Analytic Theory II:
Static Games with Incomplete Information

Branislav L. Slantchev
Department of Political Science, University of California – San Diego

February 27, 2020

Contents

The games we have discussed up to this point assume *common knowledge* about the structure of the game, the possible moves (including the probability distributions of chance events), the players, and their preferences. In fact, we assumed that the common knowledge about these things is itself common knowledge. These games of *complete information* can be viewed as rough approximations for certain situations where perhaps the level of uncertainty is low and might be considered irrelevant. Generally, however, players may not possess full information about their opponents, the situation, or how actions map into outcomes. They might have some idea, a **belief**, about the relevant factors but not know for sure. We have seen a version of this with games of imperfect information where players could not observe the *actual* actions of others but did know what the *possible* moves could be, and we then required that players form beliefs about the actions of their opponent that are *consistent* with the best-response definition of rationality we have been using. In these situations, the players are best-responding to their own beliefs about what others might be doing, and these beliefs are derived from the assumption that the others are best-responding as well. The definition of Nash equilibrium is, in effect, a requirement about the consistency of beliefs about all the players' strategies. The questions now is: Could this idea be somehow extended to deal with situations where players might not know for sure other things about each other, such as preferences, available actions, or the structure of the game itself?

It turns out that the answer is positive: in the 1960s John C. Harsanyi realized that all kinds of incomplete information could be represented with the abstract idea of **player types**: that is, if Player 1 was unsure about Player 2's preferences, he could imagine that he might be facing one from several potential "types" of Player 2, each of whom have her own specific preferences. Uncertainty about which of these "types" is actually playing the game is then represented by a probability distribution over the types; that is, Player 1's belief about Player 2's preferences boils down to a probability with which he might be interacting with a particular type. This transforms incomplete information into a move by chance at the beginning of the game—which determines which actual game is being played—and so the interaction becomes one of imperfect information, which you already know how to solve. The so-called *Harsanyi transformation* can be applied to any sort of uncertainty: if a player does not know what moves are available to him (or his opponent), then "chance" would be selecting among games with different structures, and so on.

1 A Simple Entry Game

To make matters a bit more specific, let us look at an example. There are two firms in some industry: an incumbent (player 1) and a potential entrant (player 2). Player 1 decides whether to build a plant, and simultaneously player 2 decides whether to enter. Suppose that player 2 is uncertain whether player 1's building cost is $3/2$ or 0, while player 1 knows his own cost. The payoffs are shown in Fig. 1 (p. 2).

	Enter	*Don't*
Build	$0, -1$	$2, 0$
Don't	$2, 1$	$3, 0$

high-cost

	Enter	*Don't*
Build	$3/2, -1$	$7/2, 0$
Don't	$2, 1$	$3, 0$

low-cost

Figure 1: An Entry Game with Incomplete Information.

Player 2's payoff depends on whether player 1 builds or not (but is not directly influenced by player 1's cost). Entering for player 2 is profitable only if player 1 does not build. Note that "don't build" is a dominant strategy for player 1 when his cost is high. However, player 1's optimal strategy when his cost is low depends on his prediction about whether player 2 will enter. Denote the probability that player 2 enters with y. Building is better than not building if

$$(3/2)y + (7/2)(1-y) \geq 2y + 3(1-y)$$
$$y \leq 1/2.$$

In other words, a low-cost player 1 will prefer to build if the probability that player 2 enters is less than $1/2$. Thus, player 1 has to predict player 2's action in order to choose his own action, while player 2, in turn, has to take into account the fact that player 1 will be conditioning his action on these expectations.

For a long time, game theory was stuck because people could not figure out a way to solve such games. However, in a couple of papers in 1967-68, John C. Harsanyi proposed a method that allowed one to transform the game of incomplete information into a game of imperfect information, which could then be analyzed with standard techniques. Briefly, the **Harsanyi transformation** involves introducing a prior move by Nature that determines player 1's "type" (in our example, his cost), transforming player 2's incomplete information about player 1's cost into imperfect information about the move by Nature.

Letting p denote the prior probability of player 1's cost being high, Fig. 2 (p. 3) depicts the Harsanyi transformation of the original game into one of imperfect information.

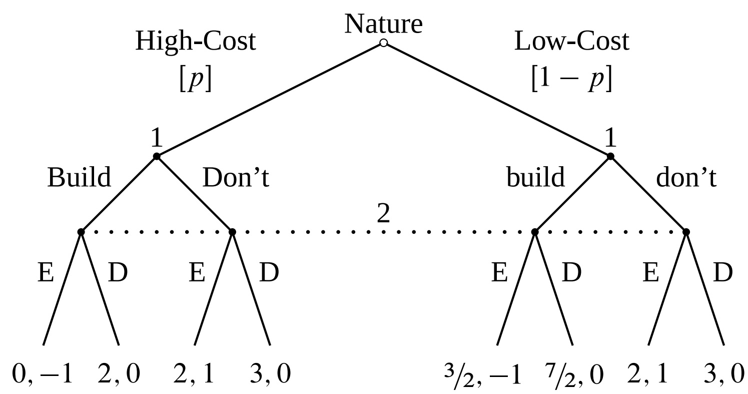

Figure 2: The Harsanyi-Transformed Game from Fig. 1 (p. 2).

Nature (or Chance) moves first and "chooses" player 1's type: with probability p the type is "high-cost" and with probability $1-p$ the type is "low-cost." It is standard to assume that both players have the same prior beliefs about the probability distribution on nature's moves. Player 1 knows his own type (i.e. he learns what the move by Nature is) but player 2 does not. Observe now that after player 1 learns his type, he has private information: all player 2 knows is that probability of him being of one type or another. It is quite important to note that here player 2's *beliefs are common knowledge*. That is, player 1 knows what she believes his type to be, and she knows that he knows, and so on. This is important because player 1 will be optimizing given what he thinks player 2 will do, and her behavior depends on these beliefs. We can now apply the Nash equilibrium solution concept to this new game. Harsanyi's **Bayesian Nash Equilibrium** (or simply Bayesian Equilibrium) is precisely the Nash equilibrium of this imperfect-information representation of the game.

Before defining all these things formally, let's solve the game in Fig. 2 (p. 3). Player 2 has one (big) information set, so her strategy will only have one component: what to do at this information set. Note now that player 1 has two information sets, so his strategy must specify what to do if his type is high-cost and what to do if his type is low-cost. One might wonder why player 1's strategy has to specify what to do in both cases, after all, once player 1 learns his type, he does not care what he would have done if he is of another type.

The reason the strategy has to specify actions for both types is roughly analogous for the reason the strategy has to specify a complete plan for action in extensive-form games with complete information: player 1's optimal action depends on what player 2 will do, which in turn depends on what player 1 would have done at information sets even if these are never reached in equilibrium. Here, player 1 knows his cost which is, say, low. So why should he bother formulating a strategy for the (non-existent) case where his cost is high? The answer is that to decide what is optimal for him, he has to predict what player 2 will do. However, player 2 does not know his cost, so she will be optimizing on the basis of her expectations about what a high-cost player 1 would optimally do and what a low-cost player 1 would optimally do. In other words, the strategy of the high-cost player 1 really represents player 2's expectations.

The Bayesian Nash equilibrium will be a *triple* of strategies: one for player 1 of the high-cost type, another for player 1 of the low-cost type, and one for player 2. In equilibrium, no deviation should be profitable.

1.1 Solution: The Strategic Form

Let's write down the strategic form representation of the game in Fig. 2 (p. 3). Player 1's pure strategies are $S_1 = \{Bb, Bd, Db, Dd\}$, where the first component of each pair tells his what to do if he is the high-cost type, and the second component if he is the low-cost type. Player 2 has only two pure strategies, $S_2 = \{E, D\}$. The resulting payoff matrix is shown in Fig. 3 (p. 4).

Player 1 \ Player 2	E	D
Bb	$3/2 - 3p/2, -1$	$7/2 - 3p/2, 0$
Bd	$2 - 2p, 1 - 2p$	$3 - p, 0$
Db	$3/2 + p/2, 2p - 1$	$7/2 - p/2, 0$
Dd	$2, 1$	$3, 0$

Figure 3: The Strategic Form of the Game in Fig. 2 (p. 3).

Db strictly dominates Bb and Dd strictly dominates Bd. This should not be surprising: since the high-cost type has a strictly dominant strategy not to build, no Nash equilibrium would permit him to build, which is why all strategies that involve him doing so just got eliminated. Removing the two strictly dominated strategies reduces the game to the one shown in Fig. 4 (p. 4).

Player 1 \ Player 2	E	D
Db	$3/2 + p/2, 2p - 1$	$7/2 - p/2, 0$
Dd	$2, 1$	$3, 0$

Figure 4: The Reduced Strategic Form of the Game in Fig. 2 (p. 3).

If player 2 chooses E, then player 1's unique best response is to choose Dd regardless of the value of $p < 1$. Hence $\langle Dd, E\rangle$ is a Nash equilibrium for all values of $p \in (0, 1)$.

Note now that E strictly dominates D whenever $2p - 1 > 0 \Rightarrow p > 1/2$, and so player 2 will never mix in equilibrium in this case. Let's then consider the cases when $p \leq 1/2$. We now also have $\langle Db, D\rangle$ as a Nash equilibrium.

Suppose now that player 2 mixes in equilibrium. Since she is willing to randomize,

$$\begin{aligned} U_2(E) &= U_2(D) \\ \sigma_1(Db)(2p-1) + (1-\sigma_1(Db))(1) &= 0 \\ \sigma_1(Db) &= \frac{1}{2(1-p)}. \end{aligned}$$

Since player 1 must be willing to randomize as well, it follows that

$$\begin{aligned} U_1(Db) &= U_1(Dd) \\ \sigma_2(E)(3/2 + p/2) + (1-\sigma_2(E))(7/2 - p/2) &= 2\sigma_2(E) + 3(1-\sigma_2(E)) \\ \sigma_2(E) &= 1/2. \end{aligned}$$

Hence, we have a mixed-strategy Nash equilibrium with $\sigma_1(Db) = \frac{1}{2(1-p)}$, and $\sigma_2(E) = 1/2$ whenever $p \le 1/2$. The upper bound on p follows from the requirement that $\sigma_1(Db) \le 1$. It is not surprising in light of the fact that most well-behaved games either have an odd number of Nash equilibria or infinitely many. Since this game only has two PSNE when $p \le 1/2$, we should expect the MSNE to exist only in these cases as well.[1]

Summarizing the results, we have the following Nash equilibria:

- Neither the high nor low cost types build, and player 2 enters;
- If $p \le 1/2$, there are two types of equilibria:
  - the high-cost type does not build, but the low-cost type does, and player 2 does not enter;
  - the high-cost type does not build, but the low-cost type builds with probability $1/[2(1-p)]$, and player 2 enters with probability $1/2$.

Intuitively, the results make sense. The high-cost type never builds, so deterring player 2's entry can only be done by the low-cost type's threat to build. If player 2 is expected to enter for sure, then even the low-cost type would prefer not to build, which in turn rationalizes her decision to enter with certainty. Deterrence fails with certainty here but at least no plant is being built (it would have been wasteful to do so given her entry). This result is independent of her prior beliefs.

If, on the other hand, her prior belief is pessimistic enough (she believes that the likelihood of the high-cost type is sufficiently low, or $p \le 1/2$), then deterrence becomes possible, and there are two equilibria. In one, the low-cost type builds for sure, and given her pessimistic priors, player 2 is unwilling to run the risk of entry, so she stays out for sure as well. In the other, the low-cost type builds with positive probability, which means that player 2 can no longer be deterred with certainty: the chances of not facing a plant are high enough to make her willing to mix as well, and the uncertainty about her action, in turn, rationalized the mixed strategy for the low-cost type of player 1. Deterrence can fail, and, moreover, it is possible that player 2 ends up entering a market where a plant has been built, giving both players the worst possible payoffs with positive probability. Thus, unlike the PSNE with deterrence failure, the MSNE does involve building a plant that *ex post* turns out to have been a waste.

[1] There is a continuum of MSNE if $p = 1/2$ but this is a knife-edge condition so we shall ignore it. A "knife-edge condition" refers to a requirement that an exogenous variable takes on specific values. Any solution that depends on such a requirement is extremely unstable since even the tiniest change in the value of that variable will wipe it out. Moreover, if we think of these variables as being drawn from continuous distributions, then the probability that they take on any specific value is zero.

1.2 Solution: Best Responses

Noting that the high-cost player 1 never builds, let x denote the probability that the low-cost player 1 builds. As before, let y denote player 2's probability of entry, and observe that the low-cost player 1 strictly prefers building to not building when the expected utility of building exceeds the expected utility of not building:

$$\begin{aligned} U_1(B|L) &\geq U_1(D|L) \\ {}^{3y}\!/_2 + {}^{7(1-y)}\!/_2 &\geq 2y + 3(1-y) \\ y &\leq {}^1\!/_2 \end{aligned}$$

Player 2 prefers to enter when the expected utility of doing so exceeds the expected utility of not entering:

$$\begin{aligned} U_2(E) &\geq U_2(D) \\ pU_2(E|H) + (1-p)U_2(E|L) &\geq pU_2(D|H) + (1-p)U_2(D|L) \\ p(1) + (1-p)(-x+1-x) &\geq 0 \\ 1 - 2x + 2px &\geq 0 \\ x &\leq \frac{1}{2(1-p)} \equiv \overline{x} \end{aligned}$$

We can write the best responses as follows:

$$BR_1(y|L) = \begin{cases} 1 & \text{if } y < {}^1\!/_2 \\ [0,1] & \text{if } y = {}^1\!/_2 \\ 0 & \text{if } y > {}^1\!/_2 \end{cases} \qquad BR_2(x) = \begin{cases} 1 & \text{if } x < \overline{x} \\ [0,1] & \text{if } x = \overline{x} \\ 0 & \text{if } x > \overline{x} \end{cases}$$

$$\text{where} \qquad \overline{x} = \frac{1}{2(1-p)}.$$

Given these best-responses, the search for a Nash equilibrium boils down to finding a pair (x, y), such that x is optimal for player 1 with low cost against player 2 and y is optimal for player 2 against player 1 given beliefs p and player 1's strategy (x for the low cost and "don't build" for the high cost). (We are, technically speaking, looking for a triple because the strategy of the high-cost type must also be included. However, since it is strictly dominant for that type to not build, we do not have to keep writing it down.)

Suppose first that $y > {}^1\!/_2$ in some equilibrium, in which case $x = BR_1(y) = 0 < \overline{x}$. But then $y = BR_2(x) = 1$, and so we have our first PSNE: $\langle 0, 1 \rangle$. This is the equilibrium, in which player 1 does not build irrespective of his type, and player 2 enters with certainty.

Suppose next that $y < {}^1\!/_2$ in some equilibrium, in which case $x = BR_1(y) = 1$. If $1 < \overline{x}$, then $y = BR_2(1) = 1$ contradicts the supposition, so there can be no equilibrium here. So it must be that $1 \geq \overline{x}$, in which case $y = BR_2(x) = 0$, and so we have our second PSNE: $\langle 1, 0 \rangle$. This only exists if $1 \geq \overline{x}$, or $p \leq {}^1\!/_2$. This is the equilibrium, in which player 1 builds with certainty if he is the low-cost type, and player 2 is deterred from entering.

Finally, suppose that $y = \frac{1}{2}$ in some equilibrium, in which case $x = BR_1(y) \in [0,1]$, so the low-cost type can mix. Since player 2 is willing to randomize, it must be that $x = \overline{x}$, which is only a valid probability if $p \leq {}^1\!/_2$. This recovers the MSNE, in which the low-cost type builds with probability $\overline{x}$, and player 2 enters with probability ${}^1\!/_2$.

This yields the full set of equilibria.

1.3 Interim vs. Ex Ante Predictions

Suppose in the two-firm example player 2 also had private information and could be of two types, "aggressive" and "accommodating." If she must predict player 1's type-contingent strategies, she must be concerned with how an aggressive player 2 might think player 1 would play for each of the possible types for player 1 and also how an accommodating player 2 might think player 1 would play for each of his possible types. (Of course, player 1 must also estimate both the aggressive and accommodating type's beliefs about himself in order to predict the distribution of strategies he should expect to face.)

This brings up the following important question: How should the different types of player 2 be viewed? On one hand, they can be viewed as a way of describing different information sets of a *single* player 2 who makes a type-contingent decision at the *ex ante* stage. This is natural in Harsanyi's formulation, which implies that the move by Nature reveals some information known only to player 2 which affects her payoffs. Player 2 makes a type-contingent plan expecting to carry out one of the strategies after learning her type. On the other hand, we can view the two types as two different "agents," one of whom is selected by Nature to "appear" when the game is played.

In the first case, the single *ex ante* player 2 predicts her opponent's play at the *ex ante* stage, so all types of player 2 would make the same prediction about the play of player 1. Under the second interpretation, the different "agents" would each make her prediction at the *interim* stage after learning her type, and thus different "agents" can make different predictions.

It is worth emphasizing that in our setup, players plan their actions *before* they receive their signals, and so we treat player 2 as a single *ex ante* player who makes type-contingent plans. Both the aggressive and accommodating types will form the same beliefs about player 1. (For more on the different interpretations, see Fudenberg & Tirole, section 6.6.1.)

2 Bayesian Nash Equilibrium

A static game of imperfect information is called a **Bayesian game**, and it consists of the following elements:

- a set of players, $N = \{1, \dots, n\}$, and, for each player $i \in N$,
- a set of actions, A_i, with the usual $A = \times_{i \in N} A_i$,
- a set of types, Θ_i, with the usual $\Theta = \times_{i \in N} \Theta_i$,
- a probability function specifying i's belief about the type of other players given his own type, $p_i : \Theta_i \to \triangle(\Theta_{-i})$,
- a payoff function, $u_i : A \times \Theta \to \mathbb{R}$.

Let's explore these definitions. We want to represent the idea that each player knows his own payoff function but may be uncertain about the other players' payoff functions. Let $\theta_i \in \Theta_i$ be some type of player i (and so Θ_i is the set of all player i types). Each type corresponds to a different payoff function that player i might have.

We specify the pure-strategy space A_i (with elements a_i and mixed strategies $\alpha_i \in \mathcal{A}_i$) and the payoff function $u_i(a_1, \dots, a_n | \theta_1, \dots, \theta_n)$. Since each player's choice of strategy can depend on his type, we let $s_i(\theta_i)$ denote the pure strategy player i chooses when his type is θ_i ($\sigma_i(\theta_i)$ is the mixed strategy). Note that in a Bayesian game, pure strategy spaces are constructed from the type and action spaces: Player i's set of possible (pure) strategies S_i is the set of all possible functions with domain Θ_i and range A_i. That is, S_i is a collection of functions $s_i : \Theta_i \to A_i$.

If player i has k possible payoff functions, then the type space has k elements, $\#(\Theta_i) = k$, and we say that player i has k types. Given this terminology, saying that player i knows his own payoff function is equivalent to saying that he knows his type. Similarly, saying that player i may be uncertain about other players' payoff functions is equivalent to saying that he may be uncertain about their types, denoted by θ_{-i}. We use Θ_{-i} to denote the set of all possible types of the other players and use the probability distribution $p_i(\theta_{-i}|\theta_i)$ to denote player i's *belief* about the other players' types θ_{-i}, given his knowledge of his own type, θ_i.[2] For simplicity, we shall assume that Θ_i has a finite number of elements.

If player i knew the strategies of the other players as a function of their type, that is, he knew $\{\sigma_j(\cdot)\}_{j \neq i}$, player i could use his beliefs $p_i(\theta_{-i}|\theta_i)$ to compute the expected utility to each choice and thus find his optimal response $\sigma_i(\theta_i)$.[3]

Following Harsanyi, we shall assume that the timing of the static Bayesian game is as follows: (1) Nature draws a type vector $\theta = (\theta_1, \ldots, \theta_n)$, where θ_i is drawn from the set of possible types Θ_i using some objective distribution p that is common knowledge; (2) Nature reveals θ_i to player i but not to any other player; (3) the players simultaneously choose actions, player i chooses from the feasible set A_i; and then (4) payoffs $u_i(a_1, \ldots, a_n|\theta)$ are received.

Since we assumed in step (1) above that it is common knowledge that Nature draws the vector θ from the prior distribution $p(\theta)$, player i can use Bayes' Rule to compute his posterior belief $p_i(\theta_{-i}|\theta_i)$ as follows:

$$p_i(\theta_{-i}|\theta_i) = \frac{p(\theta_{-i}, \theta_i)}{p(\theta_i)} = \frac{p(\theta_{-i}, \theta_i)}{\sum_{\theta_{-i} \in \Theta_{-i}} p(\theta_{-i}, \theta_i)}.$$

Furthermore, the other players can compute the various beliefs that player i might hold depending on i's type. We shall frequently assume that the players' types are independent, in which case $p_i(\theta_{-i})$ does not depend on θ_i although it is still derived from the prior distribution $p(\theta)$.

Now that we have the formal description of a static Bayesian game, we want to define the equilibrium concept for it. The notation is somewhat cumbersome but the intuition is not: each player's (type-contingent) strategy must be the best response to the other players' strategies. That is, a **Bayesian Nash equilibrium** is just a Nash equilibrium in a Bayesian game.

Given a strategy profile $s(\cdot)$ and a strategy $s'_i(\cdot) \in S_i$ (recall that this is a type-contingent strategy, with $s'_i \in S_i$, where S_i is the collection of functions $s_i : \Theta_i \to A_i$), let $(s'_i(\cdot), s_{-i}(\cdot))$ denote the profile where player i plays $s'_i(\cdot)$ and the other players follow $s_{-i}(\cdot)$, and let

$$\left(s'_i(\theta_i), s_{-i}(\theta_{-i})\right) = \left(s_1(\theta_1), \ldots, s_{i-1}(\theta_{i-1}), s'_i(\theta_i), s_{i+1}(\theta_{i+1}), \ldots, s_N(\theta_N)\right)$$

denote the value of this profile at $\theta = (\theta_i, \theta_{-i})$.

DEFINITION 1. Let G be a Bayesian game with a finite number of types Θ_i for each player i, a prior distribution p, and strategy spaces S_i. The profile $s(\cdot)$ is a (pure-strategy) **Bayesian equilibrium** of G if, for each player i and every $\theta_i \in \Theta_i$,

$$s_i(\theta_i) \in \arg\max_{s'_i \in S_i} \sum_{\theta_{-i}} u_i\left(s'_i, s_{-i}(\theta_{-i})|\theta_i, \theta_{-i}\right) p(\theta_{-i}|\theta_i),$$

that is, no player wants to change his strategy, even if the change involves only one action by one type.[4]

[2] In practice, the players' types are usually assumed to be independent, in which case $p_i(\theta_{-i}|\theta_i)$ does not depend on θ_i, and so we can write the beliefs simply as $p_i(\theta_{-i})$.

[3] This is where the assumption that Θ_i is finite is important. If there is a continuum of types, we may run into measure-theoretic problems.

[4] The general definition is a bit more complicated but we here have used the assumption that each type has a positive probability, and so instead of maximizing the *ex ante* expected utility over all types, player i maximizes his expected utility conditional on his type θ_i for each θ_i.

Each type-contingent strategy is a best response to the type-contingent strategies of the other players. Player i calculates the expected utility of playing every possible type-contingent strategy $s_i(\theta_i)$ given his type θ_i. To do this, he sums over all possible combinations of types for his opponents, θ_{-i}, and for each combination, he calculates the expected utility of playing against this particular set of opponents: The utility, $u_i(s_i', s_{-i}(\theta_{-i})|\theta_i, \theta_{-i})$, is multiplied by the probability that this set of opponents θ_{-i} is selected by Nature: $p(\theta_{-i}|\theta_i)$. This yields the optimal behavior of player i when of type θ_i. We then repeat the process for all possible $\theta_i \in \Theta_i$ and all players.

It is easy to extend this definition to infinite type spaces. For example, suppose player 1 was uncertain about player 2's payoffs, but believed they fall within some range. We would then define player 2's types as that range, $\Theta_2 = [\underline{\theta}_2, \overline{\theta}_2]$, with player 1's beliefs being represented by a well-defined density function, $f_2(\cdot)$, over that interval. We shall see an example of this shortly.

The existence of a Bayesian equilibrium is an immediate consequence of the existence of Nash equilibrium.

2.1 An Example of Notation

You are player 1 and you are playing with two opponents, A and B. Each of them has two types. Player A can be either t_A^1 with probability p_A or t_A^2 with probability $1 - p_A$, and player be can be either t_B^1 with probability p_B or t_B^2 with probability $1 - p_B$. Each of these types has two actions at his disposal. Player A can choose either a_1 or a_2, and player B can chooses either b_1 or b_2. You can choose from actions c_1, c_2, and c_3 and you can be one of two types, θ^1 or θ^2.

We let player 1 be player i and use Definition 1. First define θ_{-i}, the set of all possible combination of opponent types. Since there are two opponents with two types each, there are four combinations to consider:

$$\Theta_{-1} = \left\{\left(t_A^1, t_B^1\right), \left(t_A^1, t_B^2\right), \left(t_A^2, t_B^1\right), \left(t_A^2, t_B^2\right)\right\}$$

Of course, $\Theta_1 = \left\{\theta^1, \theta^2\right\}$. For each $\theta_1 \in \Theta_1$, we have to define $s_1(\theta_1)$ as the strategy that maximizes player 1's payoff given what the opponents do when we consider all possible combinations of opponents types, Θ_{-i}.

Note that the probabilities associated with each type of opponent allow player 1 to calculate the probability of a particular combination being realized. Since the two player types are uncorrelated, the joint probabilities are just the multiple of each individual probability, which gives us the following probabilities $p(\theta_{-1}|\theta_1)$:

$$p(t_A^1, t_B^1) = p_A p_B \qquad\qquad p(t_A^2, t_B^1) = (1 - p_A) p_B$$
$$p(t_A^1, t_B^2) = p_A(1 - p_B) \qquad\qquad p(t_A^2, t_B^2) = (1 - p_A)(1 - p_B)$$

where we suppressed the conditioning on θ^1 because the realizations are also independent from player 1's own type.

We now fix a strategy profile for the other two players to check player 1's optimal strategy for that profile. The players are using type-contingent strategies themselves. Given the number of available actions, the possible (pure) strategies are $s_A(t_A^1) = s_A(t_A^2) \in \{a_1, a_2\}$, and $s_B(t_B^1) = s_B(t_B^2) \in \{b_1, b_2\}$. So, suppose we want to find player 1's best strategy against the profile where both types of player A choose the same action, $s_A(t_A^1) = s_A(t_A^2) = a_1$, but the two types of player B choose different actions, $s_B(t_B^1) = b_1$, and $s_B(t_B^2) = b_2$.

We have to calculate the summation over all θ_{-1}, of which there are four. For each of these, we calculate the probability of this combination of opponents occurring (we did this above) and then multiply it by the payoff player 1 expects to get from his strategy if he is matched with these particular types of opponents.

This gives the expected payoff of player 1 from following his strategy against opponents of the particular type. Once we add all the terms, we have player 1's expected payoff from his strategy.

So, suppose we want to calculate player 1's expected payoff from playing $s_1(\theta^1) = c_1$:

$$u_1\left(c_1, s_A(t_A^1), s_B(t_B^1)\right) p(t_A^1, t_B^1) + u_1\left(c_1, s_A(t_A^1), s_B(t_B^2)\right) p(t_A^1, t_B^2) \\ + u_1\left(c_1, s_A(t_A^2), s_B(t_B^1)\right) p(t_A^2, t_B^1) + u_1\left(c_1, s_A(t_A^2), s_B(t_B^2)\right) p(t_A^2, t_B^2),$$

or simply:

$$u_1(c_1, a_1, b_1)\, p_A p_B + u_1(c_1, a_1, b_2)\, p_A(1 - p_B) \\ + u_1(c_1, a_1, b_1)(1 - p_A)p_B + u_1(c_1, a_1, b_2)\, p(1 - p_A)(1 - p_B).$$

We would then do this for actions c_2 and c_3, and then pick the action that yields the highest payoff from the three calculations. This is the arg max strategy. That is, it is the strategy that maximizes the expected utility.[5] This yields type θ^1 the best response to the strategy profile specified above.

We shall have to find the optimal response to this strategy profile if player 1 is of type θ^2. We then have to find player A's and player B's optimal strategies given what they know about the other players. Once all of these best responses are found, we can match them to see which constitute profiles with strategies that are mutual best responses. That is, we then proceed as before, when we found best responses and equilibria in normal form games.

2.2 From Common Priors to Uncommon Posteriors

The definition of a Bayesian game requires that players start with a *common prior belief* about the possible distribution over the chance moves ("choices" by Nature). Once each player privately learns his own type, it is possible that this information also conveys something about the probabilities of the other players' types. If this is the case, then not only the player's *posterior belief* about his own type will be different from the estimates of the other players about him, but his beliefs about the other players will also be different from his priors.

Consider first an example, in which learning one's own type does not give players any additional insight into the types of the others. Imagine there are two players, player 1 and player 2, and that for each player there are two possible types. Player i's possible types are $t_i \in \{\theta_i, \theta_i'\}$. Suppose that the types are *independently distributed* with $\Pr(t_1 = \theta_1) = p$ and $\Pr(t_2 = \theta_2) = q$. This assumption means that after learning one's own type, players retain their priors about the type of the other player. For instance, in a crisis bargaining model, the type might be a player's valuation of the issue, which has nothing to do with how much the other player values it. In this case, for a given strategy profile (s_1^*, s_2^*), the expected payoff of player 1 of type θ_1 is simply:

$$qu_1(s_1^*(\theta_1), s_2^*(\theta_2)|\theta_1, \theta_2) + (1-q)u_1(s_1^*(\theta_1), s_2^*(\theta_2')|\theta_1, \theta_2'),$$

and for a given mixed strategy profile (σ_1^*, σ_2^*), the expected payoff of player 1 of type θ_1 is:

$$q\sum_{a \in A} \sigma_1^*(a_1|\theta_1)\sigma_2^*(a_2|\theta_2)u_1(a_1, a_2|\theta_1, \theta_2) + (1-q)\sum_{a \in A} \sigma_1^*(a_1|\theta_1)\sigma_2^*(a_2|\theta_2')u_1(a_1, a_2|\theta_1, \theta_2').$$

[5] Unlike the max of an expression which denotes the expression's maximum when the expression is evaluated, the arg max operator finds the *parameter(s)*, for which the expression attains its maximum value. In our case, the arg max simply instructs us to pick the strategy that yields the highest expected payoff when matched against the profile of opponent's strategies.

In each case, we simply used the prior beliefs about the possible types of player 2 because player 1's knowledge of his own type did not tell him anything new about player 2's type. A Bayesian equilibrium will consist of *four* type-contingent strategies, one for each type of each player. Some equilibria may depend on particular values of p and q, and others may not.

Imagine now that the types of players are correlated. For instance, in a crisis bargaining model a player's type could be his probability of winning the war. If there is one true underlying probability of victory for a player, then learning something about one's own likelihood of victory also conveys something about the opponent's type. Suppose that in our example the types occur in combinations with prior probabilities as shown in Fig. 5 (p. 11). These are common knowledge.

		Player 2's type	
		θ_2	θ_2'
Player 1's type	θ_1	$1/6$	$1/3$
	θ_1'	$1/3$	$1/6$

Figure 5: Prior Distribution over the Types.

Suppose now that player 1 learns that his type is θ_1. What should his belief about player 2's type be? The prior distribution defines the joint probabilities $\Pr(t_i \cap t_{-i})$, which in turn gives the unconditional probabilities like $\Pr(t_1 = \theta_1) = \Pr(t_1 = \theta_1 \cap t_2 = \theta_2) + \Pr(t_1 = \theta_1 \cap t_2 = \theta_2') = 1/6 + 1/3 = 1/2$. The conditional probability formula (Bayes rule) yields:

$$\Pr(t_2 = \theta_2 | t_1 = \theta_1) = \frac{\Pr(t_1 = \theta_1 \cap t_2 = \theta_2)}{\Pr(t_1 = \theta_1)} = \frac{1/6}{1/2} = 1/3.$$

In other words, upon learning that his type is θ_1, player 1 must update to believe that the probability that player 2's type is θ_2 is $1/3$ from his prior belief that it was $1/2$. The other posterior beliefs are defined analogously.

3 Some Simple Games

3.1 Myerson's Exercise 3.5

Two players are not sure which of the two games, whose payoff matrices are shown in Fig. 6 (p. 11), is being played. It is common knowledge that the probability that the game is Game A is 0.9, and the information is symmetric (that is, neither player knows anything that the other player does not know).

Game A (0.9)

		Player 2: L	Player 2: R
Player 1	U	2, 2	−2, 0
	D	0, −2	0, 0

Game B (0.1)

		Player 2: L	Player 2: R
Player 1	U	0, 2	1, 0
	D	1, −2	2, 0

Figure 6: Myerson's Exercise.

Unlike our earlier example, where player 1 knows his own type, in this game he does not (note that player 2's payoffs are the same in both games). This actually makes the game even easier to solve because after the initial move by Nature that selects the game, each player has only one information set. The expected payoffs are:

$$\begin{aligned}
U_1(U, L) &= 2(.9) + 0(.1) = 1.8 & U_2(U, L) &= 2(.9) + 2(.1) = 2 \\
U_1(U, R) &= -2(.9) + 1(.1) = -1.7 & U_2(U, R) &= 0 \\
U_1(D, L) &= 0(.9) + 1(.1) = 0.1 & U_2(D, L) &= -2 \\
U_1(D, R) &= 0(.9) + 2(.1) = 0.2 & U_2(D, R) &= 0.
\end{aligned}$$

The resulting payoff matrix for the Bayesian game is shown in Fig. 7 (p. 12).

		Player 2	
		L	R
Player 1	U	$1.8, \quad 2$	$-1.7, 0$
	D	$0.1, -2$	$0.2, 0$

Figure 7: The Strategic Form of the Game from Fig. 6 (p. 11).

There are two Nash equilibria in pure strategies, $\langle U, L\rangle$ and $\langle D, R\rangle$, and a mixed strategy equilibrium $\langle {}^{1}\!/_{2}[U], {}^{19}\!/_{36}[L]\rangle$.

This is an interesting result because in two of these equilibria, player 2 chooses L with positive probability. If you look back at the original payoff matrices in Fig. 6 (p. 11), this may surprise you because $\langle D, R\rangle$ is a Nash equilibrium in both separate games (it is the unique equilibrium in Game B). On the other hand, the result is perhaps not surprising because $\langle U, L\rangle$ in Game A is Pareto-dominant, and because of the very high likelihood that this game is the one being played it is also the Pareto-dominant outcome in the Bayesian game.

Another interesting aspect of this game is that if we relax the common knowledge assumption (about the probabilities with which the games are played), then there will be no Bayesian equilibrium where player 2 would choose L. (This is a variant of Rubinstein's electronic mail game.)

3.2 The Lover-Hater Game

Suppose that player 2 has complete information and two types, L and H. Type L loves going out with player 1 whereas type H hates it. Player 1 has only one type and is uncertain about player 2's type and believes the two types are equally likely. We can describe this formally as a Bayesian game:

- Players: $N = \{1, 2\}$
- Actions: $A_1 = A_2 = \{F, B\}$
- Types: $\Theta_1 = \{x\}$, $\Theta_2 = \{l, h\}$
- Beliefs: $p_1(l|x) = p_1(h|x) = {}^{1}\!/_{2}$, $p_2(x|l) = p_2(x|h) = 1$
- Payoffs: u_1, u_2 as described in Fig. 8 (p. 13).

We shall solve this game using two different methods.

	F	B
F	$2,1$	$0,0$
B	$0,0$	$1,2$

Player 2's type is L

	f	b
F	$2,0$	$0,2$
B	$0,1$	$1,0$

Player 2's type is H

Figure 8: The Lover-Hater Battle of the Sexes.

3.2.1 Solution: Conversion to Strategic Form

We can easily convert this to strategic form, as shown in Fig. 9 (p. 13). It is immediately clear that Bb strictly dominates Ff for player 2, so she will never use the latter in any equilibrium. Finding the PSNE is easy by inspection: $\langle F, Fb\rangle$. We now look for MSNE. Observe that player 2 will always play Bb with positive probability in every MSNE. To see this, suppose that there exists some MSNE, in which $\sigma_2(Bb) = 0$. But if she does not play Bb, then F strictly dominates B for player 1, so he will choose F, to which player 2's best response is Fb. That is, we are back in the PSNE $\langle F, Fb\rangle$, and there's no mixing.

Player 1 \ Player 2		Ff	Fb	Bf	Bb
	F	$2, 1/2$	$1, 3/2$	$1,0$	$0,1$
	B	$0, 1/2$	$1/2, 0$	$1/2, 3/2$	$1,1$

Figure 9: The Lover-Hater Game in Strategic Form.

We conclude that in any MSNE, $\sigma_2(Bb) > 0$. We now have three possibilities to consider, depending on which of the remaining two pure strategies she includes in the support of her equilibrium strategy. Let p denote the probability that player 1 chooses F, and use the shortcuts $q_1 = \sigma_2(Fb)$, $q_2 = \sigma_2(Bf)$, and $q_3 = \sigma_2(Bb)$. We now examine each possibility separately:

- Suppose $q_1 = 0$ and $q_2 > 0$, which implies $q_3 = 1 - q_2$. Since player 2 is willing to mix between Bf and Bb, her expected payoffs from these pure strategies must be equal. Since $U_2(p, Bb) = 1$ and $U_2(p, Bf) = (3/2)(1-p)$, this implies $1 = 3(1-p)/2 \Rightarrow p = 1/3$. That is, player 1 must be willing to mix too. This means the payoffs from his pure strategies must be equal. Since $U_1(F, q_2) = q_2$ and $U_1(B, q_2) = (1/2)q_2 + 1(1-q_2)$, this implies $q_2 = (1/2)q_2 + 1 - q_2 \Rightarrow q_2 = 2/3$. We only need to check that $q_1 = 0$ is rational, which will be the case if $U_2(p, Fb) \leq U_2(Bb)$. Since $U_2(p, Fb) = (3/2)p = 1/2$ and $U_2(Bb) = 1$, this inequality holds. Therefore, we do have a MSNE: $\langle p = 1/3, (q_2 = 2/3, q_3 = 1/3)\rangle$. In this MSNE, player 1 chooses F with probability $1/3$ and player 2 mixes between Bf and Bb; that is, she chooses B if she is the L type, and chooses f with probability $2/3$ if she is the H type.

- Suppose $q_1 > 0$ and $q_2 = 0$, which implies $q_3 = 1 - q_1$. Since player 2 is willing to mix, it follows that $U_2(p, Fb) = (3/2)p = 1 = U_2(p, Bb)$. This implies $p = 2/3$, so player 1 must be mixing too. For him to be willing to do so, it must be the case that his payoffs from the pure strategies are equal. Since $U_1(F, q_2) = q_1$ and $U_1(B, q_2) = (1/2)q_1 + 1(1-q_1)$, this implies $q_1 = 2/3$. We only need to check if player 2 would be willing to leave out Bf. Since $U_2(p, Bf) = (3/2)p = 1 = U_2(p, Bb)$, including that strategy will not improve her expected payoff. Therefore, we do have another MSNE: $\langle p = 2/3, (q_1 = 2/3, q_3 = 1/3)\rangle$. In this MSNE, player 1 chooses F with probability $2/3$ and player

2 mixes between Fb and Bb; that is, she chooses F with probability $1/3$ if she is the L type, and chooses b if she is the H type.

- Suppose $q_1 > 0$ and $q_2 > 0$. Since player 2 is willing to mix, it follows that $U_2(p, Fb) = U_2(p, Bf) = U_2(p, Bb) = 1$. Since $U_2(p, Fb) = (3/2)p = U_2(p, Bf) = 3/2(1-p)$, it follows that $p = 1/2$. However, from $U_2(p, Bf) = U_2(p, Bb)$ we obtain $(3/2)(1 - 1/2) = 3/4 < 1 = U_2(p, Bb)$, a contradiction. Therefore there is no such MSNE.

We conclude that this game has three equilibria, one in pure strategies and the others in mixed.

3.2.2 Solution: Best Responses

Since player 1 has only one type, we suppress all references to his type from now on. Let's begin by analyzing player 2's optimal behavior for each of the two types. Let p denote the probability that player 1 chooses F, q_1 denote the probability that the L type chooses F, and q_2 denote the probability that the H type chooses f. Observe that q_1 and q_2 do *not* mean the same thing they did in the previous method (where they designated probabilities for pure strategies). In other words, in the strategic-form method, these were elements of a mixed strategy. Here, they are elements of a behavioral strategy.

Let's derive the best response for player 2. If she's type L, the expected utility from playing F is $U_L(p, F) = p$, while the expected utility from playing B is $U_L(p, B) = 2(1 - p)$. Therefore, she will choose F whenever $p \geq 2(1-p) \Rightarrow p \geq 2/3$. This yields the best response:

$$BR_L(p) = \begin{cases} q_1 = 1 & \text{if } p > 2/3 \\ q_1 \in [0,1] & \text{if } p = 2/3 \\ q_1 = 0 & \text{if } p < 2/3 \end{cases}$$

If she is of type H, the expected payoffs are $U_H(p, F) = 1 - p$ and $U_H(p, B) = 2p$. Therefore, she will choose F whenever $1 - p \geq 2p \Rightarrow p \leq 1/3$. This yields the best response:

$$BR_H(p) = \begin{cases} q_2 = 1 & \text{if } p < 1/3 \\ q_2 \in [0,1] & \text{if } p = 1/3 \\ q_2 = 0 & \text{if } p > 1/3 \end{cases}$$

Finally, we compute the expected payoffs for player 1:

$$U_1(F, q_1 q_2) = (1/2)\,[2q_1 + 0(1-q_1)] + (1/2)\,[2q_2 + 0(1-q_2)] = q_1 + q_2$$
$$U_1(B, q_1 q_2) = (1/2)\,[0q_1 + 1(1-q_1)] + (1/2)\,[0q_2 + 1(1-q_2)] = 1 - (q_1 + q_2)/2.$$

Hence, choosing F is optimal whenever $q_1 + q_2 \geq 1 - (q_1 + q_2)/2 \Rightarrow q_1 + q_2 \geq 2/3$. This yields player 1's best response:

$$BR_1(q_1 q_2) = \begin{cases} p = 1 & \text{if } q_1 + q_2 > 2/3 \\ p \in [0,1] & \text{if } q_1 + q_2 = 2/3 \\ p = 0 & \text{if } q_1 + q_2 < 2/3 \end{cases}$$

We now must find a triple (p, q_1, q_2) such that player 1's strategy is a best response to the strategies of both types of player 2 and each type of player 2's strategy is a best response to player 1's strategy. Let's check for various types of equilibria.

Can it be the case that player 1 uses a pure strategy in equilibrium? There are two cases to consider. Suppose $p = 0$, which implies $q_1 + q_2 < 2/3$ from BR_1. We now obtain $q_1 = 0$ from BR_L and $q_2 = 1$ from BR_H, which means $q_1 + q_2 = 1$, a contradiction. Hence, there is no such equilibrium.

Suppose now $p = 1$, which implies $q_1 + q_2 > 2/3$. We now obtain $q_1 = 1$ from BR_L and $q_2 = 0$ from BR_H, which means $q_1 + q_2 = 1$. This satisfies the requirement for player 1's strategy to be a best response. Therefore, we obtain a pure-strategy Bayesian equilibrium: $\langle F, Fb \rangle$.

In all remaining solutions, player 1 must mix in equilibrium. Since he is willing to mix, BR_1 implies that $q_1 + q_2 = 2/3$ (otherwise he'd play a pure strategy). This immediately means that it cannot be the case that player 2 chooses the fight, whatever her type may be. To see this, note that if some type goes to the fight for sure, $q_1 + q_2 \geq 1$, which will contradict the requirement that allows player 1 to mix. Therefore, $q_1 < 1$ and $q_2 < 1$. It also cannot be the case that neither type goes to the fight because that would imply $q_1 + q_2 = 0$, which also cannot be true in a MSNE. Therefore, either $q_1 > 0$, or $q_2 > 0$, or both. Let's consider each separately:

- Suppose $q_1 = 0$ and $q_2 > 0$: since H is willing to mix, BR_H implies that $p = 1/3$ and since $q_1 = 0$, BR_L implies $p < 2/3$. Therefore, $p = 1/3$ will make these strategies best responses. To get player 1 to mix, it has to be the case that $q_2 = 2/3$. This yields the following MSNE: $\langle p = 1/3, (q_1 = 0, q_2 = 2/3) \rangle$. In this equilibrium, player 1 chooses F with probability $1/3$, player 2 picks B if she's the L type and picks f with probability $2/3$ if she is the H type.

- Suppose $q_1 > 0$ and $q_2 = 0$: since L is willing to mix, BR_L implies that $p = 2/3$ and since $q_2 = 0$, BR_H implies that $p > 1/3$. Therefore, $p = 2/3$ will make these strategies best responses. To get player 1 to mix, it has to be the case that $q_1 = 2/3$. This yields another MSNE: $\langle p = 2/3, (q_1 = 2/3, q_2 = 0) \rangle$. In this equilibrium, player 1 chooses F with probability $2/3$, player 2 picks F with probability $2/3$ if she's the L type and picks b if she is the H type.

- Suppose $q_1 > 0$ and $q_2 > 0$: since L is willing to mix, BR_L implies that $p = 2/3$ but since H is also willing to mix, BR_H implies that $p = 1/3$, a contradiction. Therefore, there is no such MSNE.

This exhausts the possibilities, and voilà! We conclude that the game has three equilibria, one in pure strategies and two in mixed strategies. Clearly, these solutions are the same we found with the other method.

Which method should you use? Whatever works for you! In general, however, the best-response method is often more convenient despite the apparent difficulty in multiplying the types whose responses must be considered. This is because when one is considering the best responses for each type, the strategies for all other types are held constant.

3.3 The Market for Lemons

If you have bought or sold a used car, you know something about markets with asymmetric information. Typically, the seller knows far more about the car he is offering than the buyer. Generally, buyers face a significant informational disadvantage. As a result, you might expect that buyers will tend not to do very well in the market, making them cautious and loath to buy used cars, in turn making sellers worse off when the market fails due to lack of demand.

Let's model this! Suppose you, the buyer, are in the market for a used car. You meet me, the seller, through an add in the *Penny Pincher* (never a good place to look for a good car deal), and I offer you an attractive 15-year old *Firebird* for sale. You love the car, it has big fat tires, it peels rubber when you hit the gas, and it's souped up with a powerful 6-cylinder engine. It also sounds cool and has a red light in the interior. You take a couple of rides around the block and it handles like a dream.[6]

Then you suddenly have visions of the souped up engine exploding and blowing you up to smithereens, or perhaps a tire getting loose just as you screech around that particularly dangerous turn on US-1. In any case, watching the fire-fighters dowse your vehicle while you cry at the curb or watching said fire-fighters scrape you and your car off the rocks, is not likely to be especially amusing. So you tell me, "The car looks great, but how do I know it's not a lemon?"

I, being completely truthful and honest as far as used car dealers can be, naturally respond with "Oh! I've taken such good care of it. Here're are all the receipts from the regular oil changes. See? No receipts for repairs to the engine because I have *never had problems* with it! It's a peach, trust me."

You have, of course, taken my own course on repeated games and so say, "A-ha! But you will not deal with me in the future again after the sale is complete, and so you have no interest in cooperating today because I cannot punish you tomorrow for not cooperating today! You will say whatever you think will get me to buy the car."

I sigh (Blasted game theory! It was so much easier to cheat people before.) and tell you, "Fair enough. The *Blue Book* value of the car is $v > 0$ dollars. Take a look at the car, take a couple of more rides around the block if you wish and then decide whether you are willing to pay the *Blue Book* price and I will decide whether to offer you the car at that price."

We shall assume that if a car is peach, it is worth B (you, the buyer) \$3,000 and worth S (me, the seller) \$2,000. If it is a lemon, then it is worth \$1,000 to B and \$0 to S. In each case, your valuation is higher than mine, so under complete information trade should occur with the surplus of \$1,000 divided between us. However, there is asymmetric information. I know the car's condition, while you only estimate that the likelihood of it being a peach is $r \in (0, 1)$. Each of us has two actions, to trade or not trade at the market price $v > 0$. The market price is $v < \$3,000$, so a trade is, in principle, possible. (If the price were higher, the buyer would not be willing to purchase the peach even if he knew for sure that it was a peach.) We simultaneously announce what we are going to do. If we both elect to trade, then the trade takes place. Otherwise, I keep the car and you go home to deplore the evils of simultaneous-moves games. The situation is depicted in Fig. 10 (p. 16) with rows for the buyer and columns for the seller (payoffs in thousands).

		S	
		T	N
B	T	$3 - v, v$	$0, 2$
	N	$0, 2$	$0, 2$

peach

		S	
		t	n
B	T	$1 - v, v$	$0, 0$
	N	$0, 0$	$0, 0$

lemon

Figure 10: The Market for Lemons.

Let's derive the best responses. Here S can be thought of as having two types, L if her car is a lemon, and P if her car is a peach. Fix an arbitrary strategy for the buyer, let p denote the probability that he elects to trade, and calculate seller's best responses (the probability that she chooses trade) as a function of this strategy. Let q_1 denote the probability that a seller with a peach trades and q_2 denote the probability that a seller with a lemon trades.

The seller with a peach will get $U_P(p, T) = pv + 2(1 - p)$ if she trades and $U_P(p, N) = 2$ if she does not trade. Therefore, she will trade whenever $pv + 2(1 - p) \geq 2 \Rightarrow p(v - 2) \geq 0$. This yields her best

[6] Any resemblance to the second car I drove in college is purely coincidental.

response:

$$BR_P(p) = \begin{cases} q_1 = 1 & \text{if } p > 0 \text{ and } v > 2 \\ q_1 \in [0,1] & \text{if } p = 0 \text{ or } v = 2 \\ q_1 = 0 & \text{if } p > 0 \text{ and } v < 2 \end{cases}$$

The seller with a lemon will get $U_L(p,t) = pv$ if she trades, and $U_2(p,n) = 0$ if she does not trade. Therefore, she will trade whenever $pv \geq 0$. Since $v > 0$, this yields her best response:

$$BR_L(p) = \begin{cases} q_2 = 1 & \text{if } p > 0 \\ q_2 \in [0,1] & \text{if } p = 0 \end{cases}$$

Observe, in particular, that for any $p > 0$ (that is, whenever the buyer is willing to trade with positive probability), she always puts the lemon on the market. Turning now to the buyer, we see that his expected payoff from trading is:

$$U_B(T, q_1 q_2) = (3-v)rq_1 + (1-v)(1-r)q_2.$$

Since his expected payoff from not trading is $U_B(N, q_1 q_2) = 0$, he will trade whenever $(3-v)rq_1 + (1-v)(1-r)q_2 \geq 0$. Letting $R = r/(1-r) > 0$, this yields $R(3-v)q_1 \geq (v-1)q_2$. Hence, the best response is:

$$BR_B(q_1 q_2) = \begin{cases} p = 1 & \text{if } R(3-v)q_1 > (v-1)q_2 \\ p \in [0,1] & \text{if } R(3-v)q_1 = (v-1)q_2 \\ p = 0 & \text{if } R(3-v)q_1 < (v-1)q_2 \end{cases}$$

As before, we must find a profile (p, q_1, q_2), along with some possible restrictions on r, such that the buyer strategy is a best-response to the strategies of the two types of sellers, and each seller type's strategy is a best response to the buyer's strategy. We shall look for equilibria where trade occurs with positive probability, that is where $p > 0$.[7] This immediately means that the seller with the lemon always trades in equilibrium, so $q_2 = 1$. Observe further that for the seller of the peach to mix in a trading equilibrium, $v = 2$ is necessary. This is a knife-edge condition on the *Blue Book* price and the solution is not interesting because it will not hold for any v slightly different from 2. Therefore, we shall suppose that $v \neq 2$. This immediately means that we only have two cases to consider, both in pure strategies: the seller with the peach either trades or does not.

Suppose $q_1 = 0$, so the seller with the peach never trades. Since $p > 0$, this implies that $v < 2$. Looking at the condition in $BR_B(01)$, we see that the best response is $p = 0$ if $v > 1$ and $p = 1$ if $v < 1$. Since we are looking for a trade equilibrium, we conclude that if $v \leq 1$, there exists an equilibrium in which only the lemon is brought to the market; the seller with the peach stays out and the buyer obtains the lemon at the low price $v < 1$. The PSNE is $\langle T, Nt \rangle$.

Suppose now $q_1 = 1$, so both peaches and lemons are traded. Since the seller of the peach is willing to trade, this means $v > 2$. In other words, a necessary condition for the existence of this equilibrium is that the *Blue Book* price of the car exceeds the seller's valuation of the peach. (If it did not, he would never trade at that price.) Looking now at $BR_B(11)$, we find that $p = 1$ whenever $R(3-v) > v - 1$, which is satisfied whenever:

$$R > \frac{v-1}{3-v} = x \quad \Rightarrow \quad r > \frac{x}{1+x} \quad \Leftrightarrow \quad r > \frac{v-1}{2} > \frac{1}{2},$$

[7] There is an equilibrium where B buys with probability 0 and both types of S sell with probability 0, but it is not terribly interesting because it relies on knife-edge indifference conditions.

where the last inequality follows from $v > 2$. If this is satisfied, then the PSNE is $\langle T, Tt \rangle$. In other words, this equilibrium exists only if the prior probability of the car being a peach is sufficiently high.

We conclude that if the *Blue Book* price is too low ($v < 1$), then in equilibrium only the lemon is traded. If, on the other hand, the price is sufficiently high ($v > 2$), then in equilibrium both the lemon and the peach are traded provided the buyer is reasonably confident that the car is a peach ($r > 1/2$). If the price is intermediate, $1 < v < 2$, then no trade will occur in equilibrium.[8]

The results are not very encouraging for the buyer: It is not possible to obtain an equilibrium where only peaches are traded and lemons are not. Whenever trade occurs, either both types of cars are sold, or only the lemon is sold. Furthermore, if $r < 1/2$, then only lemons are traded in equilibrium. Thus, with asymmetric information markets can sometimes fail.

3.4 The Game of Chicken with Two Types

Consider the standard two-player game of Chicken, in which the drivers simultaneously choose whether to continue (C) or swerve (S). If both swerve, both are chicken and get no respect but neither is harmed, so their payoffs are 0 each. If i continues while $-i$ swerves, then i neither is harmed but i gains respect with a payoff of $R > 0$, whereas $-i$ is declared chicken with a payoff $-R$. If both continue, they split respect but get themselves into an accident, which ends with punishment by their parents that imposes a personal cost k, so the individual payoffs to the teenagers are $R/2 - k$. Imagine now that the parents of each teen can be either lenient or harsh, so that the costs they impose are either low, $k = 0$, or high, $k = K$, with $K > R$. It is common knowledge that the parents are either lenient or harsh with equal probability.

We can think of each player being of two types, $t_i \in \{\text{lenient}, \text{harsh}\}$, depending on what parents they have, so that the prior distribution of types is just ¼ for each pair, as shown in Fig. 11 (p. 18).

		Player 2's type: lenient	Player 2's type: harsh
Player 1's type	lenient	$1/4$	$1/4$
	harsh	$1/4$	$1/4$

Figure 11: Prior Distribution over the Types of Parents.

Each player knowing the type of their own parents tells them nothing about the type of parents of the other player. Thus, $\Pr(t_{-i} = \text{harsh}|t_i = \text{harsh}) = \Pr(t_{-i} = \text{harsh}|t_i = \text{lenient}) = 1/2$. This makes it easy to calculate the expected payoffs for each type of teen. We shall do this for player 1 since player 2's payoffs are computed analogously. In each calculation, we write player 2's strategy as a pair $(a_L a_H)$, which represents what action she will take if she has lenient parents (a_L) or harsh ones (a_H). If player 1 of type t_1 continues, his expected payoffs will be:

$$U_1(C, CC|t_1) = R/2 - k$$
$$U_1(C, CS|t_1) = U_1(C, SC|t_1) = (1/2)\,(R/2 - k) + (1/2)R = 3R/4 - k/2$$
$$U_1(C, SS|t_1) = R$$

[8]To see this, observe that $v < 2$ implies $q_1 = 0$ because the seller of the peach will not bring it to the market. But then $R(3 - v)q_1 = 0 < (v - 1)q_1$ for any value of $q_1 > 0$ because $v > 1$. Therefore, $p = 0$, and the buyer is not willing to trade. We can actually find equilibria with $p = 0$ and $q_1 > 0$ and $q_2 > 0$ as well. To see this, note that if the buyer is sure not to trade, both types of sellers can mix with $v < 2$. Hence, any pair (q_1, q_2) that satisfies $R(3 - v)q_1 < (v - 1)q_2$ will actually work. Obviously, it will have to be the case that $q_1 < q_2$ for this to work; that is, the seller of the peach is less likely to trade than the seller of the lemon. None of these equilibria are particularly illuminating beyond the fact that no trade occurs in any of them.

Note that since player 1 swerves, no accident ever occurs, which means that no parental punishment is inflicted, and so the payoffs are the same irrespective of type. Thus, if player 1 swerves, his expected payoffs will be:

$$U_1(S, CC) = -R$$
$$U_1(S, CS) = U_1(S, SC) = (1/2)(0) + (1/2)(-R) = -R/2$$
$$U_1(S, SS) = 0.$$

We can now use these payoffs to construct the strategic form, where we note that the *ex ante* expected payoff for player 1 (that is, before he learns his type) is just the expected payoff from each of his type-contingent payoffs we just calculated above times their probability (which, you should recall, is 1/2). For instance, the expected payoff from the strategy CC (continue irrespective of type) against CC for player 2 is:

$$U_1(CC, CC) = (1/2)U_1(C, CC|\text{lenient}) + (1/2)U_1(C, CC|\text{harsh}) = \frac{R-K}{2},$$

whereas the payoff from the strategy CS against the same strategy for player 2 is:

$$U_1(CS, CC) = (1/2)U_1(C, CC|\text{lenient}) + (1/2)U_1(S, CC|\text{harsh}) = -\frac{R}{4}.$$

The corresponding payoffs against the CS strategy for player 2 are:

$$U_1(CC, CS) = (1/2)U_1(C, CS|\text{lenient}) + (1/2)U_1(C, CS|\text{harsh}) = \frac{3R-K}{4}$$
$$U_1(CS, CS) = (1/2)U_1(C, CS|\text{lenient}) + (1/2)U_1(S, CS|\text{harsh}) = \frac{R}{8}.$$

The payoff matrix of the resulting Bayesian game is shown in Fig. 12 (p. 19).

	CC	CS	SC	SS
CC	$\frac{R-K}{2}, \frac{R-K}{2}$	$\frac{3R-K}{4}, -\frac{R}{4}$	$\frac{3R-K}{4}, -\frac{R}{4} - \frac{K}{2}$	$R, -R$
CS	$-\frac{R}{4}, \frac{3R-K}{4}$	$\frac{R}{8}, \frac{R}{8}$	$\frac{R}{8}, \frac{R}{8} - \frac{K}{4}$	$\frac{R}{2}, -\frac{R}{4}$
SC	$-\frac{R}{4} - \frac{K}{2}, \frac{3R-K}{4}$	$\frac{R}{8} - \frac{K}{4}, \frac{R}{8}$	$\frac{R}{8} - \frac{K}{4}, \frac{R}{8} - \frac{K}{4}$	$\frac{R}{2}, -\frac{R}{4}$
SS	$-R, R$	$-\frac{R}{4}, \frac{R}{2}$	$-\frac{R}{4}, \frac{R}{2}$	$0, 0$

Figure 12: Game of Chicken with Incomplete Information.

Observe now that CS strictly dominates SS, so we can eliminate SS. With this strategy removed, CS becomes strictly dominant over SC as well. We obtain the reduced form shown in Fig. 13 (p. 19).

	CC	CS
CC	$\frac{R-K}{2}, \frac{R-K}{2}$	$\frac{3R-K}{4}, -\frac{R}{4}$
CS	$-\frac{R}{4}, \frac{3R-K}{4}$	$\frac{R}{8}, \frac{R}{8}$

Figure 13: Game of Chicken with Incomplete Information, Reduced Form.

Suppose first that the harsh punishment is sufficiently nasty relative to reputation: $2K > 5R$. Then each player has a strictly dominant strategy, CS, so the game has a unique PSNE: $\langle CS, CS\rangle$. Teenagers with lenient parents continue but those with harsh parents swerve. Before the game begins, we would expect the probability of an accident to be

$$\begin{aligned}
\Pr(\text{accident}) &= \Pr(\text{Player 1 continues}) \times \Pr(\text{Player 2 continues}) \\
&= \Big[\Pr(\text{Player 1 continues}|t_1 = \text{lenient}) \Pr(t_1 = \text{lenient}) \\
&\qquad + \Pr(\text{Player 1 continues}|t_1 = \text{harsh}) \Pr(t_1 = \text{harsh}) \Big] \\
&\quad \times \Big[\Pr(\text{Player 2 continues}|t_2 = \text{lenient}) \Pr(t_2 = \text{lenient}) \\
&\qquad + \Pr(\text{Player 2 continues}|t_2 = \text{harsh}) \Pr(t_2 = \text{harsh}) \Big] \\
&= \Big[(1)(^1\!/\!_2) + (0)(^1\!/\!_2) \Big] \times \Big[(1)(^1\!/\!_2) + (0)(^1\!/\!_2) \Big] \\
&= {}^1\!/\!_4.
\end{aligned}$$

Analogous calculations yield the *ex ante* probability distribution over outcomes in this PSNE, as shown in Fig. 14 (p. 20).

		Player 2 continues	Player 2 swerves
Player 1	continues	$^1\!/\!_4$	$^1\!/\!_4$
	swerves	$^1\!/\!_4$	$^1\!/\!_4$

Figure 14: Probability Distribution over Outcomes in PSNE.

Observe that unlike the original Chicken game with complete information, the PSNE here generates positive probabilities over *all* outcomes (much like the MSNE in the original). In particular, there is a positive probability that disaster will occur even though each type is choosing a pure strategy. The uncertainty comes from not knowing the other player's type, which in turn encourages the teens with lenient parents to run a risk of disaster in return for the possibility of a reputational gain. On the other hand, the teens with harsh parents choose to avoid that risk and swerve but even for them there is a chance that they will not suffer reputational losses.

The point that pure-strategy Nash equilibria in a game of incomplete information generate unpredictability about behaviors, which in turn generates a probability distribution over various outcomes that we would normally only expect in mixed-strategy equilibria in games of complete information might lead one to wonder: is there some relationship between PSNE in incomplete information games and MSNE in "similar" complete information games? The answer, turns out, is positive, as we shall see in the next section.

Suppose now the harsh punishment is moderately nasty relative to reputation: $5R > 2K > 3R$. Then, we obtain the following best responses:

$$BR_i = \begin{cases} CS & \text{if } s_{-i} = CC \\ CC & \text{if } s_{-i} = CS \end{cases}$$

That is, if the opponent is certain to continue (CC), then the punishment is sufficiently harsh to deter the teen who expects to be punished with it. However, if the risk is smaller because there's a chance the opponent will swerve (CS), then this punishment is no longer sufficient as a deterrent, and even the teen with harsh parents will choose to continue.

3.5 Study Groups

Two players have to hand in a joint assignment, and simultaneously choose either to work hard (W) or to slack off (S). Working involves incurring a cost $c \in (0, 1)$. The assignment is completed successfully if at least one of the students works and fails otherwise. Students have private information about their own valuation of a successfully completed assignment: v_i^2, where each v_i is independently and uniformly distributed over the interval $[0, 1]$. These distributions are common knowledge. Each values a failed assignment at 0.

This is a game with a continuum of types for each player. Writing the extensive form will not be particularly helpful, and writing the strategic form is impossible because each player has an infinite number of type-contingent strategies despite there being only two possible actions. (Recall that a strategy maps the type to the action space, $s_i(v_i) \to \{W, S\}$.)

We do not know *a priori* how types map into actions. Let us first consider type-contingent strategies; that is, strategies that prescribe different actions for different subsets of types. Any such strategy for player $-i$, let's call it σ_{-i}, will generate a probability that this player will work, let's call this $\omega(\sigma_{-i}) \in (0, 1)$, where the open interval follows from the fact that some types work but others do not. This means that from player i's perspective, player $-i$ will work with probability $\omega(\sigma_{-i})$, which allows us to calculate that player's expected payoffs:

$$U_i(W, \sigma_{-i}|v_i) = v_i^2 - c \quad \text{and} \quad U_i(S, \sigma_{-i}|v_i) = \omega(\sigma_{-i})v_i^2.$$

If in equilibrium some type is willing to work, then it must be that $U_i(W, \sigma_{-i}|v_i) \geq U_i(S, \sigma_{-i}|v_i)$, or

$$v_i \geq \sqrt{\frac{c}{1 - \omega(\sigma_{-i})}}.$$

Consider now some type $\hat{v}_i > v_i$, and note that the above inequality must be strictly satisfied for that type (the left-hand side increases while the right-hand side remains constant). This means that this higher valuation type must be willing to work in that equilibrium as well. In other words, in any equilibrium with type-contingent strategies, if some type of player i works, then all higher-valuation types work as well. This further implies that if some type shirks in equilibrium, then all lower-valuation types must shirk as well. We can then infer, that in any such equilibrium there must be a type, v_i^*, that is indifferent between working and shirking. If σ_{-i}^* is player $-i$'s equilibrium strategy, then this type is defined as:

$$v_i^* = \sqrt{\frac{c}{1 - \omega(\sigma_{-i}^*)}}. \tag{1}$$

This result greatly simplifies our task because it means that *all* equilibria in type-contingent strategies must involve pure strategies of the form: "all $v_i \leq v_i^*$ shirk, and all $v_i > v_i^*$ work." These strategies make it easy to define the expectations of the players because the probability that the other player works is just the probability that his valuation exceeds the threshold value:

$$\omega(\sigma_{-i}^*) = \Pr(v_{-i} > v_{-i}^*) = 1 - \Pr(v_{-i} \leq v_{-i}^*) = 1 - v_{-i}^*,$$

where we used the distributional assumption that $v_{-i} \sim U[0, 1]$. Since (1) defines the threshold type for player i, we can substitute for the value of $\omega(\sigma_{-i}^*) = 1 - v_{-i}^*$ to obtain the system of equations:

$$v_i^* = \sqrt{\frac{c}{v_{-i}^*}},$$

which tells us that the threshold valuations must be the same: $v_i^* = v_{-i}^* = v^*$. This should not be surprising since we assumed that the valuation range, the cost, and the payoffs are the same. Substituting this into any of the two equations in the system yields:

$$v^* = \sqrt[3]{c}.$$

We conclude that the equilibrium in type-contingent strategies is unique and the strategy for each player i is:

$$s_i^*(v_i) = \begin{cases} S & \text{if } v_i \le \sqrt[3]{c} \\ W & \text{otherwise.} \end{cases}$$

What about non-contingent strategies where all types of a player choose the same action? Suppose player i works for sure in equilibrium. Player $-i$'s best response is to shirk, but then low-valuation types of player i do not want to work because $v_i^2 - c < 0$ for all $v_i < \sqrt{c}$. Therefore, there can be no such equilibrium. If, on the other hand, player i were sure to shirk, then player $-i$'s best response is to work if $v_{-i} > \sqrt{c}$ and shirk otherwise. But this means that her strategy is type-contingent, and we just derived the equilibrium for this case. We conclude that the equilibrium we found is unique.

How does this compare to a complete information game with common knowledge valuations (v_1, v_2), whose payoff is shown in Fig. 15 (p. 22)?

		Player 2 W	Player 2 S
Player 1	W	$v_1^2 - c, v_2^2 - c$	$v_1^2 - c, v_2^2$
	S	$v_1^2, v_2^2 - c$	$0, 0$

Figure 15: Study Groups with Complete Information.

If $v_i \le \sqrt{c}$, then the unique equilibrium is $\langle S, S \rangle$, and neither player works, so the assignment never gets done. The costs of effort are just too high relative to how much players care about the assignment. Since $\sqrt{c} < \sqrt[3]{c}$, if these two types play the incomplete information game, neither will work either, and the assignment will not be done. The (lack of) information changes nothing.

If $v_i \le \sqrt{c}$ but $v_{-i} > \sqrt{c}$, then there is an asymmetric situation, in which player i does not value the assignment to put in any effort but player $-i$ does, so the unique equilibrium involves $s_i = S$ and $s_{-i} = W$. What happens if two of these types play the incomplete information game? Clearly, v_i will shirk there as well. But what about v_{-i}? There are two cases: if $v_{-i} > \sqrt[3]{c}$, then he will contribute, so the outcome would be the same as with complete information. However, if $v_{-i} \in (\sqrt{c}, \sqrt[3]{c}]$, then this type would *not* contribute under incomplete information even though he would have done so with complete information. The reason is that with complete information, this type knows for sure that shirking means that the assignment will not be done, and he prefers to get it done. With incomplete information about the other player, this type thinks that there is a positive probability that the other player will complete the assignment, and so he attempts to free ride because his valuation is not sufficiently high to induce him to ensure the completion of the assignment. In other words, in this situation coordination will fail because of incomplete information, and the assignment will not get done. In retrospect, v_{-i} will deeply regret not putting in the effort.

If both players value the assignment highly, $v_i > \sqrt{c}$, then there are multiple equilibria. As before, there are the two asymmetric PSNE: $\langle W, S \rangle$ and $\langle S, W \rangle$. No player who expects the other to do the work has any incentive to work irrespective of their valuation of the assignment. As usual in these mixed-motive games (players want to coordinate on the assignment getting done but disagree about who is going to do the work), there is also a MSNE. Since player i is willing to mix, it follows that $U_i(W, \sigma_{-i}) = U_i(S, \sigma_{-i}$, or $v_i^2 - c = v_i^2 \sigma_{-i}(W)$, so:

$$\sigma^*_{-i}(W) = \frac{v_i - c}{v_i}.$$

In this equilibrium, there is a positive probability that the assignment will not be done because of coordination failure. This probability is

$$(1 - \sigma_1^*(W))(1 - \sigma_2^*(W)) = \frac{c^2}{v_1 v_2} \in (0, 1).$$

If these types were to play under incomplete information, then there are several possibilities depending on whether one or both of them meet the work threshold. If neither does, $v_i \leq \sqrt[3]{c}$, then they will shirk and the assignment will fail with certainty, which is clearly worse than the MSNE. If, on the other hand, at least one of them has a valuation that exceeds the threshold, then the assignment will be completed with certainty. In other words, coordinating under incomplete information might help players avoid coordination failure because it reduces the incentive to free-ride in situations where either player would be willing to complete the task.

4 Purification of Mixed Strategies: The Battle of the Sexes

Consider the following modification of the Battle of the Sexes: player 1 is unsure about player 2's payoff if they coordinate on going to the ballet, and player 2 is unsure about player 1's payoff if they coordinate on going to the fight. That is, the player 1's payoff in the (F, F) outcome is $2 + \theta_1$, where θ_1 is privately known to him; and player 2's payoff in the (B, B) outcome is $2 + \theta_2$, where θ_2 is privately known to her. Assume that both θ_1 and θ_2 are independent draws from a uniform distribution $[0, x]$.[9] Formally,

- Players: $N = \{1, 2\}$;
- Actions: $A_1 = A_2 = \{F, B\}$;
- Types: $\Theta_1 = \Theta_2 = [0, x]$;
- Beliefs: $p_1(\theta_2) = p_2(\theta_1) = 1/x$ (where we used the fact that the uniform probability density function is $f(x) = 1/x$ when specified for the interval $[0, x]$);
- Payoffs: u_1, u_2 as described in Fig. 16 (p. 23).

		Player 2	
		F	B
Player 1	F	$2 + \theta_1, 1$	$0, 0$
	B	$0, 0$	$1, 2 + \theta_2$

Figure 16: Battle of the Sexes with Two-Sided Incomplete Information.

Each player has a continuum of types (Θ_i is infinite). When considering type-contingent pure strategies, we shall look for a Bayesian equilibrium in which player 1 goes to the fight if θ_1 exceeds some threshold

[9] The choice of the distribution is not important but we do have in mind that these privately known values only slightly perturb the payoffs.

type, x_1, and goes to the ballet otherwise, and player 2 goes to the ballet if θ_2 exceeds some threshold type, x_2, and goes to the fight otherwise. These threshold (or cut-point) strategies are quite simple: given an interval of types, there exists a special type, the threshold, such that all types smaller than it do one thing, and all types greater than it do another. With payoffs that are strictly monotonic in type (either always increasing or decreasing), this means that the threshold type must be indifferent between the two actions.

Why are we looking for an equilibrium in such strategies? Because we can prove that any equilibrium must, in fact, involve cut-point strategies. This follows from the fact that if in equilibrium some type θ_1 chooses F, then it must be the case that all $\hat{\theta}_1 > \theta_1$ must also be choosing F. We can prove this by contradiction. Take some Bayesian equilibrium and some θ_1 whose optimal strategy is F. Now take some $\hat{\theta}_1 > \theta_1$ and suppose that his optimal strategy is B. We shall see that this leads to a contradiction. Since θ_1 chooses F in equilibrium,

$$U_1(F, \sigma_2^*|\theta_1) = (2 + \theta_1)\sigma_2^*(F) \geq (1)(1 - \sigma_2^*(F)) = U_1(B, \sigma_2^*|\theta_1).$$

Furthermore, since $\hat{\theta}_1$ chooses B in equilibrium, it follows that:

$$U_1(B, \sigma_2^*|\hat{\theta}_1) = (1)(1 - \sigma_2^*(F)) \geq (2 + \hat{\theta}_1)\sigma_2^*(F) = U_1(F, \sigma_2^*|\hat{\theta}_1).$$

Putting these two inequalities together yields: $(2 + \theta_1)\sigma_2^*(F) \geq (2 + \hat{\theta}_1)\sigma_2^*(F)$. If $\sigma_2^*(F) = 0$, player 1's best response is B regardless of type, which contradicts the supposition that θ_1 chooses F. Therefore, it must be the case that $\sigma_2^*(F) > 0$. We can therefore simplify the above inequality to obtain $2 + \theta_1 \geq 2 + \hat{\theta}_1 \Rightarrow \theta_1 \geq \hat{\theta}_1$. However, this contradicts $\hat{\theta}_1 > \theta$. We conclude that if some type of player 1 chooses F in equilibrium, then so must all higher types. A symmetric argument establishes that if some type of player 2 chooses B in equilibrium, then so must all higher types. In other words, players must be using cut-point strategies in any equilibrium.

Let's now go back to solving the game. We shall denote the (yet unknown) threshold type for player i by x_i, so the equilibrium probability of choosing the favorite entertainment is $\Pr[\theta_i > x_i]$. For simplicity (and with slight abuse of notation), let $\sigma_1(\theta_1)$ denote the probability that player 1 goes to the fight, that is:

$$\sigma_1(\theta_1) = \Pr[\theta_1 > x_1] = 1 - \Pr[\theta_1 \leq x_1] = 1 - \frac{x_1}{x}.$$

Similarly, the probability that player 2 goes to the ballet is

$$\sigma_2(\theta_2) = \Pr[\theta_2 > x_2] = 1 - \Pr[\theta_2 \leq x_2] = 1 - \frac{x_2}{x}.$$

Suppose the players play the strategies just specified. We now want to find x_1, x_2 that make these strategies a Bayesian equilibrium. Given player 2's strategy, player 1's expected payoffs from going to the fight and going to the ballet are:

$$\begin{aligned}
\mathrm{E}\left[u_1(F|\theta_1, \theta_2)\right] &= (2 + \theta_1)(1 - \sigma_2(\theta_2)) + (0)\sigma_2(\theta_2) = \frac{x_2}{x}(2 + \theta_1) \\
\mathrm{E}\left[u_1(B|\theta_1, \theta_2)\right] &= (0)(1 - \sigma_2(\theta_2)) + (1)\sigma_2(\theta_2) = 1 - \frac{x_2}{x}
\end{aligned}$$

Going to the fight is optimal if, and only if, the expected utility of doing so exceeds the expected utility of going to the ballet:

$$\begin{aligned}
\mathrm{E}\left[u_1(F|\theta_1, \theta_2)\right] &\geq \mathrm{E}\left[u_1(B|\theta_1, \theta_2)\right] \\
\frac{x_2}{x}(2 + \theta_1) &\geq 1 - \frac{x_2}{x} \\
\theta_1 &\geq \frac{x}{x_2} - 3.
\end{aligned}$$

Clearly, any type for whom this inequality is strict is going to the fight. This means that the *smallest* type that could go to the fight is the one, for whom it holds with equality. We shall call the type that's indifferent between going to either entertainment, the *threshold* type: $x_1 = x/x_2 - 3$. Player 2's expected payoffs from going to the ballet and going to the fight given player 1's strategy are:

$$\begin{aligned}
\mathrm{E}\left[u_2(B|\theta_1,\theta_2)\right] &= (0)\sigma_1(\theta_1) + (2+\theta_2)(1-\sigma_1(\theta_1)) = \frac{x_1}{x}(2+\theta_2) \\
\mathrm{E}\left[u_2(F|\theta_1,\theta_2)\right] &= (1)\sigma_1(\theta_1) + (0)(1-\sigma_1(\theta_1)) = 1 - \frac{x_1}{x}
\end{aligned}$$

and so going to the ballet is optimal if and only if:

$$\begin{aligned}
\mathrm{E}\left[u_2(B|\theta_1,\theta_2)\right] &\geq \mathrm{E}\left[u_2(F|\theta_1,\theta_2)\right] \\
\frac{x_1}{x}(2+\theta_2) &\geq 1 - \frac{x_1}{x} \\
\theta_2 &\geq \frac{x}{x_1} - 3.
\end{aligned}$$

Let $x_2 = x/x_1 - 3$ denote the threshold type for player 2. We now have the two threshold types, so we solve the following system of equations:

$$\begin{aligned}
x_1 &= x/x_2 - 3 \\
x_2 &= x/x_1 - 3.
\end{aligned}$$

The solution is $x_1 = x_2$ and $x_2^2 + 3x_2 - x = 0$. We now solve the quadratic, whose discriminant is $D = 9 + 4x$, for x_2. The solution is:

$$x_1 = x_2 = \frac{-3 + \sqrt{9+4x}}{2}.$$

The pair of strategies:

$$s_1(\theta_1) = \begin{cases} F & \text{if } \theta_1 > x_1 \\ B & \text{if } \theta_1 \leq x_1 \end{cases} \qquad\qquad s_2(\theta_2) = \begin{cases} F & \text{if } \theta_2 \leq x_2 \\ B & \text{if } \theta_2 > x_2 \end{cases}$$

where

$$x_1 = x_2 = \frac{-3 + \sqrt{9+4x}}{2}.$$

is thus a Bayesian equilibrium of this game.[10] In this equilibrium, the probability that player 1 goes to the fight equals the probability that player 2 goes to the ballet, and they are:

$$1 - \frac{x_1}{x} = 1 - \frac{x_2}{x} = 1 - \frac{-3 + \sqrt{9+4x}}{2x}. \tag{2}$$

It is interesting to see what happens as uncertainty disappears (i.e. x goes to 0). Taking the limit of the expression in Equation 2 requires an application of the L'Hôpital rule:

$$\lim_{x\to 0}\left[1 - \frac{-3+\sqrt{9+4x}}{2x}\right] = 1 - \lim_{x\to 0}\left[\frac{\frac{d}{dx}(-3+\sqrt{9+4x})}{\frac{d}{dx}(2x)}\right] = 1 - \lim_{x\to 0}\frac{2(9+4x)^{-1/2}}{2} = \frac{2}{3}$$

[10] There are two other pure-strategy equilibria, in which both players choose the same entertainment irrespective of their types. These replicate the two PSNE from the complete information game. Note also that the strategies do not specify what to do for $\theta_i = x_i$ because the probability of this occurring is 0 (the probability of any particular number drawn from a continuous distribution is zero). It is customary that one of the inequalities, it does not matter which, is weak in order to handle the case.

In other words, as uncertainty disappears, the probabilities of player 1 playing F and player 2 playing B both converge to $2/3$. But these are *exactly the probabilities of the mixed strategy Nash equilibrium of the complete information case!* That is, we have just shown that as incomplete information disappears, the players' behavior in the pure-strategy Bayesian equilibrium of the incomplete-information game approaches their behavior in the mixed-strategy Nash equilibrium in the original game of complete information.

Harsanyi (1973) suggested that player j's mixed strategy represents player i's uncertainty about j's choice of a pure strategy, and that player j's choice in turn depends on a small amount of private information. As we have just shown (and as can be proven for the general case), a mixed-strategy Nash equilibrium can almost always be interpreted as a pure-strategy Bayesian equilibrium in a closely related game with a little bit of incomplete information. The crucial feature of a mixed-strategy Nash equilibrium is not that player j chooses a strategy randomly, but rather that player i is uncertain about player j's choice. This uncertainty can arise either because of randomization or (more plausibly) because of a little incomplete information, as in the example above.

This is called **purification of mixed strategies**. That is, this example show Harsanyi's result that it is possible to "purify" the mixed-strategy equilibrium of almost any complete information game (except some pathological cases) by showing that it is the limit of a sequence of pure-strategy equilibria in a game with slightly perturbed payoffs. This a defense of mixed-strategies that does not require players to randomize deliberately, and in particular it does not require them to randomize with the required probabilities. Recall that in MSNE the player is indifferent among every mixture with the support of the MSNE strategy. There is no compelling reason why he should choose the randomization "required" by the equilibrium. Harsanyi's defense of mixed strategies gets around this problem very neatly because players here do not randomize, it is just that their behavior can appear that way to the other asymmetrically informed players.