

SIMS, C. A. (1998): "Stickiness," *Carnegie-Rochester Conference Series on Public Policy*, 49(1), 317–356.

——— (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50(3), 665–690.

——— (2005): "Rational Inattention: A Research Agenda," *Princeton University*, unpublished manuscript.

——— (2006): "Rational Inattention: Beyond the Linear-Quadratic Case," *American Economic Review: AEA Papers and Proceedings*, 96(2), 158–163.

——— (2010): "Rational Inattention and Monetary Economics," forthcoming chapter for the Handbook of Monetary Economics.