## Online Appendix

# Herd Design

Itai Arieli Ronen Gradwohl Rann Smorodinsky

In this online appendix we consider the more general model in which the sender's utility is state-dependent, and so the per-period utility is $v : A \times \Omega \to \mathbb{R}$. As in the state-independent case, for any prior $\mu$ such that $V^*(\mu) > v^*(\mu)$ the function $V^*(\mu)$ is supported on two priors, $\underline{\mu} \in [0, 1]$ and $\overline{\mu} \in [0, 1]$. Note that the function $v^*$ is piece-wise linear,[9] as is its concavification $V^*$.

Now, consider a prior $\mu \in [x_{q-1}, x_q]$. For every $\lambda \in [0, 1]$ and action $a \in A$ let $v(\lambda, a) = \lambda v(1, a) + (1 - \lambda)v(0, a)$. The piece-wise linearity of $V^*$ implies that one of the following five cases must hold:

**Case 1:** $V^*(\mu) = v^*(\mu)$.

**Case 2:** $(\mu, V^*(\mu))$ lies on the line joining $(x_{k-1}, v(x_k, a_k))$ with $(x_{m-1}, v(x_{m-1}, a_m))$ for some $k \leq q \leq m$.

**Case 3:** $(\mu, V^*(\mu))$ lies on the line joining $(x_k, v(x_k, a_k))$ with $(x_m, v(x_m, a_m))$ for some $k \leq q \leq m$.

**Case 4:** $(\mu, V^*(\mu))$ lies on the line joining $(x_{k-1}, v(x_{k-1}, a_k))$ with $(x_m, v(x_m, a_m))$ for some $k \leq q \leq m$.

**Case 5:** $(\mu, V^*(\mu))$ lies on the line joining $(x_k, v(x_k, a_k))$ with $(x_{m-1}, v(x_{m-1} a_m))$ for some $k \leq q \leq m$.

The necessary and sufficient conditions under which $V(\mu) = V^*(\mu)$ are determined by whether the beliefs $\underline{\mu}$ and $\overline{\mu}$, which support $V^*(\mu)$, are on the left or right endpoints of the respective intervals $J_k$ and $J_m$ on which they lie. In particular, if the supporting beliefs are on the left endpoints or on the right endpoints of both respective intervals, then the conditions and arguments are nearly identical to the state-independent case. In

[9] To see this, recall that in any interval of beliefs $[x_{i-i}, x_i]$ for $i \in \{1, \ldots, \ell\}$, the agents' optimal action is $a_i$. For any $\lambda \in [x_{i-i}, x_i]$, the sender's indirect utility is thus the linear $v^*(\lambda) = \lambda \cdot v(1, a_i) + (1 - \lambda) \cdot v(0, a_i)$.

contrast, if the supporting beliefs are on different endpoints of their respective intervals, then $V(\mu) = V^*(\mu)$ if and only if $\underline{\mu} = 0$ and $\overline{\mu} = 1$ (in which case full revelation is optimal). The following theorem tightly characterizes the conditions under which $V(\mu) = V^*(\mu)$.

**Theorem 2.** $V^*(\mu) = V(\mu)$ if and only if one of the following conditions is true:

- Case 1 holds.
- Case 2 holds and $|J_k| \geq |J_m|$ and $|J_m| \geq |J_i|$ for every integer $i \in (k, m)$.
- Case 3 holds and $|J_m| \geq |J_k|$ and $|J_k| \geq |J_i|$ for every integer $i \in (k, m)$.
- Case 4 holds with $k - 1 = 0$ and $m = l$.

The intuition for the proof of Theorem 2 is the following. If $\underline{\mu} = 0$ and $\overline{\mu} = 1$, then $V^*(\mu)$ can be attained by choosing the information structure that fully reveals the state. On the other hand, suppose that $\underline{\mu} > 0$, $\overline{\mu} < 1$, and $J_k < J_m$, and that $\underline{\mu}$ lies on the left endpoint of $J_k$ and $\overline{\mu}$ on the right endpoint of $J_m$. In this case, in order for the public belief to get sufficiently close to $\underline{\mu}$, there must be a sufficiently strong positive signal $s$. However, this signal will prevent the public belief from getting close to $\overline{\mu}$: whenever the belief starts to approach $\overline{\mu}$ from the left, an agent who obtains signal $s$ will take an action $a_r$ for $r > m$, thereby "overshooting" past the desired belief $\overline{\mu}$. The other cases, such as $\underline{\mu}$ lying on the right endpoint of $J_k$ and $\overline{\mu}$ on the left endpoint of $J_m$, suffer from the same problem.

We first adapt the proof of Proposition 1 to apply also to state-dependent utility.

*Proof.* Assume that the two conditions are satisfied for $\underline{\beta}, \overline{\beta}$. Let $\delta > 0$ be the value guaranteed by Lemma 2 and let $\underline{\varphi} = y_{m-1} - \overline{\eta}$, $\overline{\varphi} = y_k - \underline{\eta}$, $\underline{\alpha} = lr^{-1}(\underline{\varphi})$, and $\overline{\alpha} = lr^{-1}(\overline{\varphi})$.

Let $F$ be an information structure that is guaranteed in Lemma 4 such that $\underline{\alpha}_F = \underline{\alpha}$, $\overline{\alpha}_F = \overline{\alpha}$ and $2\epsilon < \delta$. We note that $\overline{\varphi} - \underline{\varphi} = y_k - \underline{\eta} + \overline{\eta} - y_{m-1}$. By assumption $\overline{\varphi} - \underline{\varphi} > |J_i|$ for every integer $i \in (k, m)$. Therefore, Lemma 2 implies that for every equilibrium of the corresponding game, $\mu_\infty \in [0, x_k) \cup (x_{m-1}, 1]$ with probability 1.

We claim that for a sufficiently small $\epsilon > 0$ it holds for any equilibrium $\sigma$ that if the public belief $\mu_t \leq x_{m-1}$, then $\mu_{t+1} < \overline{\beta}$. To see this note that by the choice of $\overline{\varphi}$ it holds that if $\lambda \in [0, 1]$ such that $lr(\lambda) \leq y_{m-1}$, then

$$lr(\lambda) + \overline{\varphi} = lr(\lambda) + y_k - \underline{\eta} \leq y_{m-1} + y_m - \overline{\eta} < y_m.$$

Hence if the public belief $\mu_t \leq y_{m-1}$, then the posterior probability $p_t$ of agent $t$ after receiving his private signal satisfies $p_t < x_m$. Hence with probability one agent $t$ plays an action $a_j$ such that $j \leq m$. Therefore Lemma 4 implies that $\mu_{t+1} \leq x_{m-1} + \epsilon$. Thus for a sufficiently small $\epsilon > 0$ we have $\mu_{t+1} \leq \overline{\beta}$.

Similarly, if $\mu_t \geq x_k$, then $\mu_t \geq \underline{\beta}$. Lemma 4 further guarantees that if $\mu_t \in (x_{m-1}, \overline{\beta}]$ and $a_t = a_m$, then $|\mu_{t+1} - \mu_t| \leq \epsilon$. Similarly if $\mu_t \in [\underline{\beta}, x_{m-1})$ and $a_t = a_k$, then $|\mu_{t+1} - \mu_t| \leq \epsilon$. This, together with the fact that all points $\lambda \in (\overline{\beta}, \underline{\beta})$ are continuation points, implies that $\mu_t$ reaches $(\underline{\beta} - \epsilon, \underline{\beta}) \cup (\overline{\beta}, \overline{\beta} + \epsilon)$ with probability one. Since $\epsilon < \delta$ whenever $\mu_t$ reaches $(\underline{\beta} - \epsilon, \underline{\beta}) \cup (\overline{\beta}, \overline{\beta} + \epsilon)$ the martingale stops. Therefore we must have that $\mu_\infty \in (\underline{\beta} - \epsilon, \underline{\beta}) \cup (\overline{\beta}, \overline{\beta} + \epsilon)$ for every equilibrium $\sigma$ as desired. Hence the pair $0 \leq \underline{\beta} < \mu < \overline{\beta} \leq 1$ is feasible.

□

We now prove Theorem 2.

*Proof.* We first show, using Proposition 1, that the conditions of Theorem 1 are sufficient. We begin with the case where $\mu \in (x_{q-1}, x_q)$ and $V^*(\mu) = v^*(\mu)$. In this case, it is optimal not to reveal any information. Assume next that $\mu = x_{q-1}$, where we can assume that $q > 1$. Let $\varepsilon$ be sufficiently small such that $\overline{\varphi}_\varepsilon = x_{q-1} + \varepsilon < x_q$. Let $\underline{\varphi} \in (x_{q-2}, x_{q-1})$ be such that $y_{q-1} - lr(\underline{\varphi}) < y_q - y_{q-1}$. By Proposition 1, the pair $(\underline{\varphi}, \overline{\varphi}_\varepsilon)$ is feasible for all sufficiently small $\varepsilon$. Let $F_\varepsilon$ be the information structure that guaranteed by Proposition 1, for which $\mu_\infty \in (\underline{\varphi} - \varepsilon/2, \underline{\varphi} + \varepsilon/2) \cup (\overline{\varphi}_\varepsilon - \varepsilon/2, \overline{\varphi}_\varepsilon + \varepsilon/2)$ with probability 1 in any equilibrium $\sigma$. Note that $E_\sigma[\mu_\infty] = \mu$, that the distance of $\mu$ from $(\underline{\varphi} - \varepsilon/2, \underline{\varphi} + \varepsilon/2)$ is bounded away from zero, and that the distance of $\mu$ from $(\overline{\varphi}_\varepsilon - \varepsilon/2, \overline{\varphi}_\varepsilon + \varepsilon/2)$ approaches zero when $\varepsilon \to 0$. Therefore, when $\varepsilon \to 0$ the probability that $\mu_\infty$ lies in $(\overline{\varphi}_\varepsilon - \varepsilon/2, \overline{\varphi}_\varepsilon + \varepsilon/2)$ approaches 1. This implies that the expected utility of the sender lies arbitrarily close to $V^*(\mu)$. Similar considerations can be applied when $\mu = x_q$. This concludes case 1.

If case 4 holds, the fully revealing information structure is optimal for the sender.

Assume that case 2 holds (case 3 is similar). We start with considering the case where $0 < x_{k-1}$ and $|J_k| > |J_m| > |J_i|$ for every integer $i \in (k, m)$. Let $\underline{\eta}^\delta = lr(x_{k-1}) + \delta$ and $\overline{\eta}^\delta = lr(x_{m-1}) + \delta$. Let $x^\delta_{k-1} = lr^{-1}(\underline{\eta}^\delta)$ and $x^\delta_{m-1} = lr^{-1}(\overline{\eta}^\delta)$. We claim that the pair $(x^\delta_{k-1}, x^\delta_{m-1})$ is feasible for all sufficiently small $\delta > 0$. To see this we show that the two conditions of Proposition 1 hold. The second condition holds since $y_k - \underline{\eta}^\delta + \overline{\eta}^\delta - y_{m-1}$ approaches $|J_k|$ as $\delta \to 0$. Similarly, the first condition holds for all sufficiently small $\delta > 0$ since $|J_m| > |J_k|$ and since $\overline{\eta}^\delta - y_{m-1} = \underline{\eta}^\delta - y_{k-1} = \delta$.

Therefore, $(x^\delta_{k-1}, x^\delta_{m-1})$ is feasible for all sufficiently small $\delta > 0$. By assumption, for every $\epsilon > 0$ there exists $F$ such that $\mu_\infty \in (x^\delta_{k-1} - \epsilon, x^\delta_{k-1} + \epsilon) \cup (x^\delta_{m-1} - \epsilon, x^\delta_{m-1} + \epsilon)$ with probability one in every equilibrium of the herding game. Note that $(x^\delta_{k-1}, x^\delta_{m-1})$ approaches $(x_{k-1}, x_{m-1})$ as $\delta$ goes to 0. Therefore since $E[\mu_\infty] = \mu$ we must have that $\mathbf{P}_\sigma\Big(\mu_\infty \in (x^\delta_{k-1} - \epsilon, x^\delta_{k-1} + \epsilon)\Big)$ approaches $\frac{x_{m-1}-\mu}{x_{m-1}-x_{k-1}}$ and $\mathbf{P}_\sigma\Big(\mu_\infty \in (x^\delta_{m-1} - \epsilon, x^\delta_{m-1} + \epsilon)\Big)$ approaches $\frac{\mu - x_{k-1}}{x_{m-1}-x_{k-1}}$ as $\delta$ and $\epsilon$ go to zero. This means that the actions on which the population cascades are $a_k$ with probability approaching $\frac{x_{m-1}-\mu}{x_{m-1}-x_{k-1}}$ and $a_m$ with probability approaching $\frac{\mu - x_{k-1}}{x_{m-1}-x_{k-1}}$. This approximates the Bayesian persuasion solution to any desired precision.

The case where $x_{k-1} = 0$ is shown similarly, by observing that $(\delta, x_{m-1} + \delta)$ is feasible for all sufficiently small $\delta > 0$.

We next show that the converse hold. Namely, we start with case 2 and show that if it holds and either $|J_i| > |J_m|$ for some integer $i \in (k, m)$ or $|J_m| > |J_k|$, then $V^*(\mu) > V(\mu)$ (the converse for case 3 is shown similarly).

Assume first that $|J_i| > |J_m|$ for some integer $i \in (k, m)$. Note that for an information structure $F$, in order for the event $\mu_\infty \in (x_{m-1} - \epsilon, x_{m-1} + \epsilon)$ to hold with positive probability in equilibrium $\sigma$ we must have that $\underline{\alpha}_F$ approaches $\frac{1}{2}$ with $\epsilon$. Alternatively, for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that if $\underline{\alpha}_F \leq \frac{1}{2} - \epsilon$, then $V(\mu) \leq V^*(\mu) - \delta(\varepsilon)$. Let $\epsilon_0 = 1/2 - lr^{-1}(-\frac{|J_i| - |J_m|}{2})$.

We consider two cases. If $F$ is such that $\underline{\alpha}_F \leq \frac{1}{2} - \epsilon_0$, then by the above we have that $V(\mu) \leq V^*(\mu) - \delta(\varepsilon_0)$. Otherwise, $\underline{\alpha}_F > \frac{1}{2} - \epsilon_0/2$ and $lr(\underline{\alpha}_F) \geq -\frac{|J_i| - |J_m|}{2}$. We note that $lr(\overline{\alpha}_F) \leq |J_m|$ for otherwise, by Lemma 1 we must have that $\mu_\infty \notin [x_{m-1}, x_m]$ with probability one and the utility for the is bounded away from $V^*(\mu)$. Hence, any point $\lambda \in [x_{i-1}, x_i]$ with $lr(\lambda) \in [y_{i-1} + \frac{|J_i| - |J_m|}{2}, y_i - |J_m|]$ is a cascade point. Thus, in any equilibrium, if the public belief reaches a point $\mu_t$ such that $lr(\mu_t) \in [y_{i-1} + \frac{|J_i| - |J_m|}{2}, y_i - |J_m|]$, then learning stops and $\mu_\infty = \mu_t$. Thus, whenever $\mu_t$ satisfies $lr(\mu_t) \geq y_{i-1} + \frac{|J_i| - |J_m|}{2}$, it cannot down-cross $lr^{-1}(y_{i-1} + \frac{|J_i| - |J_m|}{2})$ and reach $[x_{k-1}, x_k]$. Therefore if $\mu_t \in [x_{i-1}, x_i]$, then it holds with positive probability that $\mu_\infty \in [x_{i-1}, x_i]$. This demonstrates that $\mu_\infty \in [x_{i-1}, x_i]$ holds with positive probability. Hence the sender's utility is bounded away from $V^*(\mu)$.

We next show that if $|J_k| > |J_m|$, then $V(\mu) < V^*(\mu)$. As before, for every $\varepsilon > 0$ there exists a $\delta(\varepsilon)$ such that if $\underline{\alpha}_F \leq \frac{1}{2} - \epsilon$, then $V(\mu) \leq V^*(\mu) - \delta(\varepsilon)$. Let $\epsilon_0 = 1/2 - lr^{-1}(-\frac{|J_k|-|J_m|}{4})$. If $F$ is such that $\underline{\alpha}_F \leq \frac{1}{2} - \epsilon_0$, then by the above we have that $V(\mu) \leq V^*(\mu) - \delta(\varepsilon_0)$. Otherwise, $\underline{\alpha}_F > \frac{1}{2} - \epsilon_0/2$ and $lr(\underline{\alpha}_F) \geq -\frac{|J_k|-|J_m|}{4}$. We note that $lr(\overline{\alpha}_F) \leq |J_m|$ for otherwise we would have that $\mu_\infty \notin [x_{m-1}, x_m]$ with probability one and the sender's utility will be bounded away from $V^*(\mu)$. Hence, any point $\lambda \in [x_{k-1}, x_k]$ with $lr(\lambda) \in [y_{k-1} + \frac{|J_i|-|J_k|}{4}, y_k - |J_m|]$ is a cascade point.

Thus, if the public belief reaches a point $\mu_t$ such that $lr(\mu_t) \in [y_{k-1}+\frac{|J_i|-|J_k|}{4}, y_k-|J_m|]$, then learning stops and $\mu_\infty = \mu_t$. This implies that if $\mu \geq y_k - |J_m|$, then $\mu_t \geq y_k - |J_m| - \frac{|J_m|-|J_k|}{4}$ for every $t$. This is true since $lr(\underline{\alpha}_F) \geq -\frac{|J_k|-|J_m|}{4}$. Hence, the sender's utility is bounded away from $V^*(\mu)$.

Finally, we show that if Case 4 holds and either $k-1 \neq 0$ or $m \neq l$, then $V(\mu) \neq V^*(\mu)$. We show this for the case $k-1 \neq 0$ ($m \neq l$ is similar). Let $F$ be an information structure. If $F$ is such that $\underline{\alpha}_F \leq \frac{1}{2} - \epsilon$, then, by Lemma 1, in any equilibrium $\sigma$ the limit $\mu_\infty \notin [x_{k-1}, x_{k-1} + r)$, for sufficiently small $r > 0$, with probability one, and hence $V(\mu) \leq V^*(\mu) - \delta(\varepsilon_0)$. Similarly, if $F$ is such that $\overline{\alpha}_F \geq \frac{1}{2} + \epsilon$, then by Lemma 1 it holds for sufficiently small $r > 0$ that in any equilibrium $\sigma$ the limit $\mu_\infty \notin [x_m - r, x_m]$ with probability one and hence $V(\mu) \leq V^*(\mu) - \delta(\varepsilon_0)$. In contrast, if both $\underline{\alpha}_F \geq \frac{1}{2} - \epsilon$ and $\overline{\alpha}_F \leq \frac{1}{2} + \epsilon$, then for some constant $r(\epsilon) > 0$ all points $[x_{m-1} + r(\epsilon), x_m - r(\epsilon)]$ are cascade point. In addition, $r(\epsilon)$ goes to zero as $\epsilon$ goes to zero. This implies that for sufficiently small $\epsilon$ it holds that $\mu_\infty < x_m - r$ for some $r > 0$ and $V(\mu) < V^*(\mu)$. Hence in any case we have that $V(\mu) < V^*(\mu)$.

□