

# Game Theory Problem Set 3

a

	L	R
T	$\frac{3}{4}\epsilon_2$ 0	1
B	0 $\frac{3}{4}\epsilon_1$	0

Best responses underlined

By inspection, (T,L) and (B,R) are the only pure NE where players play mutual best responses.

Suppose there is an NE  $\sigma^*$  st P1.1 mixes T and B, then by definition of NE P1.1 has no profitable deviation by reallocating probability mass from T to B, between T and B, so  $\pi_1(T, \sigma_2^*) = \pi_1(B, \sigma_2^*)$ . Let  $q$  denote the probability assigned to L by  $\sigma_2^*$ .

$$q = (1-q)(3+\epsilon_1), (4+\epsilon_1)q = 1, q = \frac{1}{4+\epsilon_1}, q \neq 0, q \neq 1$$

so if P1.1 mixes, P1.2 also mixes. By symmetry, if P1.2 mixes, P1.1 also mixes. So there are no hybrid NE.

P1.2 mixes, then by definition of NE, P1.2 has no profitable deviation, so  $\pi_2(L, \sigma_1^*) = \pi_2(R, \sigma_1^*)$ . Let  $p$  denote the probability  $\sigma_1^*$  assigns to T.

$$p(3+\epsilon_2) = (1-p), p = \frac{1}{4+\epsilon_2}$$

so the only mixed NE is  $(\frac{1}{4+\epsilon_2}T + \frac{3\epsilon_2}{4+\epsilon_2}B, \frac{1}{4+\epsilon_1}L + \frac{3\epsilon_1}{4+\epsilon_1}R)$

$$(\frac{1}{4+\epsilon_2}T + (1-p)B, \frac{1}{4+\epsilon_1}L + (1-q)R) \text{ where } p = \frac{1}{4+\epsilon_2} \text{ and } q = \frac{1}{4+\epsilon_1}$$

$$\text{When } \epsilon_1 = \epsilon_2 = 0, p = \frac{1}{4}, q = \frac{1}{4}, \pi_1(\sigma^*) = \frac{1}{4} \cdot \frac{3}{4} \cdot 1 + \frac{3}{4} \cdot \frac{1}{4} \cdot 0$$

$$= \frac{3}{16}, \pi_2(\sigma^*) = \frac{1}{4} \cdot \frac{3}{4} \cdot 0 + \frac{3}{4} \cdot \frac{1}{4} \cdot 1 = \frac{3}{16}$$

Each player has lower expected payoff at the mixed NE than at either pure NE because of the non-zero ( $\frac{1}{16}$ ) probability of miscoordination such that each player has zero payoff.

$$q = P(\hat{\epsilon}_2 > \frac{1}{2}) = 1 - P(\hat{\epsilon}_2 < \frac{1}{2}) = 1 - \frac{1}{2 + \sqrt{4 + \epsilon}}$$

$$\lim_{\epsilon \rightarrow 0} q = \frac{3}{4}$$

The ex ante probabilities of each action of each player converge to the probabilities of each action of each player in the strategic form game where  $\epsilon_1 = \epsilon_2 = 0$ , i.e. the unperturbed game. This is the result of the Harsanyi purification theorem.

Mixed NE are unsatisfactory because at the mixed NE, each player has no strict incentive to mix as the mixed NE prescribes since the payoff of each action the mixed NE prescribes this player mix over yields equal payoff, so any probability distribution over these actions yields equal payoffs

According to Harsanyi's purification theorem, the probability distributions induced by the pure NE of the perturbed game where each player plays a threshold strategy based on some private payoff shock converges to the probability distribution of the mixed NE in the unperturbed game as the perturbation becomes vanishingly small.

Harsanyi purification can be interpreted as supposing that there is some small private fact that affects each player's preferences that is unmodeled. Then, if each player in fact plays a threshold strategy that he has strict incentive to play based on this fact, it appears to all other players and observers that each player mixes. Because any game is a model of some real strategic interaction, and any model simplifies reality, it is reasonable to suppose that there are small private unmodeled shocks.

2a Players:  $N = \{St, G\}$

Actions:  $A_{St} = \{L, R\}, A_G = \{l, r\}$

States:  $\Omega = \{SHL, SHR\}$

Signals:  $t_{St}(\omega) = \omega$  for  $\omega \in \Omega$ ,  $t_G(\omega) = 0$  for  $\omega \in \Omega$

Beliefs:  $P_{St}(\omega = SHL | t_{St}) = 1$  iff  $t_{St} = SHL$ ,  $P_{St}(\omega = SHR | t_{St}) = 1$  iff  $t_{St} = SHR$ ,  $P_G(\omega = SHL) = P_G(\omega = SHR) = \frac{1}{2}$

Since St is either type L or R, St's pure strategies are  $s_{St} \in \{LL, LR, RL, RR\}$ , where the first action is St's action if he is type L and the second is his action if he is type R. G has only one type, so G's pure strategies are  $s_G \in A_G$ .

b Let  $s^*$  be the pure BNE st P1.1 plays T iff  $\epsilon_1 > \hat{\epsilon}_1$  and P1.2 plays L iff  $\epsilon_2 > \hat{\epsilon}_2$ .

$$\pi_1(T, s_2^*) = \frac{3}{4} \cdot \frac{\epsilon_2}{\epsilon_2} (1 - \frac{\hat{\epsilon}_2}{\epsilon_2}) + \frac{\hat{\epsilon}_2}{\epsilon_2} (3 + \epsilon_1)$$

$$\pi_1(B, s_2^*) = (1 - \frac{\hat{\epsilon}_2}{\epsilon_2})$$

$$\pi_1(T, s_2^*) \geq \pi_1(B, s_2^*) \iff$$

$$1 - \frac{\hat{\epsilon}_2}{\epsilon_2} \geq \frac{\hat{\epsilon}_2}{\epsilon_2} (3 + \epsilon_1)$$

$$1 - \frac{\hat{\epsilon}_2}{\epsilon_2} \geq \frac{\hat{\epsilon}_2}{\epsilon_2} (3 + \epsilon_1)$$

$$1 - \frac{\hat{\epsilon}_2}{\epsilon_2} \geq \frac{\hat{\epsilon}_2}{\epsilon_2} (3 + \epsilon_1)$$

$$\epsilon_1 \leq \frac{\epsilon_2}{\hat{\epsilon}_2} - 4$$

By definition of BNE, P1.1 plays at  $s^*$ , P1.1 plays T iff  $\pi_1(T, s_2^*) \geq \pi_1(B, s_2^*)$ , so  $\hat{\epsilon}_1 = \frac{\epsilon_2}{\epsilon_2} - 4$ .

By symmetry,  $\hat{\epsilon}_2 = \frac{\epsilon_1}{\epsilon_1} - 4$

$$\text{Then, } \hat{\epsilon}_1 \hat{\epsilon}_2 = \frac{\epsilon_2}{\epsilon_2} - 4 \cdot \frac{\epsilon_1}{\epsilon_1} - 4, \hat{\epsilon}_1 \hat{\epsilon}_2 = \frac{\epsilon_2}{\epsilon_2} - 4 \cdot \frac{\epsilon_1}{\epsilon_1} - 4, \text{ so } \hat{\epsilon}_1 = \frac{\epsilon_2}{\epsilon_2} - 4$$

$$\hat{\epsilon}_1 = \frac{\epsilon_2}{\epsilon_2} - 4, \hat{\epsilon}_1^2 + 4\hat{\epsilon}_1 - \epsilon_2 = 0, \hat{\epsilon}_1 = \frac{-4 \pm \sqrt{16 - 4\epsilon_2}}{2}$$

$$= -2 \pm \sqrt{4 - \epsilon_2} = -2 + \sqrt{4 - \epsilon_2} \text{ since } \hat{\epsilon}_1 \in [0, \frac{\epsilon_2}{\epsilon_2}]$$

$$\text{so } \hat{\epsilon}_1 = \hat{\epsilon}_2 = -2 + \sqrt{4 - \epsilon}$$

c Let  $p$  and  $q$  denote the probability that P1.1 plays T and the probability that P1.2 plays L respectively.

$$p = P(\epsilon_1 < \hat{\epsilon}_1) = \frac{\hat{\epsilon}_1}{\epsilon_1} = \frac{-2 + \sqrt{4 - \epsilon}}{\epsilon} = \frac{1}{2 + \sqrt{4 - \epsilon}}$$

$$\lim_{\epsilon \rightarrow 0} p = \frac{1}{4}$$



ex ante payoffs

	l	r
LL	0.35 0.1	0.65 0.9
LR	0.15 0.15	0.85 0.85
RL	0.3 0.3	0.7 0.7
RR	0.1 0.35	0.9 0.65

Best responses underlined

By inspection, there are no ~~BNE~~ pure BNE on ex ante payoffs, where players play mutual best responses. Pure BNE on interim payoffs coincide with pure BNE on ex ante payoffs, so there are no pure BNE on interim payoffs.

b Suppose that let  $\sigma^*$  denote a hybrid BNE where St plays pure strategy  $s_{St}^* \in \{LL, LR, RL, RR\}$  and G plays mixed strategy  $\sigma_G^*$  which is some probability distribution (non-degenerate) probability distribution over ~~l and r~~ Ag.

Suppose that  $s_{St}^* = LL$ , then  $\pi_G(l, \sigma^*) = 0.35 > \pi_G(r, \sigma^*) = 0.1$ , so G has incentive to deviate from  $\sigma_G^*$  to the ~~the~~ pure strategy l, by reductio, and  $\sigma^*$  is not a BNE. By reductio,  $s_{St}^* \neq LL$ .

Similarly, if  $s_{St}^* = RR$ , then  $\pi_G(r, \sigma^*) = 0.35 > \pi_G(l, \sigma^*) = 0.1$ , so  $\sigma^*$  fails to deviation  $(s_{St}^*, r)$ . By reductio,  $s_{St}^* \neq RR$ .

Suppose that  $s_{St}^* = RL$ , then  $\pi_{St}(s_{St}^*, \sigma_G^*) = 0.7 > \pi_{St}(LR, \sigma_G^*) = 0.15$ , so  $\sigma^*$  fails to deviation  $(LR, \sigma_G^*)$ . By reductio,  $s_{St}^* \neq RL$ .

so  $s_{St}^* = LR$ , then  $\pi_G(l, s_{St}^*) = 0.15 = \pi_G(r, s_{St}^*) = 0.15$ , so G has no profitable deviation. Let q denote the probability that  $\sigma_G^*$  assigns to l. St has no profitable deviation if  $\pi_{St}(s_{St}^*, \sigma_G^*) = 0.85 \geq \pi_{St}(LL, \sigma_G^*)$  and  $\geq \pi_{St}(RR, \sigma_G^*)$ . Since LR strictly dominates RL ex ante.  $0.85 \geq 0.65q + 0.9(1-q) = 0.9 - 0.25q$ ,  $0.25q \geq 0.05$ ,  $q \geq 0.2$ .  $0.85 \geq 0.9q + 0.65(1-q) = 0.65 + 0.25q$ ,  $0.25q \leq 0.20$ ,  $q \leq 0.8$ . G has no profitable deviation since  $\pi_G(l, LR) = \pi_G(r, LR)$ . So any  $\sigma^* = (s_{St}^*, \sigma_G^*)$  where  $s_{St}^* = LR$  and  $\sigma_G^* = (ql + (1-q)r)$  for  $q \in (0.2, 0.8)$  is a hybrid BNE.

c Let  $\sigma^*$  denote a mixed BNE where St plays mixed strategy  $\sigma_{St}^* = (\sigma_{StL}^*, \sigma_{StR}^*) = ((p_L L + (1-p_L)R), (p_R L + (1-p_R)R))$ , i.e. plays L with  $p_L$  type L and with  $p_R$  type R, and G plays  $\sigma_G^* = (ql + (1-q)r)$ , where  $q \in [0, 1]$ .

Suppose  $\sigma_G^* = l$ , then by inspection of the payoff tables,  $\pi_{St}(RR, \sigma_G^* = l) > \pi_{St}(\sigma_{St}^*, \sigma_G^* = l)$  for any  $p_L, p_R \in (0, 1)$ . So  $\sigma^*$  fails to deviation ~~to~~  $(\sigma_{St}^*, \sigma_G^* = l)$ , and  $\sigma^*$  is not a BNE. By reductio,  $\sigma_G^* \neq l$ .

By symmetry,  $\sigma_G^* \neq r$ . so if St mixes at any BNE G mixes l and r. Intuitively this is because if G plays a pure strategy, St has a pure best response. At any BNE that St mixes, G also mixes. Intuitively, this is because if G plays a pure strategy, St has a best response in pure strategy strict best response.

$$\pi_{St}(\sigma^* | \sigma_G^* = l) = \cancel{p_L} 0.7p_L + p_R(1-p_L) + 0.8(1-p_L)q + 0.6(1-p_L)(1-q)$$

$$\pi_{St}(L, \sigma_G^*) = 0.7q + 1(1-q) = 1 - 0.3q$$

$$\pi_{St}(R, \sigma_G^*) = 0.8q + 0.6(1-q) = 0.6 + 0.2q$$

type L St has no profitable deviation from  $\sigma_{StL}^*$  if

$$\pi_{St}(L, \sigma_G^*) = \pi_{St}(R, \sigma_G^*), \quad 1 - 0.3q = 0.6 + 0.2q, \quad q = 0.8$$

$$\pi_{StR}(L, \sigma_G^*) = 0.6q + 0.8(1-q) = 0.8 - 0.2q$$

$$\pi_{StR}(R, \sigma_G^*) = q + 0.7(1-q) = 0.7 + 0.3q$$

type R St has no profitable deviation from  $\sigma_{StR}^*$  if

$$\pi_{StR}(L, \sigma_G^*) = \pi_{StR}(R, \sigma_G^*), \quad 0.8 - 0.2q = 0.7 + 0.3q, \quad q = 0.2$$

$$\pi_G(l, \sigma_{St}^*) = \frac{1}{2}(0.7p_L + 0.8(1-p_L)) + \frac{1}{2}(0.6p_R + 1(1-p_R)) = 0.4 - 0.05p_L + 0.5 - 0.05p_R = 0.2$$

$$\pi_G(r, \sigma_{St}^*) = \frac{1}{2}(\dots)$$

$$\pi_G(l, \sigma_{St}^*) = \frac{1}{2}(0.3p_L + 0.2(1-p_L)) + \frac{1}{2}(0.4p_R + 0(1-p_R)) = 0.1 + 0.05p_L + 0.2p_R$$

$$\pi_G(r, \sigma_{St}^*) = \frac{1}{2}(0.8p_L + 0.4(1-p_L)) + \frac{1}{2}(0.2p_R + 0.3(1-p_R)) = 0.2 - 0.2p_L + 0.15 - 0.05p_R$$

G has no profitable deviation from  $\sigma_G^*$  if

$$\pi_G(l, \sigma_{St}^*) = \pi_G(r, \sigma_{St}^*), \quad 0.1 + 0.05p_L + 0.2p_R = 0.2 - 0.2p_L + 0.15 - 0.05p_R$$

From above, ~~not both~~ type L St mixes <sup>and</sup> type R St mixes ~~but not both~~. From  $p_L + p_R = 1$ , if type L St mixes, so does type R St. By reductio, there is no BNE such that ~~one type~~ at least one type of St mixes.