Public Goods, Social Pressure, and the Choice Between Privacy and Publicity

By Andrew F. Daughety and Jennifer F. Reinganum

TECHNICAL APPENDIX

This technical appendix includes proofs of comparative statics results; the proof of the claim made in the text that if $\beta' > \beta$, then $H'(\delta; \beta')$ first-order stochastically dominates $H'(\delta; \beta)$; computational results, and detailed analysis of the interim preferences over policies.

A. Comparative Statics

The functions $g^P(\theta_i)$ and $g^O(\theta_i)$ depend on $\theta_i$, $\beta$, $\gamma$, and $p$; they are independent of $\alpha$.

Comparative statics of $g^P(\theta_i)$. Since $g^P(\theta_i) = g_{\min} + \theta_i$ and $g_{\min} = \gamma - p$, it is obvious that $g^P(\theta_i)$ is an increasing function of $\theta_i$, and that the function $g^P(\theta_i)$ shifts upward with an increase in $\gamma$, and shifts downward with an increase in $p$. Finally, the function $g^P(\theta_i)$ is always independent of $\beta$. Since the utility function is quasilinear, $g^P(\theta_i)$ is independent of income, $I$.

Comparative statics of $g^O(\theta_i)$. Since $g^O(0) = g^P(0)$, $g^O(0)$ behaves as described above with respect to the parameters. Thus, in what follows, we will consider only $\theta_i > 0$. Let $RHS = g_{\min} + \theta_i + \beta(1 - \exp[-(g^O(\theta_i) - g_{\min})/\beta])$. For any parameter $m$, the implicit function in Proposition 1(i) can be differentiated to obtain $\partial g^O/\partial m = (\partial RHS/\partial m) + (\partial RHS/\partial g^O)(\partial g^O/\partial m)$. Collecting terms implies that $\partial g^O/\partial m = (\partial RHS/\partial m)(1 - \exp[-(g^O(\theta_i) - g_{\min})/\beta])$. Since the denominator is positive, the sign of $\partial g^O/\partial m$ is the same as the sign of $(\partial RHS/\partial m)$. To save on notation, it will be useful to define the function $z^O(\theta_i) = (g^O(\theta_i) - g_{\min})/\beta$, and to use $z$ to denote an arbitrary (positive) value.

Since $\partial RHS/\partial \theta_i = 1$, it follows that $g^O(\theta_i) = 1/(1 - \exp[-z^O(\theta_i)]) > 0$; that is, the equilibrium action under a policy of publicity (openness) is increasing in type.

Since the parameters $\gamma$ and $p$ appear only in $g_{\min}$, and $(\partial RHS/\partial g_{\min}) = (1 - \exp[-z^O(\theta_i)])$, it is straightforward to show that $\partial g^O(\theta_i)/\partial g_{\min} = 1$. Therefore $\partial g^O(\theta_i)/\partial \gamma = 1$ and $\partial g^O(\theta_i)/\partial p = -1$. 
Differentiating and collecting terms yields $\frac{\partial g^O(\theta)}{\partial \beta} = (1 - \exp[-z^O(\theta)]) - z^O(\theta)\exp[-z^O(\theta)]/(1 - \exp[-z^O(\theta)])$. The function $1 - \exp[-z] - z\exp[-z]$ is easily shown to be positive for $z > 0$; thus, $\frac{\partial g^O(\theta)}{\partial \beta} > 0$.

*Comparative statics of the action differential* $g^O(\theta) - g^P(\theta)$.

Let $\delta(\theta; \beta) = g^O(\theta) - g^P(\theta) = \beta(1 - \exp[-(z^O(\theta))])$ denote the action differential as a function of $\theta$. This difference is increasing in type; that is, $\delta'(\theta; \beta) = \exp[-(z^O(\theta))]g^{O'}(\theta) > 0$. Thus, the highest type inflates his action the most. We have already seen that $\frac{\partial g^O(\theta)}{\partial g_{\min}} = 1$; this yields the immediate result that $\frac{\partial^2 g^O(\theta)}{\partial g_{\min}^2} = (\frac{\partial g^O(\theta)}{\partial \beta} - g_{\min})/\frac{\partial g_{\min}}{\partial \beta} = 0$. This implies that the action differential $\delta(\theta; \beta)$ is independent of the parameters $\gamma$ and $p$. Since $g^O(\theta)$ is independent of $\beta$, then $\frac{\partial \delta(\theta; \beta)}{\partial \beta} = \frac{\partial g^O(\theta)}{\partial \beta} > 0$.

**B. Proof of Claim that if $\beta' > \beta$, then $H^\delta(\beta' \delta; \beta')$ First-order Stochastic Dominates $H^\delta(\delta; \beta')$**

Recall that $\delta(\theta; \beta) = \beta(1 - \exp[-(g^O(\theta) - g_{\min})/\beta])$, and let $\bar{\delta}(\beta) = \delta(\bar{\theta}; \beta)$ for any given $\beta$; since $\delta(\theta; \beta)$ is increasing in $\beta$, so is $\bar{\delta}(\beta)$. Therefore the support of $H^\delta(t; \beta)$ induced by $H(\theta)$ and $\delta(\theta; \beta)$ is $[0, \bar{\delta}(\beta)]$. Then, fixing $\beta$:

$$H^\delta(t; \beta) = \Pr\{\delta(\theta; \beta) \leq t\} = \Pr\{\theta \leq (g^O)^{-1}(\beta\ln(\beta/(\beta - t) + g_{\min})\} = H((g^O)^{-1}(\beta\ln(\beta/(\beta - t) + g_{\min})).$$

Thus, $\frac{\partial H^\delta(t; \beta)}{\partial \beta} = h(t)((g^O)^{-1}(\beta\ln(\beta/(\beta - t) + g_{\min}))[\ln \beta + 1 - \ln(\beta - t) - \beta/(\beta - t)]$, so that $\frac{\partial H^\delta(t; \beta)}{\partial \beta} < 0$ if and only if $\ln \beta + 1 - \ln(\beta - t) - \beta/(\beta - t) < 0$. Note that $H^\delta(0; \beta) = 0$ and $H^\delta(\bar{\delta}(\beta); \beta) = \Pr\{\delta(\bar{\theta}; \beta) \leq \bar{\delta}(\beta)\} = 1$ for any given value of $\beta$, so we are interested in $\frac{\partial H^\delta(t; \beta)}{\partial \beta}$ for $t \in (0, \bar{\delta}(\beta))$.1

---

1 Note that increasing $\beta$ increases the right end-point, so this means we must extend $H^\delta(t; \beta)$ to be 1 on the interval $[\bar{\delta}(\beta), \bar{\delta}(\beta')]$ when we compare it to the distribution $H^\delta(t; \beta')$, so that they are on the same support.
Note that \( \ln \beta + 1 - \ln(\beta - t) - \beta/(\beta - t) < 0 \) if and only if \( \ln(\beta/(\beta - t)) < 1 - \beta/(\beta - t) \) for \( t \) in this open interval. Note that \( t < \beta \) since \( 1 - \exp[-(g^\beta(\theta) - g_{\min})/\beta]] < 1 \). Thus, we may restate the problem as: is \( \ln x < x - 1 \) for \( x \geq 1 \)? In fact, the line \( x - 1 \) is tangent to \( \ln x \) at \( x = 1 \), so \( \ln x < x - 1 \) for \( x > 1 \) and the two functions are equal at \( x = 1 \). Therefore, \( \partial \tilde{H}^\beta(t; \beta)/\partial \beta < 0 \) for \( t \in (0, \tilde{\beta}(\beta)) \), so that if \( \beta' > \beta \), then \( \tilde{H}^\beta(t; \beta') < \tilde{H}^\beta(t; \beta) \) for \( t \in (0, \tilde{\beta}(\beta)) \); that is, \( \tilde{H}^\beta(t; \beta') \) FOSD \( \tilde{H}^\beta(t; \beta) \).

C. Computational Results on the Effect of \( \beta \) on \( \alpha^{PO} \)

Table 1 below displays computational results for four density functions: 1) the Uniform density, with \( h(\theta) = 1 \); 2) the Left Triangle density, with \( h(\theta) = 2 - 2\theta \); 3) the Middle Triangle density, with \( h(\theta) = 4\theta \) when \( \theta \leq \frac{1}{2} \), and \( h(\theta) = 4 - 4\theta \) when \( \theta > \frac{1}{2} \); and 4) the Right Triangle density, with \( h(\theta) = 2\theta \). Notice that the Uniform density is a mean-preserving spread of the Middle Triangle density.

<table>
<thead>
<tr>
<th>density</th>
<th>( \beta )</th>
<th>0.5</th>
<th>1.0</th>
<th>2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td></td>
<td>0.40859</td>
<td>0.69264</td>
<td>1.14159</td>
</tr>
<tr>
<td>Left Triangle</td>
<td></td>
<td>0.36996</td>
<td>0.61131</td>
<td>0.96546</td>
</tr>
<tr>
<td>Middle Triangle</td>
<td></td>
<td>0.41363</td>
<td>0.69296</td>
<td>1.10361</td>
</tr>
<tr>
<td>Right Triangle</td>
<td></td>
<td>0.43900</td>
<td>0.75341</td>
<td>1.22101</td>
</tr>
</tbody>
</table>

Table 1 suggests that, for a given density, increasing \( \beta \) increases \( \alpha^{PO} \), so that \( \Phi^{PO}(\alpha^{PO}) \) shifts up, associating more values of \( \alpha \) with privacy than were associated with the lower value of \( \beta \). Also, note that, holding \( \beta \) constant, the computed values of \( E(\delta^2)/E(\delta) \) increase as we move from the Left to the Middle to the Right Triangle distributions. Thus, Table 1 is consistent with the conjecture that a shift in \( H \) to a new distribution \( H' \), where \( H' \) first-order stochastic dominates \( H \), results in higher
values of $\alpha^{PO}$ as well (i.e., upward shifts of $\Phi^{PO}$, too).

D. Material on Interim Preferences over Policies $P$ and $O$

This material pertains to Proposition 4. Two results follow from equation (5). First, comparing with equation (3), we see that $E(I^{PO}(\theta, \alpha)) = \Phi^{PO}(\alpha)$, so that when evaluated at $\alpha = \alpha^{PO}$, $E(I^{PO}(\theta, \alpha^{PO}), \alpha^{PO}) = 0$. Since differentiating $I^{PO}(\theta, \alpha)$ shows that it is a monotonically decreasing function of $\theta$, for each value of $\alpha$, this implies that $I^{PO}(0, \alpha^{PO}) > 0$ and $I^{PO}(\tilde{\theta}, \alpha^{PO}) < 0$, so that on an interim basis, if $\alpha = \alpha^{PO}$, then lower types will (interim) prefer $P$ to $O$ and higher types will (interim) prefer $O$ to $P$. Define two other values of $\alpha$, namely $\tilde{\alpha}^{PO} > 0$ such that $I^{PO}(\theta_G, \alpha^{PO}) = 0$ when $\mu > \theta_G - (\delta(\theta_G; \beta))^2/2\beta$ (that is, the value of $\alpha$ such that all types will interim prefer $P$ to $O$ for any $\alpha < \tilde{\alpha}^{PO}$; note that if $\mu < \tilde{\theta} - (\delta(\tilde{\theta}; \beta))^2/2\beta$ then no such non-negative value exists), and $\bar{\alpha}^{PO}$ such that $I^{PO}(0, \bar{\alpha}^{PO}) = 0$ (that is, the value of $\alpha$ such that all types will interim prefer $O$ to $P$ for any $\alpha > \bar{\alpha}^{PO}$). By construction, $\tilde{\alpha}^{PO} < \alpha^{PO} < \bar{\alpha}^{PO}$. Furthermore, when $\alpha \leq \tilde{\alpha}^{PO}$, the ex ante social preference for $P$ over $O$ is therefore reinforced by interim unanimity for $P$ over $O$, while when $\alpha > \bar{\alpha}^{PO}$, the ex ante social preference for $O$ over $P$ is reinforced by interim unanimity for $O$ over $P$. However, when $\alpha$ lies between $\tilde{\alpha}^{PO}$ and $\bar{\alpha}^{PO}$, lower types prefer $P$ to $O$ while higher types prefer $O$ to $P$, so that for all $\alpha$ in the interval $(\tilde{\alpha}^{PO}, \bar{\alpha}^{PO})$ there is disagreement about the preferred policy at the interim stage, and there will not be unanimous reinforcement of any ex ante policy choice.

E. Conflict Between Ex Ante and Interim Preferences

To see the possibility of conflict between $ex$ ante and interim preferences in a case wherein $O$ is $ex$ ante preferred but $P$ is interim preferred by the median type, let $\theta^{PO}(\alpha)$ be the marginal type
such that \( I^{PO}(\theta^{PO}(\alpha), \alpha) = 0 \), for \( \alpha \geq 0 \). Note that \( \theta^{PO}(\alpha) \) is decreasing in \( \alpha \), and that \( \theta^{PO}(0) > \mu \), the mean (and median) type if \( H \) is symmetric. Thus, there is an \( \alpha^* \) such that \( \theta^{PO}(\alpha^*) = \mu \). It is straightforward to show that \( \alpha^* \in (\alpha^{PO}, \alpha^{GPO}) \), so that for any value of \( \alpha \) in the interval \((\alpha^{PO}, \alpha^*)\), the \textit{ex ante} social payoff-maximizing choice of policy is \( O \), but on an \textit{interim} basis, the median type would prefer \( P \) to \( O \).

To see how the reverse conflict can occur, assume that \( \alpha = 0 \). Since \( \alpha^{PO} > 0 \), this means that society \textit{ex ante} prefers \( P \) to \( O \). Since \( \theta^{PO}(0) > \mu \), then any density \( h \) whose median is to the right of \( \theta^{PO}(0) \) implies that the median type prefers \( O \) to \( P \). Signaling type to gain esteem is sufficiently valuable to the median type (but is irrelevant in the case of the \textit{ex ante} decision) for those types to \textit{interim} prefer \( O \) to \( P \). This conflict between the \textit{ex ante} and \textit{interim} settings is summarized below.

**REMARK 2.** Conflicting \textit{Ex Ante} and \textit{Interim} Preferences over Policies.

- \( h \) symmetric: There are values of \( \alpha \) such that while a policy of publicity is \textit{ex ante} socially preferred, the alternative policy of privacy is \textit{interim}-preferred by the median type.

- \( h \) sufficiently right-weighted: There are values of \( \alpha \) such that while a policy of privacy is \textit{ex ante} socially preferred, a policy of publicity is \textit{interim}-preferred by the median type.

**PROOF OF PROPOSITION 6(a):**

Proposition 6(a) provides the following ordering of the \( \alpha \)-values at which there is \textit{ex ante} indifference between any two policies: \( 0 < \alpha^{WO} < \alpha^{PO} < \alpha^{PW} \). To see that \( 0 < \alpha^{WO} < \alpha^{PO} \), let:

\[
\eta(t) = \int_0^t (\delta(\theta; \beta))^2 h(\theta)d\theta + \int_0^t \delta(\theta; \beta) h(\theta)d\theta.
\]

Then \( \alpha^{WO} = \eta(\theta^w) \), which is clearly positive, while \( \alpha^{PO} = \eta(\tilde{\theta}) \). It is straightforward to show that sgn \( \{\eta'(t)\} = \text{sgn} \{\delta(t; \beta) \int_0^t \delta(\theta; \beta) h(\theta)d\theta - \int_0^t (\delta(\theta; \beta))^2 h(\theta)d\theta\} > 0 \) for all \( t > 0 \). Therefore, it follows that \( \alpha^{PO} = \eta(\tilde{\theta}) > \eta(\theta^w) = \alpha^{WO} \).
To see that $\alpha^{PO} < \alpha^{PW}$, let

$$v(s) = \int_s^{\tilde{\theta}} (\delta(\theta; \beta))^2 h(\theta) d\theta \int_s^{\tilde{\theta}} \delta(\theta; \beta) h(\theta) d\theta.$$ 

Then $\alpha^{PO} = v(0)$, while $\alpha^{PW} = v(\theta^W)$. It is straightforward to show that $\text{sgn} \{v'(s)\} = \text{sgn} \{\int_s^{\tilde{\theta}} (\delta(\theta; \beta))^2 h(\theta) d\theta - \delta(s; \beta) \int_s^{\tilde{\theta}} \delta(\theta; \beta) h(\theta) d\theta\} > 0$ for all $s < \tilde{\theta}$. Therefore, it follows that $\alpha^{PO} = v(0) < v(\theta^W) = \alpha^{PW}$.

F. Material on Interim Preferences over Policies P, O and W

Throughout this discussion we assume that $\theta^W \in (0, \tilde{\theta})$; if not, then the policy $W$ coincides with either $O$ or $P$ and there are not three distinct policies to be compared.

Recall that the conditional mean is $\mu(\theta) = \int_{\mathcal{F}} d(t) dt / H(\theta^W)$, where $\mathcal{F} = [0, \theta^W]$. Furthermore, let $E(g^O - g^P)$ denote the expected distortion under a policy of $O$ versus a policy of $P$, and similarly for $E(g^W - g^P)$ and $E(g^O - g^W)$. Then:

(a) $E(g^O - g^P) = \int_{\mathcal{F}} \delta(t; \beta) h(t) dt$, where the integral is taken over $[0, \tilde{\theta}]$;

(b) $E(g^W - g^P) = \int_{\mathcal{F}} \delta(t; \beta) h(t) dt$, where the integral is taken over $\mathcal{E} = [\theta^W, \tilde{\theta}]$;

(c) $E(g^O - g^W) = \int_{\mathcal{F}} \delta(t; \beta) h(t) dt$, where the integral is taken over $\mathcal{F} = [0, \theta^W]$.

The integral in part (a) reflects the fact that every type (except the lowest) distorts her action under a policy of $O$ while no type distorts her action under a policy of $P$. The integral in part (b) reflects the fact that only those types in $\mathcal{E} = [\theta^W, \tilde{\theta}]$ distort their actions. Finally, the integral in part (c) reflects the fact that only those types in $\mathcal{F} = [0, \theta^W]$ do not distort their actions.

These definitions allow us to summarize the type-specific value of one policy over another. Let $\Gamma^{PO}(\theta, \alpha) = V(x^P(\theta), \theta, \mu, G^P) - V(x^O(\theta), \theta, \theta, G^O)$ denote the type-specific value of a policy of privacy over a policy of publicity. Then:
\[ I^{\text{PO}}(\theta, \alpha) = \beta(\mu - \theta) + (\delta(\theta; \beta))^{2}/2 - \alpha ME(g^{O} - g^{I}). \]

Similarly, let \( I^{\text{PW}}(\theta, \alpha) = V_{I}(g^{P}(\theta, \mu, G^{P})) - V_{I}(g^{W}(\theta, \tilde{\theta}, G^{W})) \) denote the type-specific value of a policy of privacy over a policy of waiver. Then:

\[
I^{\text{PW}}(\theta, \alpha) = \beta(\mu - \mu(\theta^{w})) - \alpha ME(g^{W} - g^{P}) \text{ for } \theta_{i} < \theta^{w}; \text{ and } \]
\[
\beta(\mu - \theta) + (\delta(\theta; \beta))^{2}/2 - \alpha ME(g^{W} - g^{P}), \text{ for } \theta_{i} \geq \theta^{w}. \]

Finally, let \( I^{\text{WO}}(\theta, \alpha) = V_{I}(g^{W}(\theta_{i}, \theta_{i}, G^{W})) - V_{I}(g^{O}(\theta_{i}, \theta_{i}, G^{O})) \) denote the type-specific value of a policy of waiver over a policy of publicity. Then:

\[
I^{\text{WO}}(\theta, \alpha) = \beta(\mu(\theta^{W}) - \theta_{i}) + (\delta(\theta; \beta))^{2}/2 - \alpha ME(g^{O} - g^{W}), \text{ for } \theta_{i} < \theta^{W}; \text{ and } \]
\[
- \alpha ME(g^{O} - g^{W}), \text{ for } \theta_{i} \geq \theta^{W}. \]

The functions \( I^{\text{PO}}(\theta, \alpha) \), \( I^{\text{PW}}(\theta, \alpha) \), and \( I^{\text{WO}}(\theta, \alpha) \) are continuous in both arguments and strictly decreasing in \( \alpha \); the latter two functions have portions that are constant with respect to \( \theta_{i} \), but they are strictly decreasing in \( \theta_{i} \) over the non-constant regions.

We first determine conditions under which there will be non-trivial sets of types who prefer each policy in a binary comparison. In particular, let \( \tilde{\alpha}^{J} \), for \( IJ = PO, PW, WO \), be the value of \( \alpha \) for which \( \theta_{i} = 0 \) is indifferent between policy \( I \) and policy \( J \) (for this and any higher value of \( \alpha \), policy \( J \) will be preferred to policy \( I \) for all types). Then \( \tilde{\alpha}^{J} \) is defined uniquely by \( I^{IJ}(0, \tilde{\alpha}^{J}) = 0 \), yielding:

\[
\tilde{\alpha}^{PO} = \beta\mu/(ME(g^{O} - g^{P})); \\
\tilde{\alpha}^{PW} = \beta(\mu - \mu(\theta^{w}))/(ME(g^{W} - g^{P})); \\
\tilde{\alpha}^{WO} = \beta\mu(\theta^{W})/(ME(g^{O} - g^{W})).
\]

Provided that \( \alpha < \min \{ \tilde{\alpha}^{J} \} \), there will be at least some (low) types who prefer policy \( I \) to policy \( J \) in a binary comparison. In order to have at least some (high) types who prefer policy \( J \) to policy \( I \) in a binary comparison, it must be that \( I^{IJ}(\tilde{\theta}, \alpha) < 0 \); our hypothesis that \( \theta^{W} < \tilde{\theta} \) is enough to
guarantee that this holds for all \( \alpha > 0 \).

CLAIM 1: If \( 0 < \alpha < \min \{ \alpha^J \} \), then:

(i) there exists a unique \( \theta^J I J(\alpha) \in (0, \bar{\theta}) \) such that \( \Gamma^J I J(\theta^J I J(\alpha), \alpha) = 0 \);

(ii) moreover, \( \theta^WO(\alpha) < \theta^W < \theta^PW(\alpha) \) and \( \theta^WO(\alpha) < \theta^PO(\alpha) < \theta^PW(\alpha) \).

PROOF OF CLAIM 1:

By construction, if \( 0 < \alpha < \min \{ \alpha^J \} \), then \( \Gamma^J I J(0, \alpha) > 0 \) and \( \Gamma^J I J(\theta^G I J, \alpha) < 0 \), for all \( I J \). First consider \( I J = PO \). The function \( \Gamma^P O(\theta, \alpha) \) is continuous and strictly decreasing in \( \theta \); therefore there exists a unique value \( \theta^J I J(\alpha) \in (0, \bar{\theta}) \) such that \( \Gamma^J I J(\theta^J I J(\alpha), \alpha) = 0 \). Next consider \( I J = PW \). The function \( \Gamma^P W(\theta, \alpha) \) is constant at a positive level for \( \theta_i < \theta^W \), and \( \Gamma^P W(\theta, \alpha) = \Gamma^P O(\theta, \alpha) + E(g^O - g^W) \) for \( \theta_i \geq \theta^W \). Since this is a continuous and strictly decreasing function, there is a unique value \( \theta^P W(\alpha) \in (\theta^W, \bar{\theta}) \) such that \( \Gamma^P W(\theta^P W(\alpha), \alpha) = 0 \). Moreover, this implies that \( \Gamma^P O(\theta^P W(\alpha), \alpha) = -E(g^O - g^W) < 0 \), so \( \theta^PO(\alpha) < \theta^P W(\alpha) \).

Finally, consider \( I J = WO \). The function \( \Gamma^W O(\theta, \alpha) \) is constant at a negative level for \( \theta_i \geq \theta^W \); it is a continuous and strictly decreasing function for \( \theta_i < \theta^W \). Therefore, there is a unique value \( \theta^W O(\alpha) \in (0, \theta^W) \) such that \( \Gamma^W O(\theta^W O(\alpha), \alpha) = 0 \). Moreover, evaluating \( \Gamma^P O \) at this level yields \( \Gamma^P O(\theta^W O(\alpha), \alpha) = \Gamma^P W(0, \alpha) > 0 \), so \( \theta^W O(\alpha) < \theta^P O(\alpha) \).

Note that for the special case of \( \alpha = 0 \) the claim above still holds with the following minor modifications. Now the function \( \Gamma^W O(\theta, \alpha) \) starts out positive and declines to zero at \( \theta^W \); moreover, it remains constant at zero for \( \theta_i \geq \theta^W \). Thus, the equation \( \Gamma^W O(\theta^W O(\alpha), \alpha) = 0 \) is satisfied by all members of the set \([\theta^W, \bar{\theta}]\); we take the left-most element as \( \theta^W O(\alpha) \), and thus \( \theta^W O(\alpha) = \theta^W \). The rest of the claim continues to hold as stated.
Given the ordering $\theta^{WO}(\alpha) < \theta^{PO}(\alpha) < \theta^{PW}(\alpha)$ derived above, it is straightforward to show that no type finds $W$ to be the best policy. The preference orderings are as follows and are illustrated in Figure 3 in the main text:

- For $\theta \in [0, \theta^{WO}(\alpha))$: $P > W > O$ (with $W - O$ at $\theta^{WO}(\alpha)$)
- For $\theta \in (\theta^{WO}(\alpha), \theta^{PO}(\alpha))$: $P > O > W$ (with $P - O > W$ at $\theta^{PO}(\alpha)$)
- For $\theta \in (\theta^{PO}(\alpha), \theta^{PW}(\alpha))$: $O > P > W$ ($O > P - W$ at $\theta^{PW}(\alpha)$)
- For $\theta \in (\theta^{PW}(\alpha), \bar{\theta}]$: $O > W > P$

Now we relax the assumption that $\alpha < \min \{\alpha^{IJ}\}, IJ = PO, PW, WO$. It is straightforward to show that $\tilde{\alpha}^{PO}$ must lie between $\tilde{\alpha}^{PW}$ and $\tilde{\alpha}^{WO}$, but we are unable to determine in general whether $\tilde{\alpha}^{PW} < \tilde{\alpha}^{WO}$ or $\tilde{\alpha}^{WO} < \tilde{\alpha}^{PW}$ (however, if $\tilde{\alpha}^{WO} < \tilde{\alpha}^{PW}$, then $W$ can never be interim-optimal for any type because $I^{WO}(0, \alpha) < 0$, implying that $O$ is preferred to $W$ for all types).

As claimed in the text, there are conditions under which some types will most-prefer a policy of $W$; these conditions are now described. First, it can be shown that $\tilde{\alpha}^{PW} < \tilde{\alpha}^{WO}$ for the case in which $\theta$ is distributed uniformly on $[0, \bar{\theta}]$. For $\tilde{\alpha}^{PW} < \alpha < \tilde{\alpha}^{WO}$, all types strictly prefer $P$ to $W$, while those in $[0, \theta^{WO}(\alpha))$ also strictly prefer $W$ to $O$. So it is possible for some types to interim-prefer $W$ to both $P$ and $O$ (however, this set is limited by the fact that $\theta^{WO}(\alpha) < \theta^{W}$ still holds). Notice that the types who interim-prefer $W$ to both $P$ and $O$ will exercise privacy under a policy of $W$ (since they are $< \theta^{W}$), but hope to gain both from higher types who also choose privacy and from the disclosures and distortions of even higher types.