Supplement to "Attributes: Selective Learning and Influence"

Arjada Bardhi[*]

January 10, 2024

Contents

C Proofs and auxiliary results for section 3.2

Proof of Proposition 3. For any $p \in (0,2]$, $\sigma_p(a, \underline{a})$ and $\sigma_p(a, \bar{a})$ strictly increases in $a \in [\underline{\alpha}, \underline{a}]$, so $\psi^2(a)$ increases in a as well. Hence, $a = \underline{a}$ dominates any $a < \underline{a}$. By a similar argument, sampling $a > \bar{a}$ is suboptimal as well. So for any $p \in (0,2]$, $a^s \in [\underline{a}, \bar{a}]$.

(i) For $p = 1$, the statement follows from Proposition 2. Consider $p < 1$. The posterior variance satisfies the following: (i) $\lim_{a \downarrow \underline{a}} \partial \psi^2(a)/\partial a = -\infty$, (ii) $\lim_{a \uparrow \bar{a}} \partial \psi^2(a)/\partial a = \infty$, and (iii) ψ^2 is differentiable and weakly convex in $(\underline{a}, \bar{a})$. Therefore ψ^2 is maximized at the endpoints of $[\underline{a}, \bar{a}]$: only the two relevant attributes are optimal.

(ii) Let $p > 1$. The sign of $\partial \psi^2(a)/\partial a$ is determined by the sign of the function $h(a) := \sigma_p(a, \bar{a})(\bar{a} - a)^{p-1} - \sigma_p(\underline{a}, a)(a - \underline{a})^{p-1}$. Clearly, ψ^2 is strictly increasing at $a = \underline{a}$ because $h(\underline{a}) > 0$ and strictly decreasing at $a = \bar{a}$ because $h(\bar{a}) < 0$. Hence, $a^s \in (\underline{a}, \bar{a})$. The single-player sample a^s satisfies $h(a^s) = 0$, i.e.,

$$\left(\frac{a^s - \underline{a}}{\bar{a} - a^s} \right)^{p-1} = \frac{\sigma_p(a^s, \bar{a})}{\sigma_p(\underline{a}, a^s)}. \tag{1}$$

[*]Department of Economics, New York University. Email: arjada.bardhi@nyu.edu.

The function h has either a unique zero at $(\underline{a} + \bar{a})/2$, or three zeros, of which one is $(\underline{a} + \bar{a})/2$ and the other two are symmetric with respect to it. There exists at most one $a^s < (\bar{a} + \underline{a})/2$ because

$$\left. \frac{\partial h}{\partial a} \right|_{a=a^s} = \sigma_p(a^s, \bar{a})(\bar{a} - a^s)^{p-1} \left(\frac{1 + p\left(\left(\frac{\bar{a} - a^s}{\ell} \right)^p - 1 \right)}{\bar{a} - a^s} - \frac{1 + p\left(\left(\frac{a^s - \underline{a}}{\ell} \right)^p - 1 \right)}{a^s - \underline{a}} \right)$$

has the same sign over $(\underline{a}, (\bar{a} + \underline{a})/2)$. Hence, h is either globally decreasing or decreasing-increasing-decreasing over $(\underline{a}, \bar{a})$.

As $\ell \to 0$, the RHS of (1) goes to zero for any $a^s \in (\underline{a}, \bar{a})$, hence two single-player samples converge to $a^s \downarrow \underline{a}$ and $a^s \uparrow \bar{a}$ respectively. At any a^s such that $h(a^s) = 0$ and $a^s < (\underline{a} + \bar{a})/2$ (i.e., for which h crosses zero thrice), the function h is decreasing at a^s. Note that h is increasing in ℓ at such an a^s because

$$\left. \frac{\partial h}{\partial \ell} \right|_{a=a^s} = \frac{p}{\ell} \sigma_p(a^s, \bar{a})(\bar{a} - a^s)^{p-1} \left(\left(\frac{\bar{a} - a^s}{\ell} \right)^p - \left(\frac{a^s - \underline{a}}{\ell} \right)^p \right) > 0.$$

Moreover, the function h is decreasing in a at $a = a^s$ such that $h(a^s) = 0$ and $a^s < (\underline{a} + \bar{a})/2$. Thus, as ℓ increases the single-player sample to the left of $(\underline{a} + \bar{a})/2$ shifts to the right. By the mirror argument, the single-player sample that is strictly closer to $\bar{a}$ shifts to the left as ℓ increases.

For ℓ sufficiently large, the function h is strictly decreasing at $(\underline{a} + \bar{a})/2$. To see this, consider

$$\left. \frac{\partial h}{\partial a} \right|_{a=(\underline{a}+\bar{a})/2} = 2^{3-2p}(\bar{a} - \underline{a})^{p-2} e^{-2^{-p}\left(\frac{\bar{a} - \underline{a}}{\ell} \right)^p} \left(p\left(\left(\frac{\bar{a} - \underline{a}}{\ell} \right)^p - 2^p \right) + 2^p \right)$$

which is strictly negative for ℓ large because $((\bar{a} - \underline{a})/\ell)^p \to 0$ as $\ell \to +\infty$. Therefore, it must be that h is strictly decreasing over $(\underline{a}, \bar{a})$, hence the single-player sample is $a^s = (\underline{a} + \bar{a})/2$.

Finally, fix $\ell > 0$. As $p \downarrow 1$, the RHS of (1) converges to a strictly positive value whereas the LHS shrinks to 0 for any fixed sample. Therefore, the two single-player samples converge to $a^s \downarrow \underline{a}$ and $a^s \uparrow \bar{a}$ respectively. $\qquad\square$

Lemma C.1. *Suppose Assumption 2 holds. Fix a sample $\mathbf{a} = \{a_1, \ldots, a_k\}$, where $0 \leqslant a_1 < \ldots < a_k \leqslant 1$. For the singleton sample $\mathbf{a} = \{a_1\}$, $\tau(a_1) = \ell\left(2 - e^{-a_1/\ell} - e^{-(1-a_1)/\ell}\right)$. For $k \geqslant 2$, the sample realization $f(a_j)$ is weighted by*

$$\tau_j(\mathbf{a}) = \begin{cases} \ell\left(1 - e^{-a_1/\ell} + \tanh\left(\frac{a_2 - a_1}{2\ell}\right)\right) & \text{if } j = 1 \\ \ell\left(\tanh\left(\frac{a_j - a_{j-1}}{2\ell}\right) + \tanh\left(\frac{a_{j+1} - a_j}{2\ell}\right)\right) & \text{if } j = 2, \ldots, k-1 \\ \ell\left(1 - e^{-(1-a_k)/\ell} + \tanh\left(\frac{a_k - a_{k-1}}{2\ell}\right)\right) & \text{if } j = k. \end{cases}$$

Proof of Lemma C.1. Using the expressions for $\tau(a; \mathbf{a})$ derived in the proof of Lemma 2, we obtain: (i) if $a < a_1$, then $\tau_1(a; \mathbf{a}) = e^{-(a_1-a)/\ell}$ and $\tau_j(a; \mathbf{a}) = 0$ for all $j \neq 1$; (ii) if $a > a_k$, then

2

$\tau_k(a; \mathbf{a}) = e^{-|a_k - a|/\ell}$ and $\tau_j(a; \mathbf{a}) = 0$ for all $j \neq k$; (iii) if $a \in (a_i, a_{i+1})$ for $i = 1, \ldots, k-1$, then

$$\tau_i(a; \mathbf{a}) = \frac{e^{-(a-a_i)/\ell} - e^{-(2a_{i+1}-a_i-a)/\ell}}{1 - e^{-2(a_{i+1}-a_i)/\ell}} = \operatorname{csch}\left(\frac{a_{i+1} - a_i}{\ell}\right) \sinh\left(\frac{a_{i+1} - a}{\ell}\right),$$

$$\tau_{i+1}(a; \mathbf{a}) = \frac{e^{-(a_{i+1}-a)/\ell} - e^{-(a_{i+1}+a-2a_i)/\ell}}{1 - e^{-2(a_{i+1}-a_i)/\ell}} = \operatorname{csch}\left(\frac{a_{i+1} - a_i}{\ell}\right) \sinh\left(\frac{a - a_i}{\ell}\right),$$

and $\tau_j(a; \mathbf{a}) = 0$ for all $j \neq i, i+1$. Integrating these weights as in the Corollary 1, we obtain the sample weights stated in the Lemma. □

Proof of Proposition 4. We first establish that $a_1^s > 0$ and $a_k^s < 1$. Suppose, by contradiction, that $a_1^s = 0$. Differentiating $\psi^2(\mathbf{a})$ with respect to the leftmost attribute:

$$\left.\frac{\partial \psi^2(\mathbf{a})}{\partial a_1}\right|_{a_1=0} = 2\ell e^{-2a_1/\ell}\left(2e^{a_1/\ell} - 1\right) - 2\ell \operatorname{sech}^2\left(\frac{a_2 - a_1}{2\ell}\right)\bigg|_{a_1=0} = 2\ell\left(1 - \operatorname{sech}^2\left(\frac{a_2}{2\ell}\right)\right) > 0$$

for any $\mathbf{a} \setminus \{a_1\}$. This contradicts the optimality of $a_1^s = 0$; hence, $a_1^s > 0$. By a similar argument, $a_k^s < 1$. Therefore, the first-order approach is valid for all sample attributes.

Second, we show that for any $j \in \{2, \ldots, k\}$, the distance $a_j^s - a_{j-1}^s$ is constant in j. By the optimality of a_j^s, the first-order condition with respect to a_j^s is

$$\frac{\partial \psi^2(\mathbf{a})}{\partial a_j^s} = 2\ell\left(\operatorname{sech}^2\left(\frac{a_j^s - a_{j-1}^s}{2\ell}\right) - \operatorname{sech}^2\left(\frac{a_{j+1}^s - a_j^s}{2\ell}\right)\right) = 0$$

and the second order condition $\frac{\partial^2 \psi^2(\mathbf{a})}{\partial a_j^{s2}} < 0$ is satisfied. Hence, $a_j^s - a_{j-1}^s = a_{j+1}^s - a_j^s = (1 - a_1^s - a_k^s)/(k-1)$ for any $j = 2, \ldots, k-1$. By Lemma C.1, this implies that for any $j = 2, \ldots, k-1$, the sample weight is $\tau_j(\mathbf{a}^s) = 2\ell \tanh\left(\frac{1-a_1^s-a_k^s}{2\ell(k-1)}\right)$. Third, the first-order conditions with respect to a_1^s and a_k^s are respectively

$$e^{-a_1^s/\ell}\left(2 - e^{-a_1^s/\ell}\right) = \operatorname{sech}^2\left(\frac{1 - a_1^s - a_k^s}{2\ell(k-1)}\right)$$

$$e^{-(1-a_k^s)/\ell}\left(2 - e^{-(1-a_k^s)/\ell}\right) = \operatorname{sech}^2\left(\frac{1 - a_1^s - a_k^s}{2\ell(k-1)}\right)$$

Because the RHSs are equal, LHSs must be equal too. The LHS is of the form $x(2-x)$, which strictly increases in $x \in (0,1)$. Hence $a_1^s = 1 - a_k^s$, which implies $\tau_1(\mathbf{a}^s) = \tau_k(\mathbf{a}^s)$. This, along with $a_2^s, \ldots, a_{k-1}^s$ being equidistant, establishes part (i).

The FOC for the leftmost attribute a_1^s pins down the entire $\mathbf{a}^s$. We use the trigonometric identity $\operatorname{sech}^2(x) = 1 - \tanh^2(x) = (1 - \tanh(x))(1 + \tanh(x))$ and let $x := e^{-a_1^s/\ell}$ and $y := 1 - \tanh\left(\frac{1-a_1^s-a_k^s}{2\ell(k-1)}\right)$ to rewrite the FOC with respect to a_1^s as $x(2-x) = y(2-y)$, where $x, y \in [0, 1]$. Because $f(z) = z(2-z)$ is one-to-one for $z \in [0, 1]$, this implies that $x = y$, which, combined with the fact that $a_1^s = 1 - a_k^s$, gives the conditions in part (ii). The equation $1 - e^{-a_1^s/\ell} = \tanh\left(\frac{1-2a_1^s}{2\ell(k-1)}\right)$ has a unique solution because for $a_1^s \in (0, 1/2)$, the LHS is strictly increasing in a_1^s and it is zero for $a_1^s = 0$,

whereas RHS is strictly decreasing in a_1^s and it is zero for $a_1^s = 1/2$. Finally, invoking (11), note that

$$\tau_1(\mathbf{a}^s) = \tau_k(\mathbf{a}^s) = \ell\left(1 - e^{-a_1^s/\ell} + \tanh\left(\frac{1 - 2a_1^s}{2\ell(k-1)}\right)\right) = 2\ell \tanh\left(\frac{1 - 2a_1^s}{2\ell(k-1)}\right) = \tau_j(\mathbf{a}^s)$$

for any $j = 2, \ldots, k - 1$. This establishes part (iii). $\qquad\square$

Proof of Proposition 5. By Proposition 4(i), it is sufficient to establish that $|a_1^s - 1/2|$ strictly increases in ℓ for $k > 1$. For $k = 1$, $\mathbf{a}^s = \{1/2\}$ is unique for any $\ell > 0$. For $k > 1$, $a_1^s < 1/2$ by symmetry of $\mathbf{a}^s$. By implicit differentiation of the equation for $a_1^s(\ell)$ in (11) with respect to ℓ, $\frac{\partial a_1^s}{\partial \ell} < 0$ iff $2a_1^s(k-1) + (2a_1^s - 1)(2 - e^{-a_1^s/\ell}) < 0$. But $a_1^s < 1/(k+1)$ because $1 - e^{-a_1/\ell} - \tanh\left(\frac{1-2a_1}{2\ell(k-1)}\right)$ is strictly increasing in a_1 and strictly positive for $a_1 = 1/(k+1)$. Hence, $a_1^s < 1/(k+1) < (1 - 2a_1^s)/(k-1)$, which implies

$$\frac{a_1^s}{\ell} < \frac{1 - 2a_1^s}{2\ell(k-1)} \Leftrightarrow 2a_1^s(k-1) < 1 - 2a_1^s < (1 - 2a_1^s)(2 - e^{-a_1^s/\ell})$$

because $2 - e^{-a_1^s/\ell} > 1$. Therefore a_1^s is strictly decreasing in ℓ.

Next, we want to show that as $\ell \to 0$, $a_1^s \to 1/(k+1)$. Substituting the identity $(1 + \tanh(x))/(1 - \tanh(x)) = e^{2x}$ into equation (11), we obtain $2 - e^{-a_1^s/\ell} - e^{\frac{1 - a_1^s(k+1)}{\ell(k-1)}} = 0$. Because from part (i) $a_1^s < 1/(k+1)$, as $\ell \to 0$ we have $e^{-a_1^s/\ell} \to 0$. Therefore, as $\ell \to 0$, it must be that $e^{\frac{1 - a_1^s(k+1)}{\ell(k-1)}} \to 2$. The term $\ell(k-1) \to 0$ as $\ell \to 0$ and $(1 - a_1^s(k+1))/(\ell(k-1)) \to \ln(2)$, hence it must be that $1 - a_1^s(k+1) \to 0$ as well.

Finally, we want to show that as $\ell \to +\infty$, $a_1^s \to 1/(2k)$. Equation (11) implies

$$\lim_{\ell \to +\infty} \frac{1 - e^{-a_1^s/\ell}}{\tanh\left(\frac{1 - 2a_1^s}{2\ell(k-1)}\right)} = 1.$$

Because the numerator and the denominator converge to zero as $\ell \to +\infty$, we apply L'Hôpital's rule: $\lim_{\ell \to +\infty} \frac{\frac{a_1^s}{\ell^2}e^{-a_1^s/\ell}}{\frac{1 - 2a_1^s}{2\ell^2(k-1)}\operatorname{sech}^2\left(\frac{1 - 2a_1^s}{2\ell(k-1)}\right)} = \lim_{\ell \to +\infty} \frac{2a_1^s(k-1)}{1 - 2a_1^s}\frac{e^{-a_1^s/\ell}}{\operatorname{sech}^2\left(\frac{1 - 2a_1^s}{2\ell(k-1)}\right)} = 1$. As $\ell \to +\infty$, $e^{-a_1^s/\ell} \to 1$ and $\operatorname{sech}^2\left(\frac{1 - 2a_1^s}{2\ell(k-1)}\right) \to 1$. Hence, $\lim_{\ell \to +\infty} \frac{2a_1^s(k-1)}{1 - 2a_1^s} = 1$. This implies that $a_1^s \to 1/(2k)$ as $\ell \to +\infty$. $\qquad\square$

Calculations for Remark 1. Let $v \sim \mathcal{N}(\nu_0, \sigma_0^2)$, where $\sigma_0^2 > 0$ exogenous. The player has access to signals $f(a) = v + \xi(a)$ where the noise terms are correlated according to the Ornstein-Uhlenbeck covariance (with variance 1 and correlation $\exp(-|a_2 - a_1|/\ell)$). Hence, any two signals $(f(a_1), f(a_2))$ are correlated according to the Ornstein-Uhlenbeck covariance as well: their variance is $\sigma_0^2 + 1$ and their covariance is $\sigma_0^2 + \exp(-|a_2 - a_1|/\ell) = \sigma_0^2 + \sigma_{ou}(a_1, a_2)$. The covariance between v and any

$f(a)$ is σ_0^2. Let $d_j = a_{j+1} - a_j$. The sample weights for $\mathbf{a} = \{a_1, \ldots, a_k\}$ are

$$
\begin{pmatrix} \sigma_0^2 & \cdots & \sigma_0^2 \end{pmatrix}
\begin{pmatrix}
\sigma_0^2 + 1 & \sigma_0^2 + e^{-d_1/\ell} & \cdots & \sigma_0^2 + e^{-(d_1 + \ldots + d_{k-1})/\ell} \\
\sigma_0^2 + e^{-d_1/\ell} & \sigma_0^2 + 1 & \cdots & \sigma_0^2 + e^{-(d_2 + \ldots + d_{k-1})/\ell} \\
\vdots & \vdots & \ddots & \vdots \\
\sigma_0^2 + e^{-(d_1 + \ldots + d_{k-1})/\ell} & \sigma_0^2 + e^{-(d_2 + \ldots + d_{k-1})/\ell} & \cdots & \sigma_0^2 + 1
\end{pmatrix}^{-1} .
$$

From here we calculate the posterior variance as

$$
\psi^2(\mathbf{a}) = \sum_{j=1}^k \sum_{m=1}^k \tau_j(\mathbf{a}) \tau_m(\mathbf{a}) \left(\sigma_0^2 + e^{-|a_m - a_j|/\ell} \right) .
$$

For $k = 2$, the posterior variance simplifies to

$$
\left(\frac{4}{2 + \frac{1 + e^{-d_1/\ell}}{\sigma_0^2}} \right)^2
$$

which is strictly increasing in d_1. The optimal signals that maximize this posterior variance subject to $d_1 \in [0, 1]$ are $\mathbf{a}_2^* = \{0, 1\}$. Similarly, it is straightforward to verify that the optimal signals are $\mathbf{a}_3^* = \{0, 1/2, 1\}$ for $k = 3$, $\mathbf{a}_4^* = \{0, 1/3, 2/3, 1\}$ for $k = 4$, and so on. The player seeks to sample signals that are as weakly correlated as possible, so that the overlap between the information that they carry about v is as small as possible.

The following lemma establishes that the player's expected payoff is single-peaked in attribute correlation in a simple attribute setting that is close to the common-variance-common-correlation signal setting in Clemen and Winkler (1985). Suppose that the attribute space is finite: $\mathcal{A} = \{a_1, \ldots, a_N\}$. The player's value is $\sum_{j=1}^N f(a_j)$. The common variance is $\sigma(a, a) = 1$ and the common correlation is $\sigma(a, a') = \rho \in (-1/(N-1), 1)$ for any $a, a' \in \mathcal{A}$.

Lemma C.2. *For any sample $\mathbf{a}$ that consists of k attributes, the expected loss $\mathrm{var}[v] - \psi^2(\mathbf{a})$ is single-peaked in ρ with a maximum at $\rho^* > 0$ such that $(1 - \rho^*)^2 - k\rho^{*2} = k/(N-1)$.*

Proof of Lemma C.2. We calculate $\mathrm{var}[v]$ and $\psi^2(\mathbf{a})$ for the sample of k attributes $\mathbf{a} = \{a_1, \ldots, a_k\}$. This is without loss since attributes are identically distributed:

$$
\mathrm{var}[v] = N \mathrm{var}[f(a_1)] + 2 \binom{N}{2} \mathrm{cov}[f(a_1), f(a_2)] = N + N(N-1)\rho;
$$

$$
\psi^2(\mathbf{a}) = \mathrm{var}\left[\sum_{j=1}^k f(a_j) \right] \left(1 + (N-k)\frac{\rho}{1 + (k-1)\rho} \right)^2 = (k + k(k-1)\rho)\left(1 + (N-k)\frac{\rho}{1 + (k-1)\rho} \right)^2 .
$$

The expected payoff from sample $\mathbf{a}$ is

$$
V(\mathbf{a}) = \psi^2(\mathbf{a}) - \mathrm{var}[v] = -\frac{(1 - \rho)(N - k)((N-1)\rho + 1)}{(k-1)\rho + 1} .
$$

5

V is increasing in ρ if and only if $(1 - \rho)^2 - k\rho^2 \leqslant k/(N - 1)$. It is immediate to check that V is strictly decreasing at any $\rho \in \left(-\frac{1}{N-1}, 0\right]$. Moreover V is strictly increasing at $\rho = 1$. For $\rho > 0$, the term $(1 - \rho)^2 - k\rho^2$ is strictly decreasing in ρ. Therefore, there exists a unique ρ^* at which $(1 - \rho^*)^2 - k\rho^{*2} = \frac{k}{N-1}$: the payoff is strictly decreasing (resp., increasing) for $\rho < \rho^*$ (resp., $\rho > \rho^*$). Hence the expected loss is single-peaked with a peak at ρ^*. $\qquad\square$

D Proofs and auxiliary results for section 4.2

Proof of Proposition 10. Without loss, suppose $a_A < a_P$. Given a sample $\mathbf{a} = \{a\}$, the sample weight for player i is $\tau^i(a) := \tau^i_1(\mathbf{a}) = \sigma_p(a, a_i)$. We first establish that $a^* \geqslant a_A$ for any $p \in (0, 2]$. To the contrary, suppose $a^* < a_A$. Then as a increases in (a^*, a_A), both $\tau^P(a)$ and $\tau^A(a)$ increase. The agent's payoff strictly increases because

$$\frac{\partial V_A(a)}{\partial a} = 2\tau^P(a)\left(\frac{\partial \tau^A(a)}{\partial a} - \frac{\partial \tau^P(a)}{\partial a}\right) + 2\tau^A(a)\frac{\partial \tau^P(a)}{\partial a}.$$

This is strictly positive for $a < a^A$ since for both players, $\tau^i > 0$, $\partial \tau^i(a)/\partial a > 0$, and $\partial \tau^i(a)/\partial a$ decreases in a_i. The agent is strictly better off sampling a_A instead. Next, we establish that $a^* \notin ((a_A + a_P)/2, a_P)$ for any $p \in (0, 2]$. The agent's payoff is strictly decreasing in $a \in ((a_A + a_P)/2, a_P)$ because $\partial \tau^A(a)/\partial a < 0$, $\partial \tau^P(a)/\partial a < 0$ and $0 < \tau^A(a) < \tau^P(a)$. The agent is better off sampling $(a_P + a_A)/2$ instead. Third, we establish that $a^* \leqslant a_P$ for any $p \in (0, 2]$. Suppose, to the contrary, that $a^* > a_P$. Consider an alternative sample $\tilde{a} = a_P - (a^* - a_P)$. If available, i.e., if $\tilde{a} \in \mathcal{A}$, $\tau^P(\tilde{a}) = \tau^P(a^*)$ but $\tau^A(\tilde{a}) > \tau^A(a^*)$, hence $V_A(\tilde{a}) > V_A(a^*)$. If $\tilde{a} \notin \mathcal{A}$, it must be that $\tilde{a} < a_A$. But by the argument above, the agent strictly prefers a_A to any such $\tilde{a} < a_A$. This contradicts the optimality of a^*. Hence, these three observations imply that $a^* \in [a_A, (a_P + a_A)/2]$ for any $p \in (0, 2]$.

(i) Let $p \in (0, 1]$. For any $a \in [a_A, (a_P + a_A)/2]$ and any $p \in (0, 2]$, the agent's payoff $V_A(a)$ is strictly decreasing in a if and only if

$$\left(\frac{a_P - a}{a - a_A}\right)^{1-p} \frac{e^{\left(\frac{a_P - a}{\ell}\right)^p}}{e^{\left(\frac{a_P - a}{\ell}\right)^p} - e^{\left(\frac{a - a_A}{\ell}\right)^p}} > 1,$$

which holds because $0 < a - a_A < a_P - a$. Therefore, the agent prefers sampling a_A to sampling any $a \in [a_A, (a_P + a_A)/2]$.

(ii) Let $p \in (1, 2]$. At $a = a_A$, the agent's payoff is increasing because the LHS of the inequality in part (i) is zero. Moreover, the first-order condition that pins down the optimal sample $a^* \in (a_A, (a_P + a_A)/2)$ is

$$\frac{e^{\left(\frac{a_P - a^*}{\ell}\right)^p}}{e^{\left(\frac{a_P - a^*}{\ell}\right)^p} - e^{\left(\frac{a^* - a_A}{\ell}\right)^p}} = \left(\frac{a_P - a^*}{a^* - a_A}\right)^{p-1}.$$

As $\ell \to 0$, the LHS approaches 1. Therefore it must be that RHS approaches 1 as well, which implies

that a^* approaches $(a_P + a_A)/2$. Alternatively, as $\ell \to +\infty$, the LHS approaches $+\infty$, which implies that $a^* \to a_A$ so that the RHS approaches $+\infty$ as well.

Moreover, as $p \to 1$, the RHS of the FOC converges to 1, whereas the LHS converges to

$$\frac{e^{\frac{a_P - a^*}{\ell}}}{e^{\frac{a_P - a^*}{\ell}} - e^{\frac{a^* - a_A}{\ell}}} \geqslant 1.$$

In order for the FOC to hold, it must be that LHS also converges to 1, which implies that $a^* \to a_A$. $\qquad \square$

Proposition D.1. *If σ satisfies NAP, no sampling is optimal if and only if for any $a \in \mathcal{A}$, $\tau^P(a)/\tau^A(a) > 2$.*

Proof. Fix k. If no sampling is strictly optimal, then in particular $V_A(\{a_1\}) < 0$ for any singleton sample in $\mathcal{A}_1$, which is equivalent to $2\tau^A(a_1) - \tau^P(a_1) < 0$. Conversely, suppose Δ_A/Δ_P is sufficiently close to zero. If sample $\mathbf{a}^* = \{a_1, \ldots, a_n\}$ is optimal, by the previous argument all samples of size 1 must attain strictly negative payoff, hence $n \geq 2$. In particular, $\tau^A(a_j) < \tau^P(a_j)/2$ for all $a_j \in \mathbf{a}^*$. But then,

$$\alpha_2(\mathbf{a}^*) < \sum_{j=1}^n \tau_j^P(\mathbf{a}^*) \frac{\tau^P(a_j)}{2} = \frac{\alpha_1(\mathbf{a}^*)}{2}$$

hence $V_A(\mathbf{a}^*) < 0$ as well. This contradicts the optimality of $\mathbf{a}^*$. $\qquad \square$

Proof of Proposition 11. (i) Without loss, let $\bar{a}_A < \underline{a}_P$. Suppose $\mathbf{a}^* = \{a_1, a_2\}$ where $a_1 \in [\underline{a}_A, \bar{a}_A]$ and $a_2 \in [\underline{a}_P, \bar{a}_P]$. We show that $V_A(\{a_1, a_2\}) \leqslant V_A(\{a_1, \underline{a}_P\}) < V_A(\{\bar{a}_A\})$. Consider first the difference $\alpha_2(\{a_1, a_2\}) - \alpha_2(\{a_1\})$, which due to NAP equals

$$\tau_2^A(\mathbf{a}^*)\tau_2^P(\mathbf{a}^*)\left(1 - \sigma_{ou}^2(a_1, a_2)\right) = \left(\int_{a_1}^{\bar{a}_A} \sigma_{ou}(a, \underline{a}_P)(1 - \sigma_{ou}^2(a, a_1))\, \mathrm{d}a\right)\tau_2^P(\mathbf{a}^*)\sigma_{ou}(\underline{a}_P, a_2).$$

The term $\tau_2^P(\mathbf{a}^*)\sigma_{ou}(\underline{a}_P, a_2)$ strictly decreases over $a_2 \in [\underline{a}_P, \bar{a}_P]$ because its first derivative with respect to a_2 is $-2e^{-(a_1 - \underline{a}_P)/\ell}\operatorname{csch}^2((a_1 - a_2)/\ell)\sinh((a_2 - \underline{a}_P)/(2\ell))\sinh((a_2 + \underline{a}_P - 2a_1)/(2\ell)) < 0$ for $a_2 > \underline{a}_P$. Hence, $\alpha_2(\{a_1, \underline{a}_P\}) > \alpha_2(\{a_1, a_2\})$ for any $a_2 > \underline{a}_P$. On the other hand, $\psi_P^2(\{a_1, a_2\})$ is single-peaked in $a_2 \in [\underline{a}_P, \bar{a}_P]$ with the peak at $\hat{a}_2 > (\underline{a}_P + \bar{a}_P)/2$, because in the absence of a_1, ψ_P^2 would be maximized at $(\underline{a}_P + \bar{a}_P)/2$. Moreover, for any $a_2 > \hat{a}_2$, $\psi_P^2(\{a_1, a_2\}) > \psi_P^2(\{a_1, \hat{a}_2 - (a_2 - \hat{a}_2)\})$. Hence, for any $a_2 \in (\underline{a}_P, \bar{a}_P]$, $\psi_P^2(\{a_1, a_2\}) > \psi_P^2(\{a_1, \underline{a}_P\})$. Therefore, $a_2 = \underline{a}_P$ guarantees higher covariance and lower ψ_P^2, which implies $V_A(\{a_1, a_2\}) < V_A(\{a_1, \underline{a}_P\}) = V_A(\{\underline{a}_P\})$ for any $a_2 > \underline{a}_P$, where the last equality follows from $\tau_1^P(\{a_1, \underline{a}_P\}) = 0$. Now consider the alternative sample $\{\bar{a}_A\}$. Note that $\psi_P^2(\{\bar{a}_A\}) = \sigma_{ou}^2(\underline{a}_P, \bar{a}_A)\psi_P^2(\{\underline{a}_P\}) < \psi_P^2(\{\underline{a}_P\})$ and $\tau^P(\bar{a}_A)\tau^A(\bar{a}_A) = \left(\sigma_{ou}(\bar{a}_A, \underline{a}_P)\tau^P(\underline{a}_P)\right)\left(\tau^A(\underline{a}_P)/\sigma_{ou}(\bar{a}_A, \underline{a}_P)\right) = \tau^P(\underline{a}_P)\tau^A(\underline{a}_P)$. Hence, $V_A(\{\underline{a}_P\}) < V_A(\{\bar{a}_A\})$. Therefore, $\{\bar{a}_A\}$ dominates $\mathbf{a}^*$. Moreover, any sample of the form $\{a_2\}$ for $a_2 \in [\underline{a}_P, \bar{a}_P]$ is dominated by $\{\bar{a}_A\}$. Hence the optimal sample is of the form $\{a_1\}$, where $a_1 \in [\underline{a}_A, \bar{a}_A]$. Differentiating $V_A(\{a_1\})$

7

with respect to a_1,

$$\frac{\partial V_A(a_1)}{\partial a_1} = 2\ell e^{\frac{a_1 - \bar{a}_A - 2(\underline{a}_P + \bar{a}_P)}{\ell}} \left(e^{\frac{\bar{a}_P}{\ell}} - e^{\frac{\underline{a}_P}{\ell}}\right) \left(e^{\frac{a_1}{\ell}} C_1 + C_0\right)$$

where $C_1 = e^{\bar{a}_A/\ell} \left(e^{\underline{a}_P/\ell} - e^{\bar{a}_P/\ell}\right) - 2e^{(\underline{a}_P + \bar{a}_P)/\ell} < 0$ and $C_0 = 2e^{(\bar{a}_A + \underline{a}_P + \bar{a}_P)/\ell} > 0$. Therefore, the FOC that uniquely pins down a_1^*, whenever the solution is interior in $[\underline{a}_A, \bar{a}_A]$, is $e^{a_1^*/\ell} = -C_0/C_1$. The second order condition is satisfied as well because

$$\left.\frac{\partial^2 V_A(a_1)}{\partial a_1^2}\right|_{a_1 = a_1^*} = 4e^{\frac{a_1^* - \bar{a}_A - 2(\underline{a}_P + \bar{a}_P)}{\ell}} \left(e^{\frac{\bar{a}_P}{\ell}} - e^{\frac{\underline{a}_P}{\ell}}\right) \left(e^{\frac{a_1^*}{\ell}} C_1 + C_0/2\right) < 0.$$

It can be easily verified that $V_A(\{\bar{a}_A\}) < 0$. Moreover if $e^{\underline{a}_A/\ell} C_1 + C_0 > 0$ then $V_A(\{\underline{a}_A\}) > 0$. Therefore, either $V_A(\{\underline{a}_A\}) > 0 > V_A(\{\bar{a}_A\})$ and V_A is single-peaked in a_1, or $V_A(\{\underline{a}_A\}) < 0$ and V_A is strictly decreasing in a_1. The optimal attribute, if interior, is given by $a_1^* = \bar{a}_P - \ell \ln\left(\frac{1}{2}\left(2e^{(\bar{a}_P - \bar{a}_A)/\ell} + e^{(\bar{a}_P - \underline{a}_P)/\ell} - 1\right)\right)$, which simplifies to $a_1^* = -\ell \ln\left(e^{-\bar{a}_A/\ell} + \frac{e^{-\underline{a}_P/\ell} - e^{-\bar{a}_P/\ell}}{2}\right)$. The case of $\bar{a}_P < \underline{a}_A$ follows by a similar argument.

(ii) Let $\bar{a}_P < \underline{a}_A$. Equation (15) simplifies to $e^{a_1^*/\ell} - e^{\underline{a}_A/\ell} = \frac{1}{2}\left(e^{\bar{a}_P/\ell} - e^{\underline{a}_P/\ell}\right)$. By implicit differentiation with respect to ℓ, we obtain

$$\frac{\partial a_1^*(\ell)}{\partial \ell} = e^{-a_1^*/\ell} \left(\frac{a_1^*}{\ell} e^{a_1^*/\ell} - \frac{\underline{a}_A}{\ell} e^{\underline{a}_A/\ell} - \frac{\bar{a}_P}{2\ell} e^{\bar{a}_P/\ell} + \frac{\underline{a}_P}{2\ell} e^{\underline{a}_P/\ell}\right).$$

The function $g(x) := xe^x$ is above $f(x) := e^x$ for $x > 1$, below $f(x)$ for $x < 1$, and strictly more convex than $f(x)$. Hence, if $e^{a_1^*/\ell} - e^{\underline{a}_A/\ell} - \frac{1}{2}\left(e^{\bar{a}_P/\ell} - e^{\underline{a}_P/\ell}\right) = 0$ then $\frac{a_1^*}{\ell} e^{a_1^*/\ell} - \frac{\underline{a}_A}{\ell} e^{\underline{a}_A/\ell} - \frac{\bar{a}_P}{2\ell} e^{\bar{a}_P/\ell} + \frac{\underline{a}_P}{2\ell} e^{\underline{a}_P/\ell} > 0$. Therefore, $a_1^*(\ell)$ is strictly increasing in ℓ. An analogous argument applies to the case of $\bar{a}_A < \underline{a}_P$.

(iii) We take the limit of $a_1^*(\ell)$ in (15) as $\ell \to 0^+$. Let $\bar{a}_P < \underline{a}_A$. Applying L'Hôpital's rule and then dividing through by $e^{\underline{a}_A/\ell}$, we obtain

$$\lim_{\ell \to 0^+} a_1^*(\ell) = \lim_{\ell \to 0^+} \frac{2\underline{a}_A e^{\underline{a}_A/\ell} + \bar{a}_P e^{\bar{a}_P/\ell} - \underline{a}_P e^{\underline{a}_P/\ell}}{2e^{\underline{a}_A/\ell} + e^{\bar{a}_P/\ell} - e^{\underline{a}_P/\ell}} = \lim_{\ell \to 0^+} \frac{2\underline{a}_A + \bar{a}_P e^{-(\underline{a}_A - \bar{a}_P)/\ell} - \underline{a}_P e^{-(\underline{a}_A - \underline{a}_P)/\ell}}{2 + e^{-(\underline{a}_A - \bar{a}_P)/\ell} - e^{-(\underline{a}_A - \underline{a}_P)/\ell}}$$

which is just $2\underline{a}_A/2 = \underline{a}_A$. By a similar argument, if $\bar{a}_A < \underline{a}_P$ then $a_1^*(\ell) \to \bar{a}_A$ as $\ell \to 0^+$.

(iv) Let $\bar{a}_P < \underline{a}_A$. By similar steps to part (iii),

$$\lim_{\ell \to +\infty} a_1^*(\ell) = \lim_{\ell \to +\infty} \frac{2\underline{a}_A + \bar{a}_P e^{-(\underline{a}_A - \bar{a}_P)/\ell} - \underline{a}_P e^{-(\underline{a}_A - \underline{a}_P)/\ell}}{2 + e^{-(\underline{a}_A - \bar{a}_P)/\ell} - e^{-(\underline{a}_A - \underline{a}_P)/\ell}} = \frac{2\underline{a}_A + \bar{a}_P - \underline{a}_P}{2} = \underline{a}_A + \frac{\Delta_P}{2}.$$

If $\Delta_A \geqslant \Delta_P/2$, this limit is interior in $[\underline{a}_A, \bar{a}_A]$. Otherwise, $\mathbf{a}^*$ is empty. A similar argument applies to the case of $\bar{a}_A < \underline{a}_P$. $\qquad \square$

Proof of Proposition 12. Without loss, fix $0 < \underline{a}_A < \underline{a}_P < \bar{a}_A < \bar{a}_P < 1$ and let an optimal sample be $\mathbf{a}^* = \{a_1, \ldots, a_n\}$. We first show that there is no sampling in $(\bar{a}_A, \bar{a}_P]$. Suppose first that

$n = 1$ and $a_1 \in (\bar{a}_A, \bar{a}_P]$. Then, $\alpha_2(a_1) = \sigma_{ou}(\bar{a}_A, a_1)\tau^A(\bar{a}_A)\tau^P(a_1)$ and its first derivative with respect to a_1 is $-2\tau^A(\bar{a}_A)\exp((\bar{a}_A - 2a_1)/\ell)(\exp(a_1/\ell) - \exp(\underline{a}_P/\ell)) < 0$. Hence, α_2 is strictly decreasing over $(\bar{a}_A, \bar{a}_P]$. Let $a_P^s := (\underline{a}_P + \bar{a}_P)/2$. If $a_1 \in (\bar{a}_A, a_P^s]$, α_1 is strictly increasing, hence $V_A(a_1)$ is strictly decreasing. If $a_1 \in (a_P^s, \bar{a}_P]$, then due to ψ_P^2 being single-peaked at a_P^s and symmetric around it, $\tau^P(a_1) = \tau^P(a_P^s - (a_1 - a_P^s))$ and $\alpha_1(a_1) = \alpha_1(a_P^s - (a_1 - a_P^s))$. Hence, $V_A(a_1) < V_A(a_P^s - (a_1 - a_P^s))$. Next suppose $n \geq 2$ and $a_n > \bar{a}_A$. Consider first the difference $\alpha_2(\mathbf{a}^*) - \alpha_2(\mathbf{a}^* \setminus \{a_n\}) = \tau_n^A(\mathbf{a}^*)\tau_n^P(\mathbf{a}^*)(1 - \sigma_{ou}^2(a_{n-1}, a_n))$, which equals

$$\sigma_{ou}(\bar{a}_A, a_n)\tau_n^P(\mathbf{a}^*)\left(\int_{a_{n-1}}^{\bar{a}_A} \sigma_{ou}(a, \bar{a}_A)(1 - \sigma_{ou}^2(a, a_{n-1}))\,\mathrm{d}a\right).$$

The term $\sigma_{ou}(\bar{a}_A, a_n)\tau_n^P(\mathbf{a}^*)$ is strictly decreasing in a_n because its first derivative with respect to a_n is $-2\exp((\bar{a}_A - a_{n-1})/\ell)\mathrm{csch}^2((a_{n-1} - a_n)/\ell)\sinh((a_n - \underline{a}_P)/\ell)\sinh((a_n + \underline{a}_P - 2a_{n-1})/\ell) < 0$ if $a_{n-1} < \underline{a}_P$ and $-2\exp((a_n + \bar{a}_A)/\ell)/(\exp(a_{n-1}/\ell) + \exp(a_n/\ell))^2 < 0$ if $a_{n-1} \geq \underline{a}_P$. Therefore, $\alpha_2(\mathbf{a}^*)$ is strictly decreasing in $a_n \in (\bar{a}_A, \bar{a}_P]$. On the other hand, from the single-player benchmark we know that $\psi_P^2(\mathbf{a}^*)$ is single-peaked in $a_n \in (a_{n-1}, \bar{a}_P]$, with a peak at $a_n^s > (\underline{a}_P + \bar{a}_P)/2$ because in the absence of the rest of the sample, and in particular a_{n-1}, it would be maximized at $(\underline{a}_P + \bar{a}_P)/2$. If $a_n^s > \bar{a}_A$, then any attribute in $(\bar{a}_A, a_n^s)$ is dominated by $\bar{a}_A$. Moreover V_A is either single-troughed in a_n, with a trough to the right of a_n^s, or strictly decreasing in $a_n \in (\bar{a}_A, \bar{a}_P]$. Hence, $V_A((\mathbf{a}^* \setminus \{a_n\}) \cup \bar{a}_A) > V_A((\mathbf{a}^* \setminus \{a_n\}) \cup \bar{a}_A)$.

Second, to show that there is no sampling in $[\underline{a}_A, \underline{a}_P)$ for $n \geq 2$, we suppose by contradiction that $a_1 < \underline{a}_P$. For $\tau_1^P(\mathbf{a}^*) \neq 0$, it must be that $a_2 > \underline{a}_P$. Differentiate V_A with respect to a_1, we obtain

$$4\ell\left(2\cosh\left(\frac{a_1 - a_2}{\ell}\right) - 1 - \cosh\left(\frac{a_2 - \underline{a}_P}{\ell}\right)\right)\mathrm{csch}^2\left(\frac{a_1 - a_2}{\ell}\right)\sinh^2\left(\frac{a_2 - \underline{a}_P}{2\ell}\right) > 0$$

for any $a_1 \leq \underline{a}_P < a_2$ because $a_2 - a_1 > a_2 - \underline{a}_P$. Hence, V_A strictly increases in a_1. Finally, $n = k$ by Corollary 9. $\qquad\square$

E Extensions and additional results

E.1 Examples for section 5.1

Example E.1 (Inference reversal due to conflicting attributes). *Let $\mathcal{A} = [\underline{a}, \bar{a}]$ and $\omega(a) = 1$ for all $a \in \mathcal{A}$. The attribute covariance is $\sigma_{lin}(a, a') = (a - \hat{a})(a' - \hat{a})$ and the prior mean is $\mu(a) = 0$ for $a, a' \in \mathcal{A}$; note that $\sigma_{lin}(\hat{a}, \hat{a}) = 0$. This structure corresponds to a linear attribute mapping f that goes through the realization $f(\hat{a}) = 0$ for $\hat{a} \in \mathcal{A}$ and the slope of which is not known (Figure 1). Attribute variance increases quadratically with distance from $\hat{a}$. Without loss, let $\hat{a} < (\underline{a} + \bar{a})/2$. The correlation between any two attribute realizations is perfect because*

$$corr(f(a), f(a')) = \frac{\sigma_{lin}(a, a')}{\sqrt{\sigma_{lin}(a, a)\sigma_{lin}(a', a')}} = \begin{cases} +1 & \text{if } sgn(a - \hat{a}) = sgn(a' - \hat{a}) \\ -1 & \text{if } sgn(a - \hat{a}) \neq sgn(a' - \hat{a}). \end{cases}$$

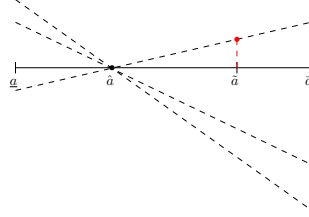*Therefore, discovering one more attribute resolves all uncertainty about f.*

Figure 1: Linear attribute mapping corresponding to σ_{lin}

Suppose $\tilde{a} \neq \hat{a}$ is discovered. The uncertainty about v prior to the discovery of $f(\tilde{a})$ is $\frac{1}{4}(\bar{a} - \underline{a})^2(\bar{a} + \underline{a} - 2\hat{a})^2$. The project is more uncertain the greater is the mass of attributes $(\bar{a} - \underline{a})$ and the farther $\hat{a}$ is from the median attribute $(\underline{a} + \bar{a})/2$, i.e., the more peripheral the known attribute $\hat{a}$ is. If $\hat{a}$ is exactly the median attribute, the uncertainty about v is zero because the uncertainty about $[\underline{a}, \hat{a}]$ cancels that about $[\hat{a}, \bar{a}]$. Given a singleton sample $\mathbf{a} = \{\tilde{a}\}$, by equation (6) the expected realization of any other attribute a is $\mathbb{E}[f(a) \mid f(\tilde{a})] = \tau_1(a; \mathbf{a})f(\tilde{a}) = (a - \hat{a})f(\tilde{a})/(\tilde{a} - \hat{a})$. Hence, from equation (8), the sample weight is $\tau_1(\tilde{a}) = \frac{1}{2}(\bar{a} - \underline{a})(\bar{a} + \underline{a} - 2\hat{a})/(\tilde{a} - \hat{a})$, which is strictly negative for $\tilde{a} < \hat{a}$. That is, a high realization for $\tilde{a} < \hat{a}$ implies low realizations for attributes in $[\hat{a}, \bar{a}]$, which is the majority of the attributes. Therefore, the sample weight of attributes to the left of $\hat{a}$ is negative even though all attributes are desirable.

Example E.2 (Inference reversal due to the presence of other sample attributes). *Let $\mathcal{A} = [0, 1]$, $\omega(a) = 1$, and the squared-exponential covariance $\sigma_2(a, a') = e^{-(a-a')^2/\ell^2}$ for all $a, a' \in [0, 1]$. Lemma E.3 shows the possibility of a reversal in the direction of inference when going from a one-attribute sample to a two-attribute one. Due to the positive attribute correlation, any singleton sample has a strictly positive sample weight. But in a two-attribute sample, one of the attributes can have a strictly negative sample weight, even though the sum of the sample weights for the two attributes must be strictly positive. Lemma E.3(ii) establishes that such a negative sample weight arises if and only if the two attributes are on the same side of the median attribute and attribute correlation is high. The attribute with a negative sample weight is the one farther away from the median attribute.*

Lemma E.3. *Let $\sigma_2(a, a') = e^{-(a-a')^2/\ell^2}$, and $\omega(a) = 1$ for all $a, a' \in [0, 1]$. For any sample $\mathbf{a}_1 = \{a_1\}$, $\tau_1(\mathbf{a}_1) > 0$. For any two-attribute sample $\mathbf{a}_2 = \{a_1, a_2\}$ such that $0 \leqslant a_1 < a_2 \leqslant 1$,*

(i) the sum of sample weights is always positive: $\tau_1(\mathbf{a}_2) + \tau_2(\mathbf{a}_2) > 0$;

(ii) one of the attributes is assigned a strictly negative if and only if a_1 and a_2 are on the same side of the median attribute and ℓ is sufficiently large.

Proof of Lemma E.3. (i) Let $g(a) := \operatorname{erf}\left(\frac{a}{\ell}\right) + \operatorname{erf}\left(\frac{1-a}{\ell}\right)$. First, note that $g(a) > 0$ because $a_1 \in [0,1]$, $\ell > 0$ and $\operatorname{erf}(x) > 0$ for any $x > 0$. For a singleton sample, equation (8) simplifies to $\tau_1(\mathbf{a}_1) = \ell\sqrt{\pi}g(a_1) > 0$. Now consider $\mathbf{a}_2 = \{a_1, a_2\}$, where $a_1 < a_2$, and let $d := a_2 - a_1$. Applying Lemma 1, the sample weights are given by

$$\tau_j(\mathbf{a}) = \frac{1}{4}\ell\sqrt{\pi}e^{-\frac{4a_1 a_2}{\ell^2}}\operatorname{csch}\left(\frac{d^2}{\ell^2}\right)\left(e^{d^2/\ell^2}g(a_j) - g(a_{-j})\right)$$

which is positive if and only if $e^{d^2/\ell^2}g(a_j) - g(a_{-j}) > 0$. Then, the sign of the sum $\tau_1(\mathbf{a}) + \tau_2(\mathbf{a})$ is determined by the sign of $g(a_1) + g(a_2)$, which is strictly positive for any $a_1, a_2 \in [0,1]$. Hence at least one of the attributes has a strictly positive sample weight.

(ii) Taking the limit of these sample weights as $\ell \to +\infty$, we obtain

$$\lim_{\ell\to+\infty}\tau_1(\mathbf{a}_2) = \frac{2a_2 - 1}{2(a_2 - a_1)}, \qquad \lim_{\ell\to+\infty}\tau_2(\mathbf{a}_2) = \frac{1 - 2a_1}{2(a_2 - a_1)}.$$

If $a_1 < a_2 < 1/2$, then $\lim_{\ell\to+\infty}\tau_1(\mathbf{a}_2) < 0$. If $1/2 < a_1 < a_2$, then $\lim_{\ell\to+\infty}\tau_2(\mathbf{a}_2) < 0$. So the conditions are sufficient. To show that they are also necessary, suppose first $a_1 < 1/2 < a_2$. Then, $e^{d^2/\ell^2}g(a_j) - g(a_{-j})$ strictly increases in the distance d for any $a_j \in \mathbf{a}_2$ and it is zero for $d = 0$. Second, suppose that $a_1 < a_2 < 1/2$. Then, $\tau_1(\mathbf{a}_2)$ as a function of ℓ is single-troughed in ℓ and crosses zero only once in ℓ, say at $\ell = \bar{\ell}$. On the other hand, $\tau_2(\mathbf{a}_2)$ as a function of ℓ is decreasing and strictly positive in ℓ. Hence, for $\tau_1(\mathbf{a}_2)$ to be strictly negative, it is necessary that $\ell > \bar{\ell}$. $\qquad\square$

E.2 Binary decision and reservation values

Proposition E.2. *Let $D = \{0,1\}$ and for each $i = A, P$, $u(1, v_i) = v_i$ and $u(0, v_i) = r_i$, where $r_i \in \mathbb{R}$ is a known outside option. The agent's expected payoff from any sample $\mathbf{a} \in \mathcal{A}_k$ is*

$$V_A(\mathbf{a}) = r_A + \left(\nu_0^A - r_A\right)\Phi\left(\frac{\nu_0^P - r_P}{\sqrt{\alpha_1(\mathbf{a})}}\right) + \frac{\alpha_2(\mathbf{a})}{\sqrt{\alpha_1(\mathbf{a})}}\phi\left(\frac{\nu_0^P - r_P}{\sqrt{\alpha_1(\mathbf{a})}}\right),$$

where α_1 and α_2 are as defined in theorem 2.

Proof of proposition E.2. Let $\rho(\mathbf{a})$ denote the correlation between $\nu_P(\mathbf{a})$ and $\nu_A(\mathbf{a})$, the joint distribution is Gaussian:

$$\begin{pmatrix}\nu^P(\mathbf{a})\\ \nu^A(\mathbf{a})\end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix}\nu_0^P\\ \nu_0^A\end{pmatrix}, \begin{pmatrix}\psi_P^2(\mathbf{a}) & \rho(\mathbf{a})\psi_A(\mathbf{a})\psi_P(\mathbf{a})\\ \rho(\mathbf{a})\psi_A(\mathbf{a})\psi_P(\mathbf{a}) & \psi_P^2(\mathbf{a})\end{pmatrix}\right).$$

Claim 1. *For any $r_P \in \mathbb{R}$,*

$$f(\nu^A(\mathbf{a}) \mid \nu^P(\mathbf{a}) \geq r_P) = \frac{\phi\left(\frac{\nu^A(\mathbf{a}) - \nu_0^A}{\psi_A(\mathbf{a})}\right)}{\psi_A(\mathbf{a})\Phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right)}\Phi\left(\frac{\nu_0^P + \rho(\mathbf{a})\frac{\psi_P(\mathbf{a})}{\psi_A(\mathbf{a})}(\nu^A(\mathbf{a}) - \nu_0^A) - r_P}{\psi_P(\mathbf{a})\sqrt{1 - \rho(\mathbf{a})^2}}\right).$$

Proof. Let x_1, x_2 be jointly Gaussian with means μ_1, μ_2, variances σ_1^2, σ_2^2 and covariance σ_{12}. Let f_1, f_2 and F_1, F_2 denote their respective pdf and cdf. Then,

$$
\begin{aligned}
f(x_1 \mid x_2 \geq \bar{x}) &= \frac{1}{1 - F_2(\bar{x})} \Pr(x_2 \geq \bar{x}) f(x_1 \mid x_2 \geq \bar{x}) \\
&= \frac{1}{1 - F_2(\bar{x})} \int_{\bar{x}}^{\infty} f(x_2 \mid x_1) f_1(x_1) dx_2 \\
&= \frac{f_1(x_1)}{1 - F_2(\bar{x})} (1 - F_{x_2 \mid x_1}(\bar{x})).
\end{aligned}
$$

The first line multiplies and divides by $\Pr(x_2 \geq \bar{x})$. The second line rewrites $\Pr(x_2 \geq \bar{x}) f(x_1 \mid x_2 \geq \bar{x})$ using the joint density and the observation that $f(x_1, x_2) = f(x_2 \mid x_1) f_1(x_1)$. The last two lines use the conditional distribution of $x_2 \mid x_1$. But,

$$
x_2 \mid x_1 \sim \mathcal{N}\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1), (1 - \rho^2)\sigma_2^2\right)
$$

and $\rho = \frac{\sigma_{12}}{\sigma_1 \sigma_2}$. Therefore, we can substitute in the expression for $F_{x_2 \mid x_1}$ to obtain

$$
f(x_1 \mid x_2 \geq \bar{x}) = \frac{f_1(x_1)}{1 - F_2(\bar{x})}\left(1 - \Phi\left(\frac{\bar{x} - \mu_2 - \rho \frac{\sigma_2}{\sigma_1}(x_1 - \mu_1)}{\sigma_2 \sqrt{1 - \rho^2}}\right)\right).
$$

Switching back to our variables of interest, let $x_1 := \nu^A(\mathbf{a}) \sim \mathcal{N}(\nu_0^A, \psi_A^2(\mathbf{a}))$, $x_2 := \nu^P(\mathbf{a}) \sim \mathcal{N}(\nu_0^P, \psi_P^2(\mathbf{a}))$ and $\bar{x} := r_P$. Therefore,

$$
f(\nu^A(\mathbf{a}) \mid \nu^P(\mathbf{a}) \geq r_P) = \frac{\phi\left(\frac{\nu^A(\mathbf{a}) - \nu_0^A}{\psi_A(\mathbf{a})}\right)}{\psi_A(\mathbf{a})\left(1 - \Phi\left(\frac{r_P - \nu_0^P}{\psi_P(\mathbf{a})}\right)\right)}\left(1 - \Phi\left(\frac{r_P - \nu_0^P - \rho(\mathbf{a})\frac{\psi_P(\mathbf{a})}{\psi_A(\mathbf{a})}(\nu^A(\mathbf{a}) - \nu_0^A)}{\psi_P(\mathbf{a})\sqrt{1 - \rho(\mathbf{a})^2}}\right)\right).
$$

$\square$

Using the claim, observe that:

$$
\begin{aligned}
\Pr(\nu^P(\mathbf{a}) \geq r_P)\mathbb{E}[\nu^A(\mathbf{a}) \mid \nu^P(\mathbf{a}) \geq r_P] &= \Phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right)\int_{-\infty}^{\infty} \nu^A(\mathbf{a}) f(\nu^A(\mathbf{a}) \mid \nu^P(\mathbf{a}) \geq r_P)\, d\nu^A(\mathbf{a}) \\
&= \int_{-\infty}^{\infty} \frac{\nu^A(\mathbf{a})}{\psi_A(\mathbf{a})}\phi\left(\frac{\nu^A(\mathbf{a}) - \nu_0^A}{\psi_A(\mathbf{a})}\right)\Phi\left(\frac{\nu_0^P + \rho(\mathbf{a})\frac{\psi_P(\mathbf{a})}{\psi_A(\mathbf{a})}(\nu^A(\mathbf{a}) - \nu_0^A) - r_P}{\psi_P(\mathbf{a})\sqrt{1 - \rho(\mathbf{a})^2}}\right) d\nu^A(\mathbf{a}) \\
&= \int_{-\infty}^{\infty} \left(x\psi_A(\mathbf{a}) + \nu_0^A\right)\phi(x)\Phi\left(\frac{\nu_0^P + \rho(\mathbf{a})\psi_P(\mathbf{a})x - r_P}{\psi_P(\mathbf{a})\sqrt{1 - \rho^2(\mathbf{a})}}\right) dx,
\end{aligned}
$$

where in the last line $x := \frac{\nu^A(\mathbf{a}) - \nu_0^A}{\psi_A(\mathbf{a})}$. From Owen (1980), we have the following Gaussian identities

(respectively, numbered 10,010.8 and 10,011.1 in Owen (1980)):

$$\int_{-\infty}^{\infty} \phi(x)\Phi(a+bx)dx = \Phi\left(\frac{a}{\sqrt{1+b^2}}\right), \quad \int_{-\infty}^{\infty} x\phi(x)\Phi(a+bx)dx = \frac{b}{\sqrt{1+b^2}}\phi\left(\frac{a}{\sqrt{1+b^2}}\right).$$

Letting $a := (\nu_0^P - r_P)/(\psi_P(\mathbf{a})\sqrt{1-\rho^2(\mathbf{a})})$ and $b := \rho(\mathbf{a})/\sqrt{1-\rho^2(\mathbf{a})}$,

$$\Pr(\nu^P(\mathbf{a}) \geq r_P)\mathbb{E}[\nu^A(\mathbf{a}) \mid \nu^P(\mathbf{a}) \geq r_P] = \nu_0^A\Phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right) + \rho(\mathbf{a})\psi_A(\mathbf{a})\phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right).$$

Therefore, the agent's payoff from sample $\mathbf{a}$ simplifies to

$$\begin{aligned}
V_A(\mathbf{a}) &= \Pr(\nu^P(\mathbf{a}) < r_P)r_A + \nu_0^A\Phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right) + \rho(\mathbf{a})\psi_A(\mathbf{a})\phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right) \\
&= r_A + (\nu_0^A - r_A)\Phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right) + \rho(\mathbf{a})\psi_A(\mathbf{a})\phi\left(\frac{\nu_0^P - r_P}{\psi_P(\mathbf{a})}\right).
\end{aligned}$$

Finally note that $\text{cov}[\nu^P(\mathbf{a}), \nu^A(\mathbf{a})] = \text{cov}[\nu^P(\mathbf{a}), v_A] = \alpha_2(\mathbf{a})$ because $\tau_j^A(\mathbf{a}) + \sum_{i \neq j} \tau_i^A(\mathbf{a})\sigma(a_i, a_j) = \tau^A(a_j)$. Substituting $\Psi_P(\mathbf{a}) = \sqrt{\alpha_1(\mathbf{a})}$ and $\rho(\mathbf{a})\psi_A(\mathbf{a}) = \alpha_2(\mathbf{a})/\sqrt{\alpha_1(\mathbf{a})}$ into $V_A(\mathbf{a})$, we obtain the desired expression. □

E.3 Noisy observations of attribute realizations

Fix a sample $\mathbf{a} \in \mathcal{A}_k$ and noisy observations $y(\mathbf{a}) = f(\mathbf{a}) + \epsilon(\mathbf{a})$, where $\epsilon(a) \sim \mathcal{N}(\mu^0(a), \eta^2(a))$ is the noise term drawn independently across attributes. Figure 2 illustrates extrapolation across noisy realizations of a Brownian sample path.
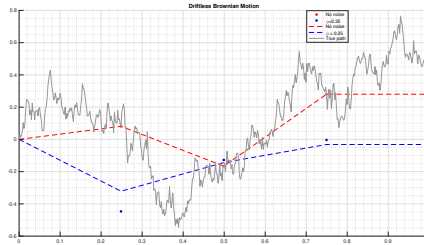

Figure 2: Extrapolation across a standard Brownian motion with $\eta = 0$ (red) and $\eta = 0.25$ (blue). Sample $\mathbf{a} = \{1/4, 1/2, 3/4\}$ and $\mathcal{A} = [0,1]$. Also, $\mu(a) = \mu^0(a) = 0$ for all $a \in \mathcal{A}$.

Corollary E.4. *The set of single-player samples does not depend on observational bias μ^0.*

Proof. Fix a sample $\mathbf{a} = \{a_1, \ldots, a_k\} \in \mathcal{A}_k$. The observations are distributed according to

$$
\begin{pmatrix} y(a_1) \\ \vdots \\ y(a_k) \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu(a_1) + \mu^0(a_1) \\ \vdots \\ \mu(a_k) + \mu^0(a_k) \end{pmatrix}, \begin{pmatrix} \sigma(a_1, a_1) + \eta^2(a_1) & \ldots & \sigma(a_1, a_k) \\ \sigma(a_2, a_1) & \ldots & \sigma(a_2, a_k) \\ \vdots & \ddots & \vdots \\ \sigma(a_k, a_1) & \ldots & \sigma(a_k, a_k) + \eta^2(a_k) \end{pmatrix} \right).
$$

Let $\Sigma(\eta)$ be this new covariance matrix. Following Lemma 1, $\tau_j(\hat{a}; \mathbf{a})$ is now the $(1, j)^{th}$ entry of the matrix $\left(\sigma(a_1, \hat{a}) \quad \ldots \quad \sigma(a_k, \hat{a}) \right) \Sigma^{-1}(\eta)$. The posterior variance is as in equation (10), where $\tau_j(\mathbf{a})$ is derived from $\tau_j(\hat{a}; \mathbf{a})$ above as in Lemma 2.1. By the same argument as in Theorem 1(iii), μ^0 enters neither the posterior variance nor the single-player sample. □

Example E.5 (Noisier observations, more uncertain attributes). *Consider the Brownian covariance $\sigma_{br}(a, a') = \min(a, a')$ over $\mathcal{A} = [0, 1]$. That is, attribute uncertainty increases from left to right and attribute $a = 0$ is the least uncertain attribute. Let $\omega(a) = 1$ for all $a \in [0, 1]$ and $k = 1$. The observations are of the form $y(a) = f(a) + \epsilon$, where $\epsilon \sim \mathcal{N}(0, \eta^2)$. For any sample $a \in [0, 1]$, the posterior variance $\psi^2(a)$ naturally decreases with the amount of noise η^2. The optimal sample $a^*(\eta)$ is pinned down by $a^*(\eta)(3a^*(\eta) - 2) - 4(1 - a^*(\eta))\eta^2 = 0$. It can be easily verified that the optimal attribute without observational noise is $a^*(0) = 2/3$. By implicit differentiation with respect to η,*
$$
\frac{\partial a^*(\eta)}{\partial \eta} = \frac{4\eta(1 - a^*(\eta))}{3a^*(\eta) + 2\eta^2 - 1} > 0 \text{ for } a^* \in (2/3, 1) \text{ and } \eta > 0 \text{ and } a^*(\eta) \text{ is strictly increasing at } \eta = 0.
$$
The higher η^2 is, the further away the single-player attribute is from $a = 0$. That is, in the presence of greater observational noise, the player samples attributes that are ex ante more uncertain.

References

Clemen, Robert T., and Robert L. Winkler. 1985. "Limits for the Precision and Value of Information from Dependent Sources." *Operations Research*, 33(2): 427–442. 5

Owen, Donald Bruce. 1980. "A Table of Normal Integrals." *Communications in Statistics - Simulation and Computation*, 9(4): 389–419. 12, 13