B Online Appendix

In this Online Appendix we provide formal statements and proofs of the claims made in Section 4.4 and 5 of "Bayesian Persuasion with Heterogeneous Priors,"by Alonso and Câmara.

B.1 Optimal Experiments to Persuade Skeptics and Believers

We complete Section 4.4 by characterizing properties of optimal experiments and describe a procedure to derive an optimal experiment. To describe this procedure, we now restrict attention to the case in which the sender is risk-neutral over the receiver's beliefs.

Proposition 8 *Suppose that **(A1)** and **(A2′)** hold, with G linear, card$(\Theta) > 2$, and that for each triplet $\theta_i, \theta_j, \theta_k \in \Theta$ of states, $(\theta_i, \theta_j, \theta_k)$ and (r_i^S, r_j^S, r_k^S) are not negatively collinear with respect to W. For each pair of states (θ_i, θ_j), define*

$$\Delta_{(i,j)} = -\left(r_j^S - r_i^S\right)(\theta_j - \theta_i). \tag{33}$$

If π^* is an optimal experiment, then, after each realization of π^*, the receiver puts positive probability in at most two states. Furthermore, for each state θ_i, there is a threshold $\xi_i \geq 0$ such that there is a realization of π^* induced by both states θ_i and θ_j if and only if $\Delta_{(i,j)} \geq \xi_i$.

Consequently, for every subset of states $\{\theta_i, \theta_j, \theta_k\}$, if either $\Delta_{(i,j)} \leq \min\{\Delta_{(i,k)}, \Delta_{(k,j)}\}$ or $\Delta_{(i,j)} < 0$, then there is no realization supported on both θ_i and θ_j.

Consider any pair $\theta_j > \theta_i$. The term $\Delta_{(i,j)}$ captures the value to the sender of "bundling" states θ_i and θ_j — the value of pooling these states into the same realization of the experiment. Pooling the states has positive value if and only if the receiver is a believer ($r_j^S < r_i^S$), conditional on the partition $\{\theta_i, \theta_j\}$. A positive-value bundle becomes more valuable when the differences $r_i^S - r_j^S$ and $\theta_j - \theta_i$ are larger. If state θ_i has more than one positive-value bundle, then the sender optimally allocates probability mass from θ_i across these bundles according to their value. Bundles with low positive value may be broken so that more probability mass can be assigned to higher-value bundles.

We now apply Proposition 8 to construct an algorithm to solve for the optimal experiment when there are three states, $\theta_1 < \theta_2 < \theta_3$ (see the proof of Proposition 8 for details):

Step 1: Compute the ratios $\frac{r_2^S - r_1^S}{\theta_2 - \theta_1}$ and $\frac{r_3^S - r_2^S}{\theta_3 - \theta_2}$. If the ratios are equal to each other and (weakly) negative, then no experimentation is optimal. Otherwise, proceed to Step 2.

Step 2: Compute the pooling values $\Delta_{(1,2)}$, $\Delta_{(2,3)}$ and $\Delta_{(3,1)}$. If all values are (weakly) negative, then a fully informative experiment is optimal. Otherwise, proceed to Step 3.

Step 3: Let θ_i and θ_j be the states with the lowest pooling value $\Delta_{(i,j)}$, and θ_k the remaining state. Construct experiment π_α as follows. There is a binary realization space $Z = \{z_i, z_j\}$. Likelihood functions are: state θ_i induces realization z_i with probability one; state θ_j induces z_j with probability one; state θ_k induces realization z_i with probability α and induces z_j with probability $1-\alpha$. The optimal experiment π_{α^*} is the one with α^* that maximizes the sender's expected payoff

$$\max_{\alpha\in[0,1]} \Pr_S[z_i|\pi_\alpha]E_R[\theta|z_i,\pi_\alpha] + \Pr_S[z_j|\pi_\alpha]E_R[\theta|z_j,\pi_\alpha]. \tag{34}$$

We can use this algorithm to solve the example from the introduction: $\Theta = \{1, 1.5, 2\}$, $p^S = (0.85, 0.10, 0.05)$ and $p^R = (0.10, 0.40, 0.55)$. The condition in Step 1 is not met, so we proceed to Step 2 and compute $\Delta_{1,1.5} = 4.125$, $\Delta_{1.5,2} = 0.075$ and $\Delta_{1,2} = 8.4$. Since they are positive, we proceed to Step 3. The lowest pooling value is $\Delta_{1.5,2}$; hence, we construct the binary realization space $Z = \{z_{1.5}, z_2\}$. State $\{1.5\}$ induces $z_{1.5}$ with probability one; state $\{2\}$ induces z_2 with probability one; and state $\{1\}$ induces $z_{1.5}$ with probability α. Given this experiment, (34) becomes

$$\begin{aligned}\max_{\alpha\in[0,1]} \quad & (\alpha 0.85 + 0.1)\left(1\frac{\alpha 0.85}{0.1+\alpha 0.85} + 1.5\frac{0.1}{0.1+\alpha 0.85}\right)\\ & + ((1-\alpha)0.85 + 0.05)\left(1\frac{(1-\alpha)0.85}{(1-\alpha)0.85+0.05} + 2\frac{0.05}{(1-\alpha)0.85+0.05}\right),\end{aligned}$$

and the sender's optimal choice is $\alpha^* = 1$.

In summary, the sender's primary concern is which bundles should be broken and which should be kept. When there are more than three states, the logic above can be used to eliminate all bundles with negative value and, for each triplet of states, eliminate the bundle with the lowest value. After all the "weak" bundles are eliminated, each group of states no longer "connected" with other groups of states can then be treated independently in the design of an optimal experiment.

Proof of Proposition 8: Proposition 5.i shows that the condition on each triplet $\theta_i, \theta_j, \theta_k \in \Theta$ implies that any realization of an optimal experiment leads to posterior beliefs supported on at most two states. For each pair (θ_i, θ_j), we now investigate under what conditions the optimal experiment has a realization induced by states θ_i and θ_j.

Denote by z_{ij} a realization induced by both states θ_i and θ_j. In particular, we allow z_{ii} to be a realization induced only by θ_i (and, thus, that fully reveals the state). For any experiment π, we have that the sender's expectation over its posterior expectations must equal the prior expectation — i.e., $\mathrm{E}_S^{\pi}\left[\mathrm{E}_S\left[\theta|z\right]\right] = \mathrm{E}_S\left[\theta\right]$. Therefore, if an experiment π^* maximizes the sender's expectation of the receiver's posterior expectation, it also maximizes the sender's expectation of the difference between the receiver's and the sender's expectation. That is, for an arbitrary π,

$$\mathrm{E}_S^{\pi^*}\left[\mathrm{E}_R\left[\theta|z\right]\right] \geq \mathrm{E}_S^{\pi}\left[\mathrm{E}_R\left[\theta|z\right]\right] \Leftrightarrow \mathrm{E}_S^{\pi^*}\left[\mathrm{E}_R\left[\theta|z\right] - \mathrm{E}_S\left[\theta|z\right]\right] \geq \mathrm{E}_S^{\pi}\left[\mathrm{E}_R\left[\theta|z\right] - \mathrm{E}_S\left[\theta|z\right]\right].$$

If a sender seeks to maximize the difference between the receiver's and her expectation of the state, her expected utility from an experiment π can be written as

$$\begin{aligned}
\mathrm{E}_S^{\pi}\left[\mathrm{E}_R\left[\theta|z\right] - \mathrm{E}_S\left[\theta|z\right]\right] &= \sum \mathrm{Pr}_S[z]\left(\left\langle q^R(z), \theta\right\rangle - \left\langle \frac{q^R(z) r^S}{\left\langle q^R(z), r^S\right\rangle}, \theta\right\rangle\right) \\
&= \sum \mathrm{Pr}_R[z]\left(\left\langle q^R(z), \theta\right\rangle\left\langle q^R(z), r^S\right\rangle - \left\langle q^R(z) r^S, \theta\right\rangle\right).
\end{aligned}$$

If an experiment induces realizations z_{ij} that are only supported on at most two states, then

$$\begin{aligned}
\left\langle q^R(z_{ij}), \theta\right\rangle\left\langle q^R(z_{ij}), r^S\right\rangle - \left\langle q^R(z_{ij}) r^S, \theta\right\rangle &= -q_i^R(z_{ij}) q_j^R(z_{ij})\left(r_j^S - r_i^S\right)\left(\theta_j - \theta_i\right) \\
&= q_i^R(z_{ij}) q_j^R(z_{ij}) \Delta_{(i,j)},
\end{aligned}$$

so that we can write

$$\mathrm{E}_S^{\pi}\left[\mathrm{E}_R\left[\theta|z\right] - \mathrm{E}_S\left[\theta|z\right]\right] = \sum \mathrm{Pr}_R\left[z_{ij}\right] q_i^R(z_{ij}) q_j^R(z_{ij}) \Delta_{(i,j)}. \tag{35}$$

Letting $\alpha_{ij}^i = \mathrm{Pr}\left[z_{ij}|\theta_i\right] \mathrm{Pr}_S\left[\theta_i\right]$, and denoting by $H(p,q)$ the harmonic mean of p and q, so that $H(p,q) = \frac{2pq}{p+q}$, we can write (35) as

$$\mathrm{E}_S^{\pi}\left[\mathrm{E}_R\left[\theta|z\right] - \mathrm{E}_S\left[\theta|z\right]\right] = \frac{1}{2} \sum H(\alpha_{ij}^i, \alpha_{ij}^j) \Delta_{(i,j)}. \tag{36}$$

As previously noted, an experiment that maximizes (35) also maximizes $\mathrm{E}_S^{\pi}\left[\mathrm{E}_R\left[\theta|z\right]\right]$. Therefore, an optimal experiment under **(A1)** and **(A2′)** also solves the following program:

$$\max \sum H(\alpha_{ij}^i, \alpha_{ij}^j) \Delta_{(i,j)}, s.t. \alpha_{ij}^i, \alpha_{ij}^j \geq 0, \sum_{\theta_k \in \Theta} \alpha_{ik}^i = p_{\theta_i}^R. \tag{37}$$

Consider a fixed state θ_i. We now investigate which realizations will be induced by θ_i. First, if $\alpha^i_{ij}, \alpha^j_{ij} > 0$, we must have $\Delta_{(i,j)} > 0$, as the sender could otherwise improve by having the experiment fully reveal θ_i and θ_j if z_{ij} is realized. Second, as

$$\frac{\partial H(\alpha^i_{ij}, \alpha^j_{ij})}{\partial \alpha^i_{ij}} = \left(\frac{\alpha^j_{ij}}{\alpha^i_{ij} + \alpha^j_{ij}}\right)^2 \leq 1,$$

the marginal return to increasing α^i_{ij} in $H(\alpha^i_{ij}, \alpha^j_{ij})$ is largest when $\alpha^i_{ij} = 0$, in which case it equals 1. Now suppose that under an optimal experiment, we have that $\alpha^i_{ij} > 0$ and $\alpha^i_{ik} = 0$. Then, we must have that $\Delta_{(i,j)} \geq \Delta_{(i,k)}$. Otherwise, if $\Delta_{(i,j)} < \Delta_{(i,k)}$, marginally increasing α^i_{ik} while reducing α^i_{ij} would generate a gain to the sender

$$\frac{\partial H(\alpha^i_{ij}, \alpha^k_{ik})}{\partial \alpha^i_{ik}} \Delta_{(i,k)} - \frac{\partial H(\alpha^i_{ij}, \alpha^j_{ij})}{\partial \alpha^i_{ij}} \Delta_{(i,j)} = \Delta_{(i,k)} - \frac{\partial H(\alpha^i_{ij}, \alpha^j_{ij})}{\partial \alpha^i_{ij}} \Delta_{(i,j)} > \Delta_{(i,k)} - \Delta_{(i,j)} \geq 0.$$

To prove the last claim, suppose by way of contradiction that $\Delta_{(i,j)} \leq \min\{\Delta_{(i,k)}, \Delta_{(k,j)}\}$ and yet $\Pr_S[z_{i,j}] > 0$. First, this requires $\Delta_{(i,j)} \geq 0$. Second, applying the first part of Proposition 8 implies that $\Delta_{(i,j)} \geq \xi_i$, and since $\Delta_{(i,k)} \geq \Delta_{(i,j)}$, we must have $\Pr_S[z_{i,k}] > 0$. Similarly, $\Delta_{(k,j)} \geq \Delta_{(i,j)} \geq \xi_j$ implies that $\Pr_S[z_{k,j}] > 0$. Finally, the fact that all elements $\Delta_{(i,j)}, \Delta_{(i,k)}$,and $\Delta_{(k,j)}$ are positive implies that r^S decreases for a higher state — i.e., for $\theta_j > \theta_i$, we must have $r^S_j < r^S_i$.

Suppose, wlog, that the three states are ordered $\theta_i < \theta_j < \theta_k$. Since (7) can be rewritten as $\lambda^S_z = \langle q^R(z), r^S \rangle$, $\Pr_S[z_{i,j}], \Pr_S[z_{j,k}] > 0$ implies $r^S_i > \lambda^S_{z_{ij}} > r^S_j > \lambda^S_{z_{jk}} > r^S_k$. Therefore, $a_{z_{ij}} < a_{z_{jk}}$, but $\lambda^S_{z_{ij}} > \lambda^S_{z_{jk}}$, which violates the conclusion of Proposition 7, and, thus, this experiment cannot be optimal. ■

B.2 Private Priors

Consider the extended model with private priors described in Section 5. As an application of (24), consider the pure persuasion model from Section 4.3. When the sender knows the receiver's prior, Proposition 5(i) provides conditions on the likelihood ratio of priors such that persuasion is valuable. Suppose that these conditions are met and the sender strictly benefits from providing experiment π to a particular receiver. By a continuity argument, the same π strictly benefits the sender when she faces another receiver whose beliefs are not too different. Consequently, even if the sender does not know the receiver's prior, persuasion remains beneficial when the receiver's possible priors are not too dispersed. Proposition B.1 provides an upper bound on how dispersed these beliefs can be. To this end, let R be the set of likelihood ratios induced by the priors in the support of $h(p^R|p^S)$,

$$R = \left\{ r^R : \{r_\theta^R = p_\theta^R / p_\theta^S\}_{\theta\in\Theta},\ p^R \in Supp(h(p^R|p^S)) \right\}. \tag{38}$$

Proposition B.1 *Suppose that r^R and $r^R\theta$ are not collinear w.r.t. W for all $r^R \in R$, and let $m = \frac{1}{2}\frac{\max\left|u_S''(a)\right|}{\min u_S'(a)} > 0$. If for all $r^R, r^{R'} \in R$*

$$\left\| r^R - r^{R'} \right\| \leq \beta, \tag{39}$$

with β given by (47), then the sender benefits from persuasion.

The condition on r^R and $r^R\theta$ implies that if the sender knew the receiver's prior, then she could find an experiment with a positive value (cf. Proposition 5). The bound β is defined below by (47), as a function of the curvature of u_S. From (39), β represents a lower bound on the cosine of the angle between any two likelihood ratios in the support of $h(p^R|p^S)$. Therefore, (39) describes how different the receiver's possible prior beliefs can be for the sender still to benefit from persuasion, by imposing an upper bound on the angle between any two likelihood ratios in R.

Proof: The proof of this Proposition will make use of the following lemma:

Lemma B.1 *Let R be defined by* (38) *and $m = \frac{1}{2}\frac{\max\left|u_S''(a)\right|}{\min u_S'(a)} > 0$, and for each $r^R \in R$, define $\Delta_S = \frac{\left\langle q^S, r^R\theta\right\rangle}{\left\langle q^S, r^R\right\rangle} - \left\langle p^R, \theta\right\rangle$, and define $l_{r^R}(\varepsilon)$ as*

$$l_{r^R}(\varepsilon) = \frac{\left\langle \varepsilon, r^R\right\rangle}{\Delta_S}. \tag{40}$$

For any ε and $r^R \in R$ such that

$$l_{r^R}(\varepsilon) < -m \text{ and } \Delta_S > 0, \text{ with } p^S + \varepsilon \in \Delta(\Theta), \tag{41}$$

there exists an experiment π with the following properties: (i) Some realization of π induces in the sender the belief $p^S + \varepsilon$; and (ii) π increases the expected utility of the sender when the receiver's associated likelihood ratio is r^R.

Proof: The function $l_{r^R}(\varepsilon)$ has an immediate interpretation as a measure of disagreement: the numerator $\langle \varepsilon, r^R \rangle$ is the difference in the probability that the receiver and sender attach to a realization inducing a posterior $q_S = p_S + \varepsilon$ on the sender, divided by the probability that the sender ascribes to such realization, while the denominator is the change in the receiver's action when the sender changes her belief to q_S. We first show that if some ε satisfies (41), then the value of information control is positive. Consider V_S defined in (11), which in this case can be written as

$$V_S(q^S) = u_S\left(\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}\right),$$

with gradient at p^S

$$\nabla V_S(p^S) = u_S'(\langle p^R, \theta \rangle)\left(r^R\theta - \langle p^R, \theta \rangle r^R\right).$$

By Corollary 1, the value of information control is positive if and only if there exists ε, with $p^S + \varepsilon \in \Delta(\Theta)$, such that

$$\langle \nabla V_S(p^S), \varepsilon \rangle < V_S(p^S + \varepsilon) - V_S(p^S). \tag{42}$$

We now show that an ε satisfying (41) also satisfies (42). Since

$$u_S\left(\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}\right) - u_S(\langle p^R, \theta \rangle) - u_S'(\langle p^R, \theta \rangle)\left(\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle} - \langle p^R, \theta \rangle\right) = \int_{\langle p^R,\theta \rangle}^{\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R,\theta \rangle}^{t} u_S''(\tau) d\tau dt,$$

we can rewrite (42) as

$$u_S'(\langle p^R, \theta \rangle)\langle \varepsilon, r^R \rangle \Delta_S < \int_{\langle p^R,\theta \rangle}^{\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R,\theta \rangle}^{t} u_S''(\tau) d\tau dt.$$

By the mean value theorem, we have

$$\int_{\langle p^R,\theta \rangle}^{\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R,\theta \rangle}^{t} u_S''(\tau) d\tau dt \geq -\max|u_S''(a)| \int_{\langle p^R,\theta \rangle}^{\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R,\theta \rangle}^{t} d\tau dt = -\frac{1}{2}\max|u_S''(a)|\,\Delta_S^2.$$

Moreover, if ε satisfies (41), then it also satisfies

$$\langle \varepsilon, r^R \rangle \min u_S'(a) < -\frac{1}{2}\max|u_S''(a)|\,\Delta_S,$$

implying that ε also satisfies (42) since

$$u_S'(\langle p^R, \theta \rangle)\langle \varepsilon, r^R \rangle \Delta_S < \langle \varepsilon, r^R \rangle \Delta_S \min u_S'(a) < -\frac{1}{2}\max|u_S''(a)|\,\Delta_S^2 \leq \int_{\langle p^R,\theta \rangle}^{\frac{\langle q^S, r^R\theta \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R,\theta \rangle}^{t} u_S''(\tau) d\tau dt.$$

For each ε satisfying (41), we now construct an experiment that improves the sender's expected utility and that has a realization that induces belief $p^S+\varepsilon$ in the sender. Let υ be the excess of the right-hand side over the left-hand side in (42),

$$\upsilon = V_S(p^S+\varepsilon) - V_S(p^S) - \left\langle \nabla V_S(p^S), \varepsilon \right\rangle > 0. \tag{43}$$

Consider the experiment $\pi(\varepsilon,\delta)$ with $Z = \{\varepsilon^+, \varepsilon^-\}$, such that $\Pr_S[z=\varepsilon^+] = \delta$, and if $z=\varepsilon^+$, then the sender's posterior is $p^S+\varepsilon$. A taylor series expansion of $V_S(q^S)$ yields

$$V_S(q^S) = V_S(p^S) + \left\langle \nabla V_S(p^S), q^S - p^S \right\rangle + L\left(q^S-p^S\right), \text{ with } \lim_{t\to 0} \frac{L\left(t\left(q^S-p^S\right)\right)}{t} = 0. \tag{44}$$

Then, the sender's gain from $\pi(\varepsilon,\delta)$ is

$$\begin{aligned}
\Delta_{\pi(\varepsilon,\delta)} &= \delta\left(V_S(p^S+\varepsilon) - V_S(p^S)\right) + (1-\delta)\left(V_S(p^S - \frac{\delta}{1-\delta}\varepsilon) - V_S(p^S)\right) \\
&= \delta\left(\upsilon + \left\langle \nabla V_S(p^S), \varepsilon\right\rangle\right) - \delta\left\langle \nabla V_S(p^S), \varepsilon\right\rangle + L\left(-\frac{\delta}{1-\delta}\varepsilon\right) \\
&= \delta\left(\upsilon - (1-\delta)\frac{L\left(-\delta\varepsilon/(1-\delta)\right)}{(-\delta/(1-\delta))}\right).
\end{aligned}$$

The convergence to zero of the second term in the parentheses when δ tends to zero and $\upsilon>0$ guarantees the existence of $\delta>0$ such that $\Delta_{\pi(\varepsilon,\delta)} > 0$. ■

Proof of Proposition B.1: First, we introduce additional notation. With $l_{r^R}(\varepsilon)$ defined as in (40), define the sets $M(r^R)$ by

$$M(r^R) = \left\{\varepsilon : l_{r^R}(\varepsilon) < -m,\ \Delta_S > 0, p^S+\varepsilon \in \Delta(\Theta)\right\}.$$

Note that r^S and θ are negatively collinear if and only if r^R and $r^R\theta$ are positively collinear. That is, the condition on Proposition 5 could instead be stated in terms of collinearity of r^R and $r^R\theta$. Moreover, if r^R and $r^R\theta$ are not collinear, then the restriction of $l_{r^R}(\varepsilon)$ to $\{\varepsilon : \langle\varepsilon, 1\rangle = 0\}$ is surjective, and, thus, the set $M(r^R)$ is non-empty.

Define the function

$$\Psi\left(\varepsilon, r^R\right) = \left\langle \varepsilon, r^R - mf^R\right\rangle + \left(\left\langle \varepsilon, r^R\right\rangle\right)^2, \text{ with } f^R = r^R\theta - \left\langle p^S, r^R\theta\right\rangle,$$

which characterizes $M(r^R)$ since for ε such that $p^S+\varepsilon \in \Delta(\Theta)$, $\Psi\left(\varepsilon, r^R\right) \le 0$ and $\left\langle \varepsilon, f^R\right\rangle \ge 0$ if and only if $\varepsilon \in M(r^R)$. Finally, let

$$\begin{aligned}
\gamma &= 2\left(1 + m\left(\max|\theta| + \|\theta\|\right) + (4 + m\|\theta\|)\sup_{r^R\in R}\left\|r^R\right\|\right), \qquad (45) \\
Z &= \min_{\varepsilon\in\{\varepsilon: p^S+\varepsilon\in\Delta(\Theta)\}, r^R\in R} \Psi\left(\varepsilon, r^R\right) \text{ s.t. } \left\langle \varepsilon, r^R\left(\theta - \left\langle p^S, r^R\theta\right\rangle\right)\right\rangle \le 0, r^R \in R. \qquad (46)
\end{aligned}$$

Under the conditions of Proposition B.1, $Z < 0$. Finally, define β in (39) as

$$\beta = \frac{|Z|}{\gamma}. \tag{47}$$

Our proof is structured in two steps that show (i) if $\cap_{r^R \in R} M(r^R)$ is non-empty, then following Lemma B.1 allows us to design an experiment π that increases the sender's expected utility for every receiver's belief in the support of $h(p^R|p^S)$; and (ii) under the conditions of Proposition B.1, $\cap_{r^R \in R} M(r^R) \neq \varnothing$.

Step (i) - Suppose that $\varepsilon \in \cap_{r^R \in R} M(r^R)$. Consider υ as defined by (43). As υ is a continuous function of r^R in the compact set R, it achieves a minimum $\underline{\upsilon}= \min_{r^R \in R} \upsilon > 0$. Then, define $\underline{\delta}$ as

$$\underline{\delta} = \min \left\{ \delta : \underline{\upsilon} + \frac{L\left(-\frac{\delta}{1-\delta}\varepsilon\right)}{\delta} \geq 0 \right\},$$

with the function L given by (44). Now, define the experiment $\pi(\varepsilon, \delta')$ as in the proof of Lemma B.1— i.e., $Z = \{\varepsilon^+, \varepsilon^-\}, q^S(\varepsilon^+) = p^S + \varepsilon$ and $\Pr_S[z = \varepsilon^+] = \delta'$, and set $\delta' = \underline{\delta}$. Then, the sender's gain from $\pi(\varepsilon, \delta')$ is positive for any receiver's prior in $Supp(h(p^R|p^S))$.

Step (ii) - Fix $p^{R'}$ with associated likelihood ratio $r^{R'} \in R$. For any $r^R \in R$ with $\eta = r^R - r^{R'}$, we have

$$\Psi\left(\varepsilon, r^R\right) - \Psi\left(\varepsilon, r^{R'}\right) = \left(1 + m\left\langle p^S, r^{R'}\theta\right\rangle + \left\langle \varepsilon, r^R + r^{R'}\right\rangle\right)\langle \varepsilon, \eta\rangle - m\langle \varepsilon, \eta\theta\rangle + m\left\langle p^S, \eta\theta\right\rangle\langle \varepsilon, r\rangle.$$

The following bounds make use of the Cauchy-Schwartz inequality (in particular, the implication that $|\langle \varepsilon, \eta\theta\rangle| \leq \|\varepsilon\| \|\eta\| \|\theta\|$—see Steele, 2004)[28] and the fact that $\left\|p^S\right\| \leq 1$ and $\|\varepsilon\| = \left\|q^S - p^S\right\| \leq 2$,

$$\begin{aligned}
\left|1 + m\left\langle p^S, r^{R'}\theta\right\rangle + \left\langle \varepsilon, r^R + r^{R'}\right\rangle\right| &\leq 1 + m \max\theta + 4 \sup_{r^R \in R} \left\|r^R\right\|, \\
|m\langle \varepsilon, \eta\theta\rangle| &\leq m\|\varepsilon\| \|\eta\| \|\theta\| \leq 2m\|\eta\| \|\theta\|, \\
\left|m\left\langle p^S, \eta\theta\right\rangle\langle \varepsilon, r\rangle\right| &\leq 2m\|\eta\| \|\theta\| \sup_{r^R \in R} \left\|r^R\right\|.
\end{aligned}$$

[28]Steele, J. M. (2004) "The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities," *Mathematical Association of America.*

From these bounds, we then obtain the following estimate

$$\begin{aligned}\left|\Psi\left(\varepsilon, r^{R}\right)-\Psi\left(\varepsilon, r^{R'}\right)\right| &\leq \left|1+m\left\langle p^{S}, r^{R'}\theta\right\rangle+\left\langle\varepsilon, r^{R}+r^{R'}\right\rangle\right|\|\varepsilon\|\|\eta\| \\ &\quad +|m\langle\varepsilon, \eta\theta\rangle|+\left|m\left\langle p^{S}, \eta\theta\right\rangle\langle\varepsilon, r\rangle\right| \\ &\leq 2\left(1+m \max\theta+4 \sup_{r^{R} \in R}\left\|r^{R}\right\|\right)\|\eta\|+2m\|\theta\|\|\eta\| \\ &\quad +2m\|\theta\| \sup_{r^{R} \in R}\left\|r^{R}\right\|\|\eta\| \\ &= \gamma\|\eta\|,\end{aligned}$$

where γ is defined by (45). Selecting ε' an $r^{R'}$ that solve the program (46) and noting that $Z<0$, we have that for any $r^{R} \in R$,

$$\Psi\left(\varepsilon', r^{R}\right)=\Psi\left(\varepsilon', r^{R'}\right)+\Psi\left(\varepsilon', r^{R}\right)-\Psi\left(\varepsilon', r^{R'}\right) \leq Z+\gamma\|\eta\| \leq Z+|Z|=0.$$

This implies that $\varepsilon' \in M(r^{R})$ for all $r^{R} \in R$. ■