D Determining the support of the signal and the message-to-action distribution

Each player $i$ receives a message $s_i$ from a rich enough message space $S_i$. The player can freely correlate her message with the fundamental. She, then, chooses a (generally mixed) strategy that she will play upon receiving each realization $s_i$ of her signal. The action space of player $i$ is denoted by $A_i = \mathbb{R}$ and the state space is denoted by $\Theta = \mathbb{R}$.

In order to determine the optimal signal support that a player $i$ will use, fix $m_{-i}$ the strategy profile of player $i$’s opponents.\footnote{The profile consists of a tuple $m_j = (s_j, a_j)$ of a fundamental-to-signal distribution $s_j$ and a signal-to-action distribution $a_j$ for each player $j$.} Let $s_i$ be the message that player $i$ obtained after having chosen channel $s_i$. Given $s_i$, player $i$ forms a posterior belief about $\theta$ and from this belief and $m_{-i}$ (by pushing forward), a posterior belief about the value of $\bar{a}$.\footnote{Formally it should be $\bar{a}_{-i}$ (i.e. the average action of all players excluding player $i$) but as the contribution of a single player to the average action of a continuum of players is zero, $\bar{a}_{-i} = \bar{a}$.} From these beliefs, player $i$ forms expectations $\theta^i(s_i; s_i)$ and $\bar{a}^i(s_i; s_i, m_{-i})$ on the respective variables.

**Lemma 11.** In a best response of player $i$, almost all messages $s_i \in S_i$ have the following property: there exists a unique action $a^i \in A_i$ such that $\Pr(a^i|s_i) = 1$.

**Proof.** For simplicity of exposition it is assumed that $P_{s_i|\theta}$ is described by a PDF $q_i(\cdot|\theta)$ and $P_{a_i|s_i}$ is described by a PDF $\lambda_i(\cdot|s_i)$. The proof is essentially the same in the more general case.

Let $p_i(\cdot|s_i; s_j)$ denote the PDF of the posterior belief that player $i$ has about the fundamental (conditional on $i$ receiving message $s_i$ while using channel $s_j$). It is calculated
by Bayes’s rule:
\[
p^i(\theta|s_i; s_i) = \frac{q_i(s_i|\theta)p(\theta)}{\int_\Theta q_i(s_i|\theta) \, p(\theta) \, d\theta}.
\]
Given player \(j\)’s strategy and upon receiving message \(s_i\), player \(i\) forms a posterior belief about \(j\)’s message applying Bayes’s rule once more. This is given by:
\[
q^i_j(s_j|s_i; s_i, s_j) = \int_\Theta q_j(s_j|\theta) \, p^i(\theta|s_i; s_i) \, d\theta.
\]
So, player \(i\)’s posterior belief about player \(j\)’s action is
\[
\lambda^j_i(a_j|s_i; s_i, m_j) = \int_{S_j} \lambda_j(a_j|s_j) \, q^i_j(s_j|s_i; s_i, s_j) \, ds_j
\]
and player \(i\)’s expectation of player \(j\)’s action is
\[
a^j_i(s_i; s_i, m_j) = \int_{A_j} a_j \, \lambda^j_i(a_j|s_i; s_i, m_j) \, da_j.
\]
Therefore, player \(i\)’s expectation of the average action of her opponents is
\[
\tilde{a}^i_{-i}(s_i; s_i, m_{-i}) = \tilde{a}^i(s_i; s_i, m_{-i}) = \int_{\Theta} \int_{A_j} a_j \, \lambda^j_i(a_j|s_i; s_i, m_j) \, da_j \, dj
\]
Notice that this expectation is equal to \(i\)’s expectation over the population-wide average action \(\tilde{a}\) as player \(i\)’s action cannot affect the mean action in a continuum population.
Finally, player \(i\)’s expectation of the value of the fundamental is
\[
\tilde{\theta}^i(s_i) = \int_\Theta \theta \, p^i(\theta|s_i; s_i) \, d\theta.
\]
Any costs player \(i\) has spent on acquiring information are sunk at the time she observes \(s_i\). So, her expected utility at that point is calculated by
\[
E_i(u_i|s_i; s_i, m_{-i}) = -(1-\gamma) \int_\Theta (a_i - \theta)^2 \, p^i(\theta|s_i; s_i, m_{-i}) \, d\theta - \int_\Theta \gamma (a_i - \tilde{a}(\theta))^2 \, p^i(\theta|s_i; s_i, m_{-i}) \, d\theta.
\]
Given that player \(i\) maximizes expected utility, any “best” action \(b_i\) that receives positive density in a best response of player \(i\) has to satisfy the following first order condition.
\[
b_i(s_i; s_i, m_{-i}) = (1-\gamma) \int_\Theta \theta \, p^i(\theta|s_i; s_i) \, d\theta + \gamma \int_\Theta \tilde{a}(\theta) \, p^i(\theta|s_i; s_i) \, d\theta
\]
As long as the integrals appearing in the right-hand side of the above equation are well-defined, there is a unique value of $b_i$ that satisfies the above condition. Therefore, a best response should put all probability to that action, i.e., $\Pr(a^i|s_i) = 1$ with $a^i$ given by the above equation.

Moreover, there should be a unique message that maps to each action.

**Lemma 12.** In a best response of player $i$, almost all actions $a_i \in A_i$ have the following property: there exists a unique message $s^{a_i} \in S_i$ such that $\Pr(a_i|s^{a_i}) = 1$.

**Proof.** Consider two strategies: $m_i = (s_i, a_i)$ under which each action has a unique message that maps to it (through $a_i$) and $m'_i = (s'_i, a'_i)$ under which a set of actions of positive measure (under the measure induced on $A_i$ by $m'_i$) have multiple messages that map to them. For each action $a_i \in A_i$, denote by $S'(a_i)$ the set of messages that map to $a_i$ i.e. $S'(a_i) = \{s'_i \in S'_i : \Pr(a_i = a_i|s'_i) = 1\}$ and by $s^{a_i}$ the (unique) message that maps to $a_i$ under $a_i$. Note that $S'(a_i)$ should be nonempty for almost all $a_i$ as a result of Lemma 11. Let also $q(s^{a_i}|\theta) = \sum_{s' \in S'(a_i)} q'(s|\theta)$. It is clear that the expected value of $-(1 - \gamma)(a_i - \theta)^2 - \gamma(a_i - \bar{a})^2$ from the two strategies will be the same as they induce the same probability distribution on $A_i$ for the same values of $\theta$. It is also true that $I(\theta, a'_i) > I(\theta, a_i)$ due to the convexity of mutual information in $q$ (see Fozunbal, McLaughlin, and Schafer 2005). Therefore, $m'_i$ is more expensive to player $i$ than $m_i$ and thus not an optimal choice.

From Lemmas 11 and 12, the action part of the strategy $a_i$ can be summarized by a bijection from $S_i$ to $A_i$ such that $a_i$ gives probability one to a unique action $a_i$ for each message $s_i \in S_i$, and for each action $a_i \in \mathbb{R}$ there exists a unique message for which $\Pr(a_i|s_i) = 1$. Thus, the message space should have the same cardinality as the action space. This should happen even if some of these messages are never used. Of course, if any of the messages is not to be used, this would immediately mean that the corresponding action would never be used by player $i$. So, player $i$’s message space can be reduced to be a space equinumerous with $A_i = \mathbb{R}$. Since signals are important only as far as they prescribe probabilities over actions, the exact choice of the message space
will not change players’ actions as long as it has the same cardinality as \( \mathbb{R} \), and \( a_i \) can be described by a bijection, as explained above.

### E Properties of SMFE

Both (12) and (13) are second-order nonlinear differential equations to which a general solution cannot be obtained in closed form. However, it is still possible to describe some properties that any SMFE should have. Two sets of results are provided. Firstly the behavior of \( b(\cdot) \) and \( \bar{a}(\cdot) \) close to \( \pm \infty \) as well as their ex-ante expected value are examined. Secondly, a relation between the variance of the best action and the variance of the fundamental \( \sigma^2 \) is established. Together with Propositions 9 and 5, these results relate to Bergemann and Morris (2013)’s agenda to identify moments of equilibrium distributions of variables.

**Proposition 13.** Let \( r \) be an SMFE with average action function \( \bar{a}(\cdot) \) and best action function \( b(\cdot) \). Then

1. \[ \int_{-\infty}^{+\infty} b(\theta)p(\theta) d\theta = \int_{-\infty}^{+\infty} \bar{a}(\theta)p(\theta) d\theta = \int_{-\infty}^{+\infty} \theta p(\theta) d\theta = \bar{\theta} \]
2. \[ \lim_{\theta \to +\infty} b(\theta) - \theta < 0 \quad \text{and} \quad \lim_{\theta \to +\infty} \bar{a}(\theta) - \theta < 0 \]
3. \[ \lim_{\theta \to -\infty} b(\theta) - \theta > 0 \quad \text{and} \quad \lim_{\theta \to -\infty} \bar{a}(\theta) - \theta > 0 \]

**Proof.** See Appendix F.1.

Bergemann and Morris (2013) derive the result of Item 1 for a normally distributed prior and point out that it should hold for any prior. This result says that the “mean error” that players make is zero. They will miss their target \( \theta \) most of the time but on average they should be correct. Items 2 and 3 show that—in equilibrium—players are biased towards the center of the distribution and take actions with extreme values (as compared to the ex-ante mean of the distribution) less often.

**Proposition 14.** Let \( (p, \gamma, \mu) \) be a beauty contest with flexible information acquisition that admits an SMFE and \( \sigma^2_b \) be the variance of the best action \( b \) in that SMFE. Then

1. \[ \sigma^2 - \sigma^2_b = \frac{\mu \gamma}{1 - \gamma} + \text{Var}(\theta - b) \]
2. $\mu < \frac{2(1-\gamma)}{1+\gamma}\sigma^2$.

**Proof.** See Appendix F.2.

An immediate consequence of Item 1 of Proposition 14 is that in an SMFE the best action is more concentrated than the fundamental ($\sigma^2 \geq \sigma_b^2$). Item 2 gives an upper bound to the value of $\mu$. This upper bound is not tight: what one can say for sure is that if $\mu$ exceeds this value, then $(p, \gamma, \mu)$ has no SMFE.

### E.1 Heterogeneous costs

**Proposition 15.** Let a population with costs distributed according to $M \in \Delta([\mu_{\min}, \mu_{\max}])$ play a beauty contest with prior $p$ and coordination motive $\gamma$. Let also $g$ be the distribution of the best action in an equilibrium. If $R_{\max} \equiv \mathcal{F}_\xi^{-1}[\exp(\mu_{\max}\pi^2\xi^2)\hat{g}(\xi)]$ is the PDF of a probability distribution, then all players follow continuous strategies in equilibrium, the inverse of the best action is given by

$$\theta(b) = b - \frac{\hat{\mu}\gamma}{2(1-\gamma)} d\left(\log(g(b))\right),$$

and $g(\cdot) = p(\theta(\cdot))\theta'(\cdot)$. In fact, this equilibrium is an SMFE.

**Proof.** By Bochner's theorem (Bochner 1933; Rudin 1962, p.19), since

$$R_{\max} \equiv \mathcal{F}_\xi^{-1}[\exp(\mu_{\max}\pi^2\xi^2)\hat{g}(\xi)]$$

is the PDF of a probability distribution, $R_{\max}$, given by $R_{\max}(\xi) = \exp(\mu_{\max}\pi^2\xi^2)\hat{g}(\xi)$, is a positive definite function. Begin by the following observation:

$$\exp(\mu\pi^2\xi^2)\hat{g}(\xi) = \exp(-(\mu_{\max} - \mu)\pi^2\xi^2)\exp(\mu_{\max}\pi^2\xi^2)\hat{g}(\xi) = \exp(-(\mu_{\max} - \mu)\pi^2\xi^2)\hat{R}_{\max}.$$

So, $\exp(\mu\pi^2\xi^2)\hat{g}(\xi)$ is a positive definite function as the product of two positive definite functions (notice that $\exp(-(\mu_{\max} - \mu)\pi^2\xi^2)$ is the Fourier transform of the normal distribution $N(0, (\mu_{\max} - \mu)/2)$ and, thus, positive definite).

Therefore, according to Lemma 2, all players’ best responses are in continuous strategies. Moreover, the expected action of a player with information cost $\mu_i$ conditional on the best action being $b$ is given by:

$$\alpha_i(b, \mu_i) = b + \frac{\mu_i}{2} d\left(\log(g(b))\right).$$
So, the average action conditional on $b$ is given by

$$a(b) = \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} \alpha_i(b, \mu_i) \, dM(\mu_i) = \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} b + \frac{\mu_i}{2} \frac{d}{db} (\log(g(b))) \, dM(\mu_i)$$

$$= b + \frac{\bar{\mu}}{2} \frac{d}{db} (\log(g(b)))$$

The final part follows from an argument identical to the one used in Proposition 6.

\[\square\]

F Proofs

F.1 Proof of Proposition 13

Item 1:
From the definition of $b$ it is clear that $\mathbb{E}(\bar{a}) = \mathbb{E}(b) \Leftrightarrow \mathbb{E}(b) = \bar{\theta}$. So, all that needs to be shown is that $\mathbb{E}(\bar{a}) = \mathbb{E}(b)$.

Let $r_i$ be the best response to $b(\cdot)$ and begin from the left-hand side of the above equation:

$$\mathbb{E}(\bar{a}) = \int_{-\infty}^{+\infty} \bar{a}(\theta) p(\theta) \, d\theta = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} a_i r_i(a_i | \theta) \, da_i p(\theta) \, d\theta = \int_{-\infty}^{+\infty} a_i \int_{-\infty}^{+\infty} r_i(a_i | \theta) p(\theta) \, d\theta \, da_i = \int_{-\infty}^{+\infty} R(a_i) a_i \, da_i = \mathbb{E}(a_i)$$

From the proof of Section C.6, since $r_i$ is a best response to $b(\cdot)$ one gets that $\mathbb{E}(a_i) = \mathbb{E}(b)$ so $\mathbb{E}(\bar{a}) = \mathbb{E}(b)$. \[\square\]

Items 2 and 3:
It is first shown that $\lim_{\theta \to +\infty} p(\theta)/b'(\theta) = 0$. Begin by integrating condition (12).

$$\int_{-\infty}^{+\infty} b(\theta) p(\theta) \, d\theta = \int_{-\infty}^{+\infty} \theta p(\theta) \, d\theta + \frac{\mu \gamma}{2(1-\gamma)} \int_{-\infty}^{+\infty} \frac{1}{b'(\theta)} \frac{d}{d\theta} \left( \log \left( \frac{p(\theta)}{b'(\theta)} \right) \right) p(\theta) \, d\theta$$

The above expression is well-defined in a smooth, monotone, full-support profile. From the proof of Item 1, $\int_{-\infty}^{+\infty} b(\theta) p(\theta) \, d\theta = \int_{-\infty}^{+\infty} \theta p(\theta) \, d\theta$. So:

$$\int_{-\infty}^{+\infty} p(\theta) \frac{d}{d\theta} \left( \log \left( \frac{p(\theta)}{b'(\theta)} \right) \right) \, d\theta = \int_{-\infty}^{+\infty} \frac{d}{d\theta} \left( \frac{p(\theta)}{b'(\theta)} \right) \, d\theta = \left[ \frac{p(\theta)}{b'(\theta)} \right]_{\theta=-\infty}^{\theta=+\infty} = 0.$$
So, \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = -\lim_{\theta \to -\infty} \frac{p(\theta)}{b'(\theta)} \) and the two limits exist. As \( p \) is a PDF, it has to be that \( \lim_{\theta \to +\infty} p(\theta) = \lim_{\theta \to -\infty} p(\theta) = 0 \). Now focus on \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)}. \) There are three possible cases:

(i) \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = +\infty. \)

As \( \lim_{\theta \to +\infty} p(\theta) = 0 \), it has to be that \( \lim_{\theta \to +\infty} b'(\theta) = 0 \). But then, there exists a \( \theta' \) such that \( b(\theta) < \theta \) for all \( \theta > \theta' \). So, from equation (10), it has to be that \( p(\theta)/b'(\theta) \) is decreasing for all \( \theta > \theta' \). This contradicts \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = +\infty. \)

(ii) \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = l > 0. \)

In this case, there exists a \( \theta'' \) such that \( p(\theta)/b'(\theta) \geq l/2 \) for all \( \theta \geq \theta'' \). Since \( b' > 0, b \) is strictly increasing and thus \( \lim_{\theta \to +\infty} b(\theta) \) is well-defined (possibly infinite). So, for \( \theta \geq \theta'' \) it has to be that \( b'(\theta) \leq (2/l)p(\theta) \) and integrating this gives \( \int_{\theta'}^{\theta} b'(x)dx \leq (2/l)\int_{\theta'}^{\theta} p(x)dx \leq 2/l \) and so \( b(\theta) - b(\theta'') \leq 2/l. \) So, \( \lim_{\theta \to +\infty} b(\theta) \leq 2/l + b(\theta'') < +\infty. \) This contradicts that \( b(\cdot) \) is bijective.

(iii) \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = 0. \)

Since the other two cases lead to contradictions, it has to be that this is the case.

A similar argument can be made for the case where \( \theta \to -\infty. \)

So, \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = \lim_{\theta \to -\infty} \frac{p(\theta)}{b'(\theta)} = 0. \)

By solving condition (12) for \( p \) one gets

\[
p(\theta) = \frac{p(\theta')}{b'(\theta')} b'(\theta) \exp\left(\frac{2(1-\gamma)}{\mu\gamma} \int_{\theta'}^{\theta} b'(t)(b(t) - t)dt\right) \tag{50}
\]

for any \( \theta' \in \mathbb{R}. \) And so

\[
\frac{p(\theta)}{b'(\theta)} = \frac{p(\theta')}{b'(\theta')} \exp\left(\frac{2(1-\gamma)}{\mu\gamma} \int_{\theta'}^{\theta} b'(t)(b(t) - t)dt\right).
\]

Now, taking the limit for \( \theta \to +\infty: \)

\[
\lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = \frac{p(\theta')}{b'(\theta')} \exp\left(\frac{2(1-\gamma)}{\mu\gamma} \int_{\theta'}^{+\infty} b'(t)(b(t) - t)dt\right).
\]

As \( \lim_{\theta \to +\infty} \frac{p(\theta)}{b'(\theta)} = 0 \) and \( p(\theta')/b'(\theta') > 0 \) for any \( \theta' \), it has to be that

\[
\int_{\theta'}^{+\infty} b'(t)(b(t) - t)dt = -\infty
\]
for all $\theta' \in \mathbb{R}$. Clearly, as $b'(\theta) > 0$ for all $\theta$ this can happen only if $\lim_{\theta \to +\infty} b(\theta) - \theta < 0$. The same arguments for $\bar{a}$ can be given if one takes into account the definition of $b(\theta) = (1 - \gamma) \theta + \gamma \bar{a}(\theta)$. A similar argument can be given for $\theta \to -\infty$. \hfill \qed

### E2 Proof of Proposition 14

**Item 1:**
From player $i$’s point of view, and given that she knows the function $b(\cdot)$, there are two random variables: $\theta$ and $a_i$. One can define more random variables, namely $y = \mathbb{E}(a_i|\theta)$ which is the (equilibrium) average action given $\theta$ and $x = (1 - \gamma) \theta + \gamma y$, which is the best action given $\theta$. Using the variance decomposition formula for $a_i$, one obtains

$$\text{Var}(a_i) = \mathbb{E}(\text{Var}(a_i|\theta)) + \text{Var}(\mathbb{E}(a_i|\theta)) = \mathbb{E}(\text{Var}(a_i|\theta)) + \text{Var}(y).$$

Using this and from equation (47) (in the proof of Lemma 2), one gets

$$\text{Var}(x) = \frac{\mu}{2} + \text{Var}(a_i) = \frac{\mu}{2} + \mathbb{E}(\text{Var}(a_i|\theta)) + \text{Var}(y). \quad (51)$$

As $y = x/\gamma + (1 - \gamma) \theta / \gamma$,

$$\text{Var}(y) = \left(\frac{1}{\gamma}\right)^2 \text{Var}(x) + \left(\frac{1 - \gamma}{\gamma}\right)^2 \text{Var}(\theta) - \frac{2(1 - \gamma)}{\gamma^2} \text{Cov}(x, \theta). \quad (52)$$

Substituting (52) into equation (51) and after calculations, one gets

$$\gamma (\text{Var}(\theta) - \text{Var}(x)) = \frac{\mu \gamma^2}{2(1 - \gamma)} + \frac{\gamma^2}{1 - \gamma} \mathbb{E}(\text{Var}(a_i|\theta)) + \text{Var}(\theta) + \text{Var}(x) - 2 \text{Cov}(x, \theta)).$$

Now, notice that

$$\text{Var}(\theta) + \text{Var}(x) - 2 \text{Cov}(x, \theta) = \text{Var}(\theta - x)$$

and thus

$$\sigma^2 - \sigma^2_b = \frac{\mu \gamma}{2(1 - \gamma)} + \frac{\gamma}{1 - \gamma} \mathbb{E}(\text{Var}(a_i|\theta)) + \frac{1}{\gamma} \text{Var}(\theta - b) \quad (53)$$

where $\sigma^2 = \text{Var}(\theta)$ and $\sigma^2_b = \text{Var}(x)$.

Now, from the result of Proposition 5:

$$\mathbb{E}(\text{Var}(a_i|\theta)) = \frac{\mu}{2} + \frac{\mu^2}{4} \int_{-\infty}^{+\infty} \frac{d^2}{db^2} (\log(g(b)))g(b) db =$$
\[
\frac{\mu}{2} + \frac{\mu^2}{4} \left\{ \frac{d}{db} (\log(g(b))) g(b) \right\}_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} \frac{d}{db} (\log(g(b))) g'(b) \, db \right\} \Rightarrow \\
\mathbb{E}(\text{Var}(a_i | \theta)) = \frac{\mu}{2} + \frac{\mu^2}{4} \int_{-\infty}^{+\infty} \frac{d}{db} (\log(g(b))) g'(b) \, db \\
\text{(54)}
\]

Moreover, from (41) and the fact that \(\mathbb{E}(\bar{a}) = \mathbb{E}(b)\):

\[
\text{Var}(\bar{a} - b) = \frac{\mu^2}{4} \int_{-\infty}^{+\infty} \frac{d}{db} (\log(g(b))) g'(b) \, db
\]

Using this into (54) and substituting into (53):

\[
\sigma^2 - \sigma_b^2 = \frac{\mu \gamma}{2(1 - \gamma)} + \frac{\gamma}{1 - \gamma} \left( \frac{\mu}{2} + \text{Var}(\bar{a} - b) \right) + \frac{1}{\gamma} \text{Var}(\theta - b)
\]

Finally, as \((b - \bar{a}) = \frac{1 - \gamma}{\gamma} \theta - b\), one gets that \(\text{Var}(b - \bar{a}) = \frac{(1 - \gamma)^2}{\gamma^2} \text{Var}(\theta - b)\). So

\[
\sigma^2 - \sigma_b^2 = \frac{\mu \gamma}{1 - \gamma} + \text{Var}(\theta - b)
\]

Item 2:

As in an SMFE, \(r_i\) is the best response to a smooth, monotone, full-support strategy profile that induces a best action function \(b(\cdot)\), it has to be that \(\text{Var}(\theta) > \mu/2\). Using this along with the result of Proposition 14, Item 1, one gets that

\[
\frac{\mu}{2} < \sigma^2 - \frac{\mu \gamma}{1 - \gamma} - \text{Var}(\theta - b) \Rightarrow \\
\mu < \frac{2(1 - \gamma)}{1 + \gamma} \sigma^2 - \frac{2(1 - \gamma)}{1 + \gamma} \text{Var}(\theta - b) < \frac{2(1 - \gamma)}{1 + \gamma} \sigma^2
\]

\(\Box\)

G More general payoffs

In this appendix a class of games broader than the one of the main text is considered. Propositions analogous to Lemma 2 and proposition 6 are derived.

As in the main model, there is a large population of individuals. The decision problem facing each individual is a “tracking problem” à la Jung et al. (2019) but there are also strategic considerations. In particular, each individual \(i\) gets payoff that is given by

\[-u(a_i - b) - \bar{u}(b, \theta)\]
where $b$ is the “best action” and is given by $b := f(\bar{a}, \theta)$ for some function $f : \mathbb{R}^2 \to \mathbb{R}$. The partial derivative $\partial_1 f$ is positive everywhere, so that the game is a coordination game. Moreover, the best action depends monotonically on the state of the world and so, without loss of generality, $\partial_2 f > 0$ everywhere.

The function $u : \mathbb{R} \to \mathbb{R}$ is assumed to be even, analytic, and $u'(x) < 0$ for $x < 0$ whereas $u'(x) > 0$ for $x > 0$. Moreover, $\int_{-\infty}^{+\infty} \exp(-u(x)/\mu) \, dx < \infty$. Players’ payoffs are also affected by $\tilde{u} : \mathbb{R}^2 \to \mathbb{R}$ but no individual player can affect its value. As $\tilde{u}$ will not enter the individuals’ decision-making problem, no assumptions are imposed on it. This model includes beauty contests (analyzed in the main text) as well as all linear-quadratic large games as special cases. In particular, it includes all examples of Bergemann and Morris (2013) with infinite players and one-dimensional action sets.23

Individuals can acquire information about $\theta$ paying a cost linear in the reduction of Shannon entropy, as in the main model.

**Best response**

A strategy profile of player $i$’s opponents defines an average action function $\bar{a} : \mathbb{R} \to \mathbb{R}$ to which player $i$ needs to best respond. The best action that player $i$ needs to track is given by $B(\theta) := f(\bar{a}(\theta), \theta)$. A smooth, monotone, full-support profile of player $i$’s opponents requires that $B$ is strictly increasing, which imposes the condition that $\bar{a}'(\theta) > -\partial_1 f(\bar{a}(\theta), \theta)/\partial_2 f(\bar{a}(\theta), \theta)$ at all $\theta$.

Slightly abusing notation, let $\theta(\cdot)$ denote the inverse of $B(\cdot)$. Since the fundamental is distributed according to the full-support density $p$, the best action is distributed according to

$$g(b) = p(\theta(b)) \, \theta'(b)$$

which has full support. Let also $z : \mathbb{R} \to \mathbb{R}_{>0}$ be defined through

$$z(x) := \frac{\exp(-u(x)/\mu)}{\int_{-\infty}^{+\infty} \exp(-u(x)/\mu) \, dx}.$$ 

A generalization of Lemma 2 can now be formulated (see also Jung et al. 2019, Proposition 2 for this result).
**Proposition 16.** Player $i$ has a continuous best response to a smooth monotone full-support profile of her opponents if and only if

$$R_i := \mathcal{F}^{-1}[\hat{g} / \hat{z}]$$

is the PDF of a probability distribution. \hfill (55)

This continuous strategy is her unique best response and is given by

$$r_i(a_i | b) = R_i(a_i) \frac{z(a_i - b)}{g(b)}.$$

where $R_i(a_i)$ is the marginal density of action $a_i$.

**Corollary.** If player $i$’s best response is continuous, her posterior belief about the best action is given by

$$\tau(b | a_i) = z(a_i - b).$$

This is the analogue of Proposition 4. Notice that this posterior belief is fully determined by the payoff function $u(\cdot)$ and the cost $\mu$. In particular, it does not depend on the best action’s distribution $g(\cdot)$ (or, for that matter $p(\cdot)$) nor does it depend on how this best action is being calculated (the particular function $f(\cdot)$).

**Equilibrium**

The characterization of SMFE in the general case — analogous to Proposition 6 — is provided below.

**Proposition 17.** The following two statements are equivalent

(A) $\theta(\cdot)$ is the inverse of the best action function and $g(\cdot)$ is the PDF of the distribution of the best action in an SMFE.

(B) $\theta : \mathbb{R} \to \mathbb{R}$ is a strictly increasing bijection, $\mathcal{F}^{-1}[\hat{g} / \hat{z}]$ is a probability distribution,

$$b = f \left( b - \frac{(g * Z)(b)}{g(b)}, \theta(b) \right)$$

and

$$g(b) = p(\theta(b))\theta'(b).$$ \hfill (56) \hfill (57)

Where $Z(x) := x \mathcal{F}^{-1}[\log \hat{z}](x)$.  

11
### H Proofs

#### H.1 Proof of Proposition 16

Following essentially the same process as in the proof of Lemma 2, one calculates the best response of player $i$ to $g$:

$$r_i(a_i|b) = \exp\left(\frac{-\bar{u}(b, \theta_0(b))}{\mu}\right) \exp\left(\frac{-k(b)}{\mu}\right) R(a_i) \exp\left(\frac{-u(a_i-b)}{\mu}\right)$$

where $k(b)$ ($b \in \mathbb{R}$) are Lagrange multipliers. In order to have a continuous best response, the solution should also satisfy the constraints $\int_{-\infty}^{+\infty} r(a_i|b) g(b) \, db = R(a_i)$ for all $a_i \in \mathbb{R}$, which gives

$$\int_{-\infty}^{+\infty} \exp\left(\frac{-\bar{u}(b, \theta_0(b))}{\mu}\right) \exp\left(\frac{-k(b)}{\mu}\right) g(b) \exp\left(\frac{-u(a_i-b)}{\mu}\right) \, db = 1$$

or, simply,

$$\int_{-\infty}^{+\infty} G(b) \exp(-u(a_i-b)/\mu) \, db = 1 \Rightarrow (G \ast \exp(-u/\mu))(a_i) = 1 \quad \text{for all } a_i \in \mathbb{R}.$$  

Using the convolution theorem, and after calculations, one gets that

$$G(b) = \left(\int_{-\infty}^{+\infty} \exp(-u(x)/\mu) \, dx\right)^{-1} =: G,$$

i.e., a normalizing constant. Using the definition of $z$, i.e., that $z(x) = G \exp(-u(x)/\mu)$, the best response of player $i$ can be written as

$$r_i(a_i|b) = \frac{R(a_i)}{g(b)} z(a_i - b).$$

All that remains is for $R$ to be determined. This comes from the constraint $\int_{-\infty}^{+\infty} r(a_i|b) \, da_i = 1$ for (almost) all $b$ which, in turn, yields

$$\int_{-\infty}^{+\infty} z(a_i - b) R(a_i) \, da_i = g(b) \Rightarrow (R \ast z)(b) = g(b),$$

as $z$ is even. This means that $\hat{R} = \hat{g}/\hat{z}$. So, if $g$ can be deconvolved with $z$ as a kernel, then the problem has a solution in continuous strategies with best response given by

$$r_i(a_i|b) = \mathcal{F}^{-1}[\hat{g}/\hat{z}](a_i) \frac{z(a_i-b)}{g(b)}.$$

$\square$
H.2 Proof of Proposition 17

Begin by calculating $\alpha(b)$, the expected action of player $i$ conditional on the best action being $b$ when she is best-responding. This can be calculated using the property:

$$\alpha(b) = \frac{1}{-2\pi i} \left( \mathcal{F}_a[r(a|b)] \right)'(0)$$

Calculations give:

$$\mathcal{F}_a[r(a|b)](\xi) = \mathcal{F}_a\left[ \frac{R(a)\hat{z}(a-b)}{g(b)} \right](\xi) = \frac{1}{g(b)} \mathcal{F}_a[R(a)\hat{z}(a-b)](\xi)$$

$$= \frac{1}{g(b)} \left( \hat{R} \ast \mathcal{F}_a[\hat{z}(a-b)] \right)(\xi) = \frac{1}{g(b)} \left( \hat{R} \ast \exp(-2\pi ib\xi)\hat{z} \right)(\xi)$$

$$= \frac{1}{g(b)} \int_{-\infty}^{+\infty} \frac{\hat{g}(y)}{\hat{z}(y)} \exp(-2\pi ib(\xi - y))\hat{z}(\xi - y) \, dy$$

and the first derivative at $\xi = 0$ gives:

$$\left( \mathcal{F}_a[r(a|b)] \right)'(0) = \frac{1}{g(b)} \int_{-\infty}^{+\infty} \frac{\hat{g}(y)}{\hat{z}(y)} \exp(2\pi iby)(\hat{z}'(-y) - 2\pi ib\hat{z}(-y)) \, dy$$

As $\hat{z}$ is an even, purely real function, $\hat{x}$ is also even and purely real. Thus $\hat{x}'$ is a purely real, odd function. So,

$$\left( \mathcal{F}_a[r(a|b)] \right)'(0) = \frac{1}{g(b)} \left( -\mathcal{F}^{-1} \left[ \hat{g}(y)(\log \circ \hat{z})'(y) \right](b) - 2\pi ibg(b) \right)$$

and, finally,

$$\alpha(b) = b - \frac{(g \ast Z)(b)}{g(b)}$$

where $Z(x) := x \mathcal{F}^{-1} [\log \circ \hat{z}](x)$. Notice that $Z$ is fully determined by the payoff function $u$ and the information cost $\mu$.

Having calculated $\alpha$ (as was done in Lemma 9), the rest of the proof is identical to the proof of Proposition 6, using

$$b = f(\alpha(b), \theta(b))$$

as the fixed point condition. \qed
References

Bergemann, D. and S. Morris (2013). “Robust Predictions in Games With Incomplete Information”. *Econometrica* 81 (4), pp. 1251–1308. DOI: 10.3982/ECTA11105.

Bochner, S. (1933). “Monotone Funktionen, Stieltjessche Integrale und harmonische Analyse”. *Mathematische Annalen* 108 (1), pp. 378–410. DOI: 10.1007/BF01452844.

Fozunbal, M., S. W. McLaughlin, and R. W. Schafer (2005). “Capacity analysis for continuous-alphabet channels with side information, part I: A general framework”. *IEEE Transactions on Information Theory* 51 (9), pp. 3075–3085.

Jung, J., J. H. J. Kim, F. Matějka, and C. A. Sims (2019). “Discrete Actions in Information-Constrained Decision Problems”. *The Review of Economic Studies* 86 (6), pp. 2643–2667. DOI: 10.1093/restud/rdz011.

Rudin, W. (1962). *Fourier analysis on groups*. Vol. 121967. Wiley Online Library.