Econometrica Supplementary Material

SUPPLEMENT TO "PERSUASION OF A PRIVATELY INFORMED RECEIVER"

(*Econometrica*, Vol. 85, No. 6, November 2017, 1949–1964)

ANTON KOLOTILIN
School of Economics, UNSW Business School

TYMOFIY MYLOVANOV
Department of Economics, University of Pittsburgh and Kyiv School of Economics

ANDRIY ZAPECHELNYUK
School of Economics and Finance, University of St Andrews

MING LI
Department of Economics, Concordia University and CIREQ

APPENDIX A: MULTIPLE ACTIONS

IN THIS APPENDIX, we allow the receiver to make a choice among multiple actions. We characterize the implementable receiver's interim utilities and show that the sender can generally implement a strictly larger set of the receiver's interim utilities by persuasion mechanisms than by experiments. We also formulate the sender's optimization problem and show that the sender can achieve a strictly higher expected utility by persuasion mechanisms than by experiments.

A.1. *Preferences*

Let $A = \{0, 1, \dots, n\}$ be a finite set of actions available to the receiver. The state $\omega \in \Omega = [0, 1]$ and the receiver's type $r \in R = [0, 1]$ are independent and have distributions F and G. We continue to assume that the receiver's utility is linear in the state for every type and every action.

It is convenient to define the receiver's and sender's utilities, $u(\omega, r, a)$ and $v(\omega, r, a)$, recursively by the utility difference between each two consecutive actions. For each $a \in \{1, \dots, n\}$,

$$u(\omega, r, a) - u(\omega, r, a-1) = b_a(r)\big(\omega - x_a(r)\big),$$
$$v(\omega, r, a) - v(\omega, r, a-1) = z_a(r) + \rho(r)\big(u(\omega, r, a) - u(\omega, r, a-1)\big),$$

and the utilities from action $a = 0$ are normalized to zero, $u(\omega, r, 0) = v(\omega, r, 0) = 0$ for all ω and all r.

For each $a \in \{1, \dots, n\}$, the receiver's and sender's utilities can be expressed as

$$u(\omega, r, a) = \left(\sum_{i=1}^{a} b_i(r)\right)\omega - \left(\sum_{i=1}^{a} b_i(r)x_i(r)\right),$$

Anton Kolotilin: akolotilin@gmail.com
Tymofiy Mylovanov: mylovanov@gmail.com
Andriy Zapechelnyuk: azapech@gmail.com
Ming Li: ming.li@concordia.ca

 DOI: 10.3982/ECTA13251

$$v(\omega, r, a) = \left(\sum_{i=1}^{a} z_i(r)\right) + \rho(r)u(\omega, r, a).$$

We assume that $b_a(r) > 0$ for all r and all $a \in \{1, \dots, n\}$. This assumption means that every type r prefers higher actions in higher states. Note that $x_a(r)$ is the cutoff state at which the receiver of type r is indifferent between two consecutive actions $a-1$ and a. Define $x_0(r) = -\infty$ and $x_{n+1}(r) = \infty$.

Denote by $\bar{x}_a(r)$ the cutoff truncated to the unit interval,

$$\bar{x}_a(r) = \max\bigl\{0, \min\bigl\{1, x_a(r)\bigr\}\bigr\}.$$

We assume that the cutoffs are ordered on $[0, 1]$ such that

$$\bar{x}_1(r) \le \bar{x}_2(r) \le \cdots \le \bar{x}_n(r) \quad \text{for all } r \in R.$$

Thus, type r optimally chooses action a on the interval of states $(\bar{x}_a(r), \bar{x}_{a+1}(r))$.[1]

A.2. *Experiments*

Because the receiver's utility is linear in the state for every type and every action, every experiment σ can be equivalently described by the probability that the posterior mean state is below a given value $x \in \Omega$,

$$H_\sigma(x) = \int_\Omega \sigma(x|\omega)\,\mathrm{d}F(\omega).$$

In fact, as in Blackwell (1951), Rothschild and Stiglitz (1970), and Gentzkow and Kamenica (2016), it is convenient to describe an experiment by a convex function $C_\sigma : \mathbb{R} \to \mathbb{R}$ defined as

$$C_\sigma(x) = \int_x^\infty \bigl(1 - H_\sigma(m)\bigr)\,\mathrm{d}m.$$

Observe that by (9), for every experiment σ, we have $C_\sigma(r) = U_\sigma(r)$ for all r, where $U_\sigma(r)$ is the receiver's interim utility under σ in the problem of Section 2, with two actions and $u(\omega, r, a) = a(\omega - r)$. Hence, by Theorem 1, the set of all C_σ is equal to

$$\mathcal{C} = \{C : \underline{C} \le C \le \overline{C} \text{ and } C \text{ is convex}\},$$

where $\overline{C}$ and $\underline{C}$ correspond to the full and no disclosure experiments,

$$\overline{C}(x) = \int_x^\infty \bigl(1 - F(m)\bigr)\,\mathrm{d}m,$$

$$\underline{C}(x) = \max\bigl\{\mathbb{E}[\omega] - x, 0\bigr\}.$$

[1]This assumption ensures that the actions that can be optimal for type r are consecutive. If actions $a-1$ and $a+1$ are optimal for type r under states ω' and ω'', then there must be a state between ω' and ω'' where action a is optimal. This assumption simplifies the exposition. Relaxing this assumption poses no difficulty: it will only require us, for each type r, to omit from the analysis the actions that are never optimal for this type.

A.3. *Implementable Interim Utilities*

The expected utility of type r under experiment σ is equal to

$$U_\sigma(r) = \int_\Omega \Big(\max_{a\in A} u(m, r, a)\Big)\, \mathrm{d}H_\sigma(m) \quad \text{for all } r \in R.$$

PROPOSITION 1: *U is implementable by an experiment if and only if there exists $C \in \mathcal{C}$ such that*

$$U(r) = \sum_{a=1}^{n} b_a(r) C\big(x_a(r)\big) \quad \textit{for all } r \in R. \tag{12}$$

PROOF: The receiver's interim utility under experiment σ is

$$U_\sigma(r) = \sum_{a=0}^{n} \int_{x_a(r)}^{x_{a+1}(r)} u(m, r, a)\, \mathrm{d}H_\sigma(m) = \sum_{a=1}^{n} \int_{x_a(r)}^{\infty} b_a(r)\big(m - x_a(r)\big)\, \mathrm{d}H_\sigma(m),$$

where we used $u(m, r, 0) = 0$. For each $a \in \{1, \dots, n\}$, integration by parts yields

$$\int_{x_a(r)}^{\infty} b_a(r)\big(m - x_a(r)\big)\, \mathrm{d}H_\sigma(m) = -b_a(r)\big(m - x_a(r)\big)\big(1 - H_\sigma(m)\big)\Big|_{x_a(r)}^{\infty} + b_a(r)\int_{x_a(r)}^{\infty} \big(1 - H_\sigma(m)\big)\, \mathrm{d}m = b_a(r) C_\sigma\big(x_a(r)\big).$$

Summing up the above over $a \in \{1, \dots, n\}$ yields (12).

It follows from Section A.2 that the set of the receiver's interim utilities implementable by experiments is equal to the set of functions that satisfy (12) for every $C \in \mathcal{C}$. *Q.E.D.*

A persuasion mechanism can be described by a (possibly, infinite) menu of experiments, Σ. The receiver of type r chooses one experiment from the menu and then observes messages only from this experiment. Obviously, the receiver chooses the experiment that maximizes his expected utility,

$$U_\Sigma(r) = \max_{\sigma\in\Sigma} U_\sigma(r) \quad \text{for all } r \in R.$$

By Proposition 1, it is immediate that the receiver's interim utility U is implementable if and only if there exists a menu $\mathcal{C}_\Sigma \subset \mathcal{C}$ such that

$$U(r) = \max_{C\in\mathcal{C}_\Sigma} \left\{ \sum_{a=1}^{n} b_a(r) C\big(x_a(r)\big) \right\} \quad \text{for all } r \in R.$$

Theorem 1 shows that the sender can implement the same set of receiver's interim utilities by experiments as by persuasion mechanisms. With more than two actions, however, the sender can generally implement a strictly larger set of interim utilities by persuasion mechanisms than by experiments, as shown in Example 2.

EXAMPLE 2: Let $A = \{0, 1, 2\}$, F admit a strictly positive density, and $u(\omega, r, a)$ be continuous in r for all ω and all a. Furthermore, suppose that there exist two types $r' < r''$, such that for all $r \in (r', r'')$,

$$0 < x_1(r') < x_1(r) < x_1(r'') < x_2(r') < x_2(r) < x_2(r'') < 1.$$

Consider a persuasion mechanism consisting of the menu of two experiments represented by partitions $\{P', P''\}$, where P' and P'' are the first-best partitions for types r' and r'',

$$P' = \{[0, x_1(r')), [x_1(r'), x_2(r')), [x_2(r'), 1]\},$$
$$P'' = \{[0, x_1(r'')), [x_1(r''), x_2(r'')), [x_2(r''), 1]\}.$$

Types r' and r'' choose, respectively, P' and P'' and get their maximum possible utilities $\overline{U}(r')$ and $\overline{U}(r'')$. By the continuity of $u(\omega, r, a)$ in r, there exists a type $r^* \in (r', r'')$ who is indifferent between choosing P' and P''. By this indifference,

$$L_1(P') + L_2(P') = L_1(P'') + L_2(P''),$$

where, for each $a \in \{1, 2\}$,

$$L_a(P') = \int_{x_a(r^*)}^{x_a(r')} \left(u(\omega, r^*, a) - u(\omega, r^*, a-1)\right) \mathrm{d}F(\omega)$$

denotes the utility loss of type r^* from using cutoff $x_a(r')$ rather than his first-best cutoff $x_a(r^*)$ to decide between actions $a-1$ and a. Analogously, for each $a \in \{1, 2\}$, we define $L_a(P'')$.

Figure 4 illustrates this example. The three blue lines depict the utility of the receiver of type r^* from taking action a, $u(\omega, r^*, a)$, for each $a = 0, 1, 2$. The kinked solid blue line is the utility of type r^* from taking the optimal action, $\max_{a\in\{0,1,2\}} u(\omega, r^*, a)$. In Figure 4, the loss of type r^* from experiment P' relative to the first best, $L_1(P') + L_2(P')$, is the total area of the two shaded triangles (assuming that ω is uniformly distributed). Similarly, the loss of type r^* from experiment P'' relative to the first best, $L_1(P'') + L_2(P'')$, is the total area of the two hatched triangles. For type r^*, these shaded and hatched areas are equal, so type r^* is indifferent between the two experiments.

An experiment that gives the maximum possible utilities $\overline{U}(r')$ and $\overline{U}(r'')$ to types r' and r'' must at least communicate the common refinement of partitions P' and P''. Therefore, the utility of type r^* under such an experiment is at least

$$\overline{U}(r^*) - \min\{L_1(P'), L_1(P'')\} - \min\{L_2(P'), L_2(P'')\},$$

which is strictly larger than his utility under the persuasion mechanism,

$$\overline{U}(r^*) - L_1(P') - L_2(P'),$$

unless $L_1(P') = L_1(P'')$ and $L_2(P') = L_2(P'')$.

In Figure 4, for type r^*, the loss $L_1(P'')$ (left hatched triangle) is smaller than the loss $L_1(P')$ (left shaded triangle). Similarly, the loss $L_2(P')$ (right shaded triangle) is smaller than the loss $L_2(P'')$ (right hatched triangle). Hence, the total loss is smaller under the experiment that is the coarsest common refinement of P' and P'' (the area of the smaller shaded and hatched triangles) than under either experiment P' or experiment P''.

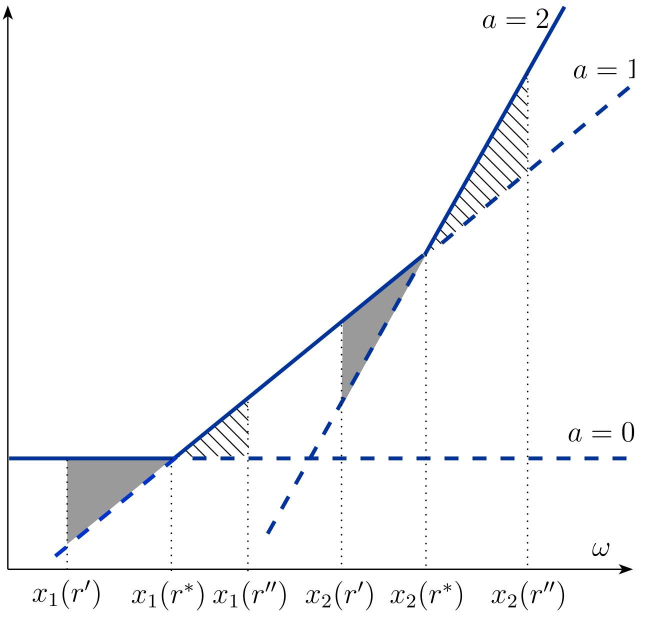

FIGURE 4.—The utility of type r^* in Example 1.

A.4. *Sender's Problem*

In this section, we impose the following additional assumptions. For each $a \in \{0, 1, \dots, n\}$, function $x_a(r)$ is strictly increasing, and its range contains Ω. Let function $r_a(x)$ be the inverse of function $x_a(r)$. Moreover, for each $a \in \{0, 1, \dots, n\}$, function $z_a(r)$ is differentiable, and function $G(r_a(x))$ is twice differentiable.

For a given experiment σ, the sender's expected utility conditional on the receiver's type being r is[2]

$$V_\sigma(r) = \rho(r)U(r) + \sum_{a=1}^{n}\left(\sum_{i=1}^{a} z_i(r)\right)\big(H_\sigma\big(x_{a+1}(r)\big) - H_\sigma\big(x_a(r)\big)\big).$$

We now express the sender's expected utility as a function of C_σ, and all the model parameters are summarized in function I, the same way as in Lemma 2 in Section 4.

LEMMA 3: *For every experiment* σ,

$$\int_R V_\sigma(r)\,\mathrm{d}G(r) = K + \int_\Omega C_\sigma(x) I(x)\,\mathrm{d}x,$$

where K *is a constant independent of* σ *and*

$$I(x) = \sum_{a=1}^{n}\left(\frac{\mathrm{d}}{\mathrm{d}x}\left(z_a\big(r_a(x)\big)\frac{\mathrm{d}}{\mathrm{d}x}G\big(r_a(x)\big)\right) + \rho\big(r_a(x)\big)b_a\big(r_a(x)\big)\frac{\mathrm{d}}{\mathrm{d}x}G\big(r_a(x)\big)\right).$$

[2]For each r where $H_\sigma(x_{a+1}(r))$ is discontinuous, this formula assumes that type r breaks the indifference in favor of action a if the posterior mean state is $x_{a+1}(r)$. This assumption is innocuous because G admits a density, and there are at most countably many discontinuities of H_σ.

PROOF: We have

$$\sum_{a=1}^{n}\left(\sum_{i=1}^{a} z_i(r)\right)\bigl(H_\sigma\bigl(x_{a+1}(r)\bigr)-H_\sigma\bigl(x_a(r)\bigr)\bigr)=\sum_{a=1}^{n} z_a(r)\bigl(1-H_\sigma\bigl(x_a(r)\bigr)\bigr)$$
$$=-\sum_{a=1}^{n} z_a(r)C'_\sigma\bigl(x_a(r)\bigr),$$

where we used $x_{n+1}(r)=\infty$ (hence, $H_\sigma(x_{n+1}(r))=1$) and $1-H_\sigma(x)=-C'_\sigma(x)$. By Proposition 1, we, thus, obtain

$$V(r)=\rho(r)U(r)+\sum_{a=1}^{n}\left(\sum_{i=1}^{a} z_i(r)\right)\bigl(H_\sigma\bigl(x_a(r)\bigr)-H_\sigma\bigl(x_{a+1}(r)\bigr)\bigr)$$
$$=\rho(r)U(r)-\sum_{a=1}^{n} z_a(r)C'_\sigma\bigl(x_a(r)\bigr)=\sum_{a=1}^{n}\bigl(\rho(r)b_a(r)C_\sigma\bigl(x_a(r)\bigr)-z_a(r)C'_\sigma\bigl(x_a(r)\bigr)\bigr).$$

Fix $a\in\{1,\dots,n\}$ and define the variable $x=x_a(r)$. Hence, $r=r_a(x)$. Using this variable change, we have

$$\int_R \bigl(\rho(r)b_a(r)C_\sigma\bigl(x_a(r)\bigr)-z_a(r)C'_\sigma\bigl(x_a(r)\bigr)\bigr)\,\mathrm{d}G(r)$$
$$=\hat{K}_a+\int_\Omega \bigl(\rho\bigl(r_a(x)\bigr)b_a\bigl(r_a(x)\bigr)C_\sigma(x)-z_a\bigl(r_a(x)\bigr)C'_\sigma(x)\bigr)\,\mathrm{d}G\bigl(r_a(x)\bigr),$$

where $\hat{K}_a$ is a constant independent of σ because $\Omega=[0,1]\subset[x_a(0),x_a(1)]$, and, for all σ and all $x\notin\Omega$, we have $C_\sigma(x)=\max\{0,\mathbb{E}[\omega]-x\}$. Now we integrate by parts

$$\int_\Omega z_a\bigl(r_a(x)\bigr)C'_\sigma(x)\,\mathrm{d}G\bigl(r_a(x)\bigr)=\tilde{K}_a-\int_\Omega C_\sigma(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(z_a\bigl(r_a(x)\bigr)\frac{\mathrm{d}}{\mathrm{d}x}G\bigl(r_a(x)\bigr)\right)\mathrm{d}x$$
$$=\tilde{K}_a-\int_\Omega C(x)\frac{\mathrm{d}}{\mathrm{d}x}\left(z_a\bigl(r_a(x)\bigr)\frac{\mathrm{d}}{\mathrm{d}x}G\bigl(r_a(x)\bigr)\right)\mathrm{d}x,$$

where $\tilde{K}_a$ is a constant independent of σ because, for all σ, we have $C_\sigma(0)=\mathbb{E}[\omega]$ and $C_\sigma(1)=0$. Thus, we obtain

$$\int_R \bigl(\rho(r)b_a(r)C_\sigma\bigl(x_a(r)\bigr)-z_a(r)C'_\sigma\bigl(x_a(r)\bigr)\bigr)\,\mathrm{d}G(r)$$
$$=K_a+\int_\Omega\left(\rho\bigl(r_a(x)\bigr)b_a\bigl(r_a(x)\bigr)\frac{\mathrm{d}G\bigl(r_a(x)\bigr)}{\mathrm{d}x}+\frac{\mathrm{d}}{\mathrm{d}x}\left(z_a\bigl(r_a(x)\bigr)\frac{\mathrm{d}}{\mathrm{d}x}G\bigl(r_a(x)\bigr)\right)\right)C_\sigma(x)\,\mathrm{d}x,$$

where $K_a=\hat{K}_a+\tilde{K}_a$. Summing the above over $a\in\{1,\dots,n\}$, we obtain $K+\int_\Omega C_\sigma(x)\times I(x)\,\mathrm{d}x$, where $K=\sum_a K_a$, and I is defined in Lemma 3. *Q.E.D.*

The sender's optimal experiment is described by a function $C\in\mathcal{C}$ that solves

$$\max_{C\in\mathcal{C}}\int_\Omega C(x)I(x)\,\mathrm{d}x.$$

The solutions to this problem are characterized by Theorem 2.

As shown in Section A.3, when the receiver has more than two actions, the sender can implement a strictly larger set of receiver's interim utilities by persuasion mechanisms than by experiments. We now show that the set of implementable interim actions is also strictly larger under persuasion mechanisms. Therefore, even if the sender cares only about the receiver's action, and not his utility, $\rho(r) = 0$ for all r, the sender can achieve a strictly larger expected utility under persuasion mechanisms.

EXAMPLE 2—Continued: In addition, let there exist $x_2^* \in (x_1(r''), x_2(r'))$ such that $\mathbb{E}[\omega|\omega \geq x_2^*] = x_2(r'')$ and $\mathbb{E}[\omega|\omega < x_2^*] < x_1(r')$.

An experiment that maximizes the probability of action $a = 2$ for type r'' must send message $x_2(r'')$ if and only if $\omega \in [x_2^*, 1]$. Under any such experiment, type r' takes action $a = 2$ if and only if $\omega \in [x_2^*, 1]$, because for $\omega < x_2^*$, this experiment must generate messages distinct from $x_2(r'')$ and, thus, below x_2^*, which is in turn below $x_2(r')$.

Consider now a persuasion mechanism consisting of the menu of two experiments represented by the following partitions:

$$P' = \{[0, x_2^* - \varepsilon)\backslash[x_1(r'), x_1(r'')), [x_1(r'), x_1(r'')), [x_2^* - \varepsilon, 1]\},$$
$$P'' = \{[0, x_2^* - \varepsilon), [x_2^* - \varepsilon, x_2^*), [x_2^*, 1]\},$$

where $\varepsilon > 0$ is sufficiently small. Type r' strictly prefers P' (to P'') because, for a sufficiently small ε, the benefit of taking action $a = 1$ (rather than $a = 0$) on $[x_1(r'), x_1(r''))$ exceeds the cost of taking action $a = 2$ (rather than $a = 1$) on $[x_2^* - \varepsilon, x_2^*)$. Type r'' is indifferent between P'' and P', because under both partitions he weakly prefers to take action $a = 0$ on $[0, x_2^* - \varepsilon)$ and action $a = 1$ on $[x_2^* - \varepsilon, 1]$. Therefore, under this mechanism, type r'' takes action $a = 2$ if and only if $\omega \in [x_2^*, 1]$, but type r' takes action $a = 2$ if and only if $\omega \in [x_2^* - \varepsilon, 1]$. As shown above, these probabilities of $a = 2$ for types r' and r'' cannot be achieved by any experiment.

Finally, an optimal persuasion mechanism need not be an experiment. Suppose that the sender cares only about action $a = 2$, that is, $\rho(r) = z_0(r) = z_1(r) = 0$ and $z_2(r) = 1$ for all r. Also suppose that the support of G contains only r' and r'' with r'' being likely enough, so that the sender's optimal experiment maximizes the probability of action $a = 2$ for type r''. The persuasion mechanism constructed above gives a strictly larger expected utility to the sender than any experiment.

APPENDIX B: NONLINEAR UTILITIES

In this appendix, we allow the sender's and receiver's utilities to be nonlinear in the state. We characterize conditions under which persuasion mechanisms are equivalent to experiments, and show, in particular, that cutoff mechanisms are equivalent to experiments. We also show that the equivalence of implementation by persuasion mechanisms and by experiments generally fails.

B.1. *Preferences*

As in Section 2, the receiver has two actions, $A = \{0, 1\}$, the set of states is $\Omega = [0, 1]$, and the set of receiver's types is $R = [0, 1]$. The receiver's utility, however, is

$$u(\omega, r, a) = au(\omega, r),$$

where $u(\omega, r)$ is differentiable, strictly increasing in ω, and strictly decreasing in r. We also normalize the utility such that, for each $\omega \in \Omega$,

$$u(\omega, \omega) = 0.$$

The sender's utility is

$$v(\omega, r, a) = av(\omega, r),$$

where $v(\omega, r)$ is differentiable. State ω and type r are independent and have distributions F and G.[3]

B.2. *Characterization of Experiments*

We start with the characterization of persuasion mechanisms that are equivalent to experiments.

PROPOSITION 2: *An incentive-compatible persuasion mechanism π is equivalent to an experiment if and only if $\pi(\omega, r)$ is nonincreasing in r for every $\omega \in \Omega$.*

Intuitively, because for each experiment σ, the distribution $\sigma(r|\omega)$ of r conditional on each state ω is nondecreasing in r, each $\pi(\omega, r) \in [1 - \sigma(r|\omega), 1 - \sigma(r_-|\omega)]$ (see (6)) is nonincreasing in r.

PROOF: Consider a mechanism π that is equivalent to an experiment σ. Since $\sigma(r|\omega)$ is a distribution function of r conditional on ω, it is nondecreasing in r for each ω. Then, by (6), $\pi(\omega, r)$ is nonincreasing in r for each ω.

Conversely, let $\pi(r, \omega)$ be nonincreasing in r for all ω. For every ω and r, define $\sigma(r|\omega) = 1 - \pi(r_+, \omega)$, where $\pi(r_+, \omega)$ denotes the right limit of $\pi(\cdot, \omega)$ at r. Since $\pi(r_+, \omega) \in [0, 1]$ is nonincreasing and right-continuous in r, the function $\sigma(r|\omega)$ is a distribution, which describes the distribution of messages for every given state ω. Thus, σ is an experiment. It remains to verify that the constructed experiment is direct and induces the same action by the receiver as mechanism π, that is, when the experiment sends a message r, then type r is indifferent between the two actions. For all r,

$$\begin{aligned} U_\pi(r) &= \int_\Omega u(\omega, r)\pi(\omega, r)\,\mathrm{d}F(\omega) = \int_\Omega u(\omega, r_+)\pi(\omega, r_+)\,\mathrm{d}F(\omega) \\ &= \int_\Omega u(\omega, r)\pi(\omega, r_+)\,\mathrm{d}F(\omega) = \int_\Omega u(\omega, r)\big(1 - \sigma(r|\omega)\big)\,\mathrm{d}F(\omega) = U_\sigma(r), \end{aligned}$$

where the first equality holds by the definition of U_π, the second by the absolute continuity of U_π (Theorem 1 of Milgrom and Segal (2002)), the third by the continuity of u in r, the fourth by the definition of σ, and the last by the definition of U_σ for direct experiments.

[3]Note that if ω and r are correlated, the analysis below applies if we impose strict monotonicity on function $\tilde{u}(\omega, r) = u(\omega, r)g(r|\omega)/g(r)$ rather than on u, where $g(r)$ and $g(r|\omega)$ denote, respectively, the marginal density of r and the conditional density of r for a given ω. This is because the receiver's interim utility under a mechanism π can be written as $U(r) = \int_\Omega \tilde{u}(\omega, r)\pi(\omega, r)\,\mathrm{d}F(\omega)$.

There exist left and right derivatives of U_π for all r (Theorem 3 of Milgrom and Segal (2002)) that satisfy

$$U'_\pi(r_+) = \int_\Omega \frac{\partial u(\omega, r)}{\partial r} \pi(r_+, \omega)\, \mathrm{d}F(\omega),$$
$$U'_\pi(r_-) = \int_\Omega \frac{\partial u(\omega, r)}{\partial r} \pi(r_-, \omega)\, \mathrm{d}F(\omega).$$

Since $U_\pi(r) = U_\sigma(r)$ and $\sigma(r|\omega) = 1 - \pi(r_+, \omega)$ for all r, we have

$$U'_\sigma(r_+) = \int_\Omega \frac{\partial u(\omega, r)}{\partial r} \big(1 - \sigma(r|\omega)\big)\, \mathrm{d}F(\omega),$$
$$U'_\sigma(r_-) = \int_\Omega \frac{\partial u(\omega, r)}{\partial r} \big(1 - \sigma(r_-|\omega)\big)\, \mathrm{d}F(\omega),$$

showing that type r is indifferent between the two actions upon receiving message r.

Q.E.D.

B.3. *Binary State*

Here we apply Proposition 2 to show that if there are only two states in the support of the prior F, then every incentive-compatible mechanism is equivalent to an experiment.

COROLLARY 2: *Let the support of F consist of two states. Then every incentive-compatible mechanism π is equivalent to an experiment.*

PROOF: Consider F whose support consists of two states, without loss of generality, $\{0, 1\}$, and let π be an incentive-compatible persuasion mechanism. By Proposition 2, it is sufficient to show that π is nonincreasing in r for all $r \in (0, 1)$. Incentive compatibility implies that for all $r, \hat{r} \in (0, 1)$,

$$\sum_{\omega=0,1} u(\omega, r)\big(\pi(\omega, r) - \pi(\omega, \widehat{r})\big) \Pr(\omega = 1) \geq 0. \tag{13}$$

Rewriting (13) twice, with $(r, \widehat{r}) = (r_2, r_1)$ and $(r, \widehat{r}) = (r_1, r_2)$, yields the inequalities

$$-\frac{u(0, r_2)}{u(1, r_2)} \delta(r_2, r_1, 0) \leq \delta(r_2, r_1, 1) \leq -\frac{u(0, r_1)}{u(1, r_1)} \delta(r_2, r_1, 0), \tag{14}$$

where $\delta(r_2, r_1, \omega) = (\pi(\omega, r_2) - \pi(\omega, r_1)) \Pr(\omega = 1)$. Because $u(0, r) < 0$ and $u(1, r) > 0$ for $r = r_1, r_2$, the monotonicity of u in r implies that

$$0 < -\frac{u(0, r_2)}{u(1, r_2)} \leq -\frac{u(0, r_1)}{u(1, r_1)} \quad \text{for } r_2 \leq r_1. \tag{15}$$

Combining (14) and (15) gives $\pi(\omega, r_2) \geq \pi(\omega, r_1)$ if $r_2 \leq r_1$ for each $\omega = 0, 1$. *Q.E.D.*

Note that if F has a two-point support, then the receiver's utility is linear in the state without loss of generality, and, hence, Theorem 1 applies. However, Corollary 2 makes a stronger statement, because it asserts that every incentive-compatible mechanism is equivalent to an experiment, not just implements the same receiver's interim utility.

B.4. *Beyond Binary State*

Suppose now that the support of the prior F consists of three states $\omega_1 < \omega_2 < \omega_3$ and let $f_i = \Pr(\omega_i) > 0$ for $i = 1, 2, 3$.

When there are at least three states and the utility of the receiver is nonlinear in (any transformation of) the state, then the posterior distribution of the state induced by an experiment can no longer be parameterized by a one-dimensional variable—such as the posterior mean state in the case of linear utilities, the posterior probability of one of the states in the case of binary-valued state, and the cutoff value in the case of cutoff mechanisms.

As a consequence, the interim action $q(r)$ and, hence, the sender's interim utility $V(r)$ are no longer pinned down by the receiver's interim utility $U(r)$.

PROPOSITION 3: *Let π_1 and π_2 be two mechanisms that are distinct for each $r \in (\omega_1, \omega_3)$ but implement the same differentiable receiver's interim utility U. Then, the interim action q is the same for π_1 and π_2 if and only if there exist functions b, c, and d such that $u(\omega, r) = c(r) + b(r)d(\omega)$ for each $(\omega, r) \in \{\omega_1, \omega_2, \omega_3\} \times (\omega_1, \omega_3)$.*

PROOF: For all $r \in (\omega_1, \omega_3)$ and $j = 1, 2$, we have

$$U(r) = \sum_{i=1}^{3} u(\omega_i, r)\pi_j(\omega_i, r) f_i,$$

$$U'(r) = \sum_{i=1}^{3} \frac{\partial u(\omega_i, r)}{\partial r} \pi_j(\omega_i, r) f_i,$$

where the first line holds by the definition of U, and the second line by the incentive compatibility of π.

The expected action, $q_{\pi_j}(r) = \sum_{i=1}^{3} \pi_j(\omega_i, r) f_i$, is the same across $j = 1, 2$ for each r if and only if the vectors $u(\omega, r)$, $\frac{\partial u(\omega,r)}{\partial r}$, and $\mathbf{1}$ are linearly dependent for each r. That is, for each $(\omega, r) \in \{\omega_1, \omega_2, \omega_3\} \times (\omega_1, \omega_3)$, there exist functions $\gamma(r)$ and $\mu(r)$ such that

$$\frac{\partial u(\omega, r)}{\partial r} + \mu(r) u(\omega, r) = \gamma(r). \tag{16}$$

The solution of differential equation (16) is given by

$$u(\omega, r) = e^{-\int_{\omega_1}^{r} \mu(x)\,\mathrm{d}x} \left(\eta(\omega) + \int_{\omega_1}^{r} \gamma(x) e^{\int_{\omega_1}^{x} \mu(y)\,\mathrm{d}y}\,\mathrm{d}x \right),$$

where function $\eta(\omega)$ satisfies the (initial) normalization condition $u(\omega, \omega) = 0$. This completes the proof with $b(r)$, $c(r)$, and $d(\omega)$ given by

$$\big(b(r), c(r), d(\omega)\big) = \left(e^{-\int_{\omega_1}^{r} \mu(x)\,\mathrm{d}x} \int_{\omega_1}^{r} \gamma(x) e^{\int_{\omega_1}^{x} \mu(y)\,\mathrm{d}y}\,\mathrm{d}x, e^{-\int_{\omega_1}^{r} \mu(x)\,\mathrm{d}x}, \eta(\omega) \right).$$

Q.E.D.

When the receiver's utility is nonlinear, the sender can implement a strictly larger set of the receiver's interim actions by persuasion mechanisms than by experiments. Therefore, the sender can achieve a strictly higher expected utility by persuasion mechanisms, even if her utility v is state-independent.

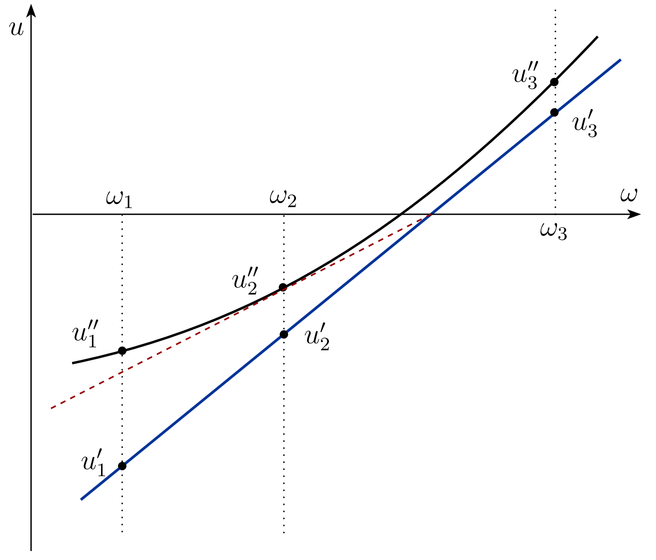

FIGURE 5.—The utilities of types r' (blue) and r'' (black) in Example 2.

EXAMPLE 3: Let there be three states, $\omega_1 < \omega_2 < \omega_3$, and two types of the receiver, $r' > r''$. Denote $u_i' = u(\omega_i, r')$ and $u_i'' = u(\omega_i, r'')$ for $i \in \{1, 2, 3\}$. Assume $u_2'' < 0 < u_3'$. Moreover, assume $u_1'/u_1'' > u_2'/u_2''$, which means that point (ω_1, u_1'') lies above the dashed line in Figure 5. Finally, assume that the probability masses (f_1, f_2, f_3) on the states satisfy $f_1 u_1' + f_3 u_3' < 0$ and $f_2 u_2' + f_3 u_3' < 0$.

Let Σ' be the set of all experiments that maximize the probability of action for type r'. It is easy to check that any $\sigma' \in \Sigma'$ induces type r' to act with probability 1 if $\omega = \omega_3$, with probability $-f_3 u_3'/(f_2 u_2')$ if $\omega = \omega_2$, and with probability 0 if $\omega = \omega_1$.

Observe that by the monotonicity of u in r, each message of $\sigma' \in \Sigma'$ that induces type r' to act also induces type $r'' < r'$ to act. Moreover, by the definition of Σ', each message of $\sigma' \in \Sigma'$ that induces type r' not to act can be sent only in states ω_1 or ω_2 where the utility of type r'' is negative by assumption, $u_1'' < u_2'' < 0$, so type r'' does not act either. Thus, for each experiment σ under which type r' acts with probability $f_3(1 - u_3'/u_2')$ (i.e., $\sigma \in \Sigma'$), type r'' acts with the same probability as type r'.

We now construct a persuasion mechanism that also maximizes the probability that type r' acts, but induces type r'' to act with a different probability. Let Σ'' be the set of all experiments σ'' that induce type r' to act with probability 1 if $\omega = \omega_3$, with probability 0 if $\omega = \omega_2$, and with probability $-f_3 u_3'/(f_1 u_1')$ if $\omega = \omega_1$. Consider a persuasion mechanism that consists of a menu of two experiments $\{\sigma', \sigma''\}$ with $\sigma' \in \Sigma'$ and $\sigma'' \in \Sigma''$. Notice that type r' is indifferent between σ' and σ'' as he obtains zero expected utility in either case. However, type r'' strictly prefers σ'' to σ', because $u_1'/u_1'' > u_2'/u_2''$ by assumption. Therefore, under this persuasion mechanism, type r' acts with probability $f_3(1 - u_3'/u_2')$, but type r'' acts with different probability $f_3(1 - u_3'/u_1')$.

Finally, when the receiver's utility is nonlinear, the set of receiver's interim utilities implementable by persuasion mechanisms (as compared to experiments) can be strictly larger.

EXAMPLE 3—Continued: In addition, let $r^* \in (r', r'')$ be such that $u_1'/u_1^* < u_2'/u_2^*$, where $u_i^* = u(\omega_i, r^*)$ for $i \in \{1, 2, 3\}$.

It is easy to check that Σ' and Σ'' are the sets of experiments σ' and σ'' that maximize the utility of types r^* and r'', respectively, subject to the constraint that type r' gets utility $\underline{U}(r') = 0$. Because Σ' and Σ'' do not intersect, no experiment can achieve the interim utility induced by a persuasion mechanism that consists of the menu of two experiments $\sigma' \in \Sigma'$ and $\sigma'' \in \Sigma''$.

APPENDIX C: LINEAR UTILITIES

This appendix extends our main results to the class of utility functions that are linear in the state. Specifically, we consider the model defined in Section 2, with the modification that the utilities are linear in the state and are arbitrary functions of the receiver's type.

Let the receiver's and sender's utilities be normalized to zero if the receiver does not act, $a = 0$, and be linear in the state if the receiver acts, $a = 1$,

$$u(\omega, r, a) = a \cdot b(\omega - t),$$

$$v(\omega, r, a) = a \cdot \big(c(\omega - t) + d\big),$$

where $r = (b, c, d, t) \in \mathbb{R}^4$ denotes the receiver's type. The type has distribution G that admits a differentiable density g, which is strictly positive on a compact set in $\mathbb{R}^4$ and zero everywhere else. The state $\omega \in \Omega = [0, 1]$ is independent of r and has distribution F.

Let H_σ be the distribution of the posterior mean induced by an experiment σ. As in Section A.2, it is convenient to describe σ by

$$C_\sigma(t) = \int_t^\infty \big(1 - H_\sigma(m)\big)\,\mathrm{d}m.$$

PROPOSITION 4: *For each experiment σ, the receiver's interim utility is*

$$U_\sigma(r) = |b| C_\sigma(t) + \min\{0, b\}\big(\mathbb{E}[\omega] - t\big). \tag{17}$$

There exist $K \in \mathbb{R}$ and $I : \mathbb{R} \to \mathbb{R}$ such that, for each σ, the sender's expected utility is

$$V_\sigma = K + \int_{t \in \mathbb{R}} C_\sigma(t) I(t)\,\mathrm{d}t. \tag{18}$$

Proposition 4 allows us to extend Theorems 1 and 2 to this setting. Recall from Section A.2 that the set of all C_σ is equal to

$$\mathcal{C} = \{C : \underline{C} \le C \le \overline{C} \text{ and } C \text{ is convex}\},$$

where $\overline{C}$ and $\underline{C}$ correspond to the full and no disclosure experiments. Each persuasion mechanism can be described by a (possibly, infinite) menu of experiments, Σ, which the receiver chooses from. By (17), for a given menu Σ, the receiver's interim utility is

$$\max_{\sigma \in \Sigma} U_\sigma(r) = |b| \Big(\max_{\sigma \in \Sigma} C_\sigma(t)\Big) + \min\{0, b\}\big(\mathbb{E}[\omega] - t\big).$$

Notice that $\max_{\sigma \in \Sigma} C_\sigma$ is the upper envelope of convex functions $C_\sigma \in \mathcal{C}$ and hence it is in $\mathcal{C}$. Therefore, by Proposition 4, any implementable pair of the sender's and receiver's expected utilities is implementable by an experiment. Moreover, the sender's problem can be expressed as

$$\max_{C\in\mathcal{C}} \int_{\mathbb{R}} C(t)I(t)\,\mathrm{d}t,$$

and Theorem 2 holds with U replaced by C.

PROOF OF PROPOSITION 4: Fix a type $r = (b, c, d, t) \in \mathbb{R}^4$ and evaluate

$$U_\sigma(r) = \int_0^1 \max\{0, b(m-t)\}\,\mathrm{d}H_\sigma(m).$$

Clearly, if $b = 0$, then $U_\sigma(r) = 0$. We now consider two cases, $b > 0$ and $b < 0$.

Case 1: $b > 0$. Given a posterior mean m, the receiver acts if and only if $t < m$. By integration by parts, the receiver's interim utility is

$$U_\sigma(r) = \int_t^1 b(m-t)\,\mathrm{d}H_\sigma(m) = bC_\sigma(t).$$

Again, by integration by parts, the sender's interim utility is

$$V_\sigma(r) = \int_t^1 \big(c(m-t)+d\big)\,\mathrm{d}H_\sigma(m) = cC_\sigma(t) - dC'_\sigma(t).$$

Case 2: $b < 0$. Given a posterior mean m, the receiver acts if and only if $t \geq m$. By integration by parts, the receiver's interim utility is

$$U_\sigma(r) = \int_0^t b(m-t)\,\mathrm{d}H(m) = -bC(t) + b\big(\mathbb{E}[\omega]-t\big).$$

Again, by integration by parts, the sender's interim utility is

$$V_\sigma(r) = \int_0^t \big(c(m-t)+d\big)\,\mathrm{d}H(m) = -cC(t) + dC'(t) + d + c\big(\mathbb{E}[\omega]-t\big).$$

We, thus, obtain (17).

We now show that the sender's expected utility is given by (18). Let $g(b, c, d|t)$ be the density of (b, c, d) conditional on t, and let $g_t(t)$ be the marginal density of t. Define

$$c_+(t) = \int cg(b, c, d|t)\mathbf{1}_{b>0}\,\mathrm{d}(b, c, d),$$

and

$$c_-(t) = \int cg(b, c, d|t)\mathbf{1}_{b<0}\,\mathrm{d}(b, c, d).$$

Similarly, define $d_+(t)$ and $d_-(t)$.

Fix t and take the expectation with respect to (b, c, d) on the set of $b > 0$:

$$\int_{(b,c,d)} V_\sigma(b, c, d, t)\mathbf{1}_{b>0}g(b, c, d|t)\,\mathrm{d}(b, c, d) = c_+(t)C_\sigma(t) - d_+(t)C'_\sigma(t).$$

Now, integrating with respect to t,

$$\int_{(b,c,d,t)} V_\sigma(b,c,d,t)\mathbf{1}_{b>0}\,\mathrm{d}G(b,c,d,t) = \int_t \big(c_+(t)C_\sigma(t) - d_+(t)C'_\sigma(t)\big)g_t(t)\,\mathrm{d}t$$
$$= \int_t \left(c_+(t)g_t(t) + \frac{\mathrm{d}}{\mathrm{d}t}\big[d_+(t)g_t(t)\big]\right)C_\sigma(t)\,\mathrm{d}t.$$

Similarly, for $b<0$,

$$\int_{(b,c,d)} V_\sigma(b,c,d,t)\mathbf{1}_{b<0}g(b,c,d|t)\,\mathrm{d}(b,c,d)$$
$$= -c_-(t)C_\sigma(t) + d_-(t)C'_\sigma(t) + d_-(t) + c_-(t)\big(\mathbb{E}[\omega]-t\big).$$

Now, integrating with respect to t,

$$\int_{(b,c,d,t)} V_\sigma(b,c,d,t)\mathbf{1}_{b<0}\,\mathrm{d}G(b,c,d,t) = -\int_t \big(c_-(t)C_\sigma(t) - d_-(t)C'_\sigma(t)\big)g_t(t)\,\mathrm{d}t + K$$
$$= -\int_t \left(c_-(t)g_t(t) + \frac{\mathrm{d}}{\mathrm{d}t}\big[d_-(t)g_t(t)\big]\right)C_\sigma(t)\,\mathrm{d}t + K,$$

where

$$K = \int_t \big(d_-(t) + c_-(t)\big(\mathbb{E}[\omega]-t\big)\big)g_t(t)\mathbf{1}_{b<0}\,\mathrm{d}t$$

is a constant independent of C_σ.

Since the measure of types with $b=0$ is zero, we obtain

$$\int_r V_\sigma(r)\,\mathrm{d}G(r) = \int_r V_\sigma(r)(\mathbf{1}_{b>0}+\mathbf{1}_{b<0})\,\mathrm{d}G(r) = \int_t I(t)C_\sigma(t)\,\mathrm{d}t + K,$$

where

$$I(t) = \big(c_+(t)-c_-(t)\big)g_t(t) + \frac{\mathrm{d}}{\mathrm{d}t}\big[\big(d_+(t)-d_-(t)\big)g_t(t)\big]. \qquad \textit{Q.E.D.}$$

REFERENCES

BLACKWELL, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 93–102. [2]

GENTZKOW, M., AND E. KAMENICA (2016): "A Rothschild-Stiglitz Approach to Bayesian Persuasion," *American Economic Review, Papers & Proceedings*, 106, 597–601. [2]

MILGROM, P., AND I. SEGAL (2002): "Envelope Theorems for Arbitrary Choice Sets," *Econometrica*, 70, 583–601. [8,9]

ROTHSCHILD, M., AND J. STIGLITZ (1970): "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2, 225–243. [2]

Co-editor Joel Sobel handled this manuscript.

Manuscript received 23 February, 2015; final version accepted 26 June, 2017; available online 28 June, 2017.