Perfect Bayesian Equilibrium

For an important class of extensive games, a solution concept is available that is simpler than sequential equilibrium, but with similar properties.

In a **Bayesian extensive game with observable actions**, nature moves first and independently selects a type for each player. Afterwards, the actions chosen by players are observed by all.

Definition 231.1: A **Bayesian extensive game with observable actions** is a tuple, $\langle \Gamma, (\Theta_i), (p_i), (u_i) \rangle$ where

1. $\Gamma = \langle N, H, P \rangle$ is an extensive game form with perfect information and (possibly) simultaneous moves

2. The set of types for player i, Θ_i, is a finite set, and we denote $\Theta = \times_{i \in N} \Theta_i$

3. p_i is a probability measure on Θ_i, where $p_i(\theta_i) > 0$ for all $\theta_i \in \Theta_i$, and these measures are independent

4. $u_i : \Theta \times Z \to \mathbb{R}$ is a von Neumann-Morgenstern utility function.

The set of histories is $\{\varnothing\} \cup (\Theta \times H)$

The information set for player $i \in P(h)$ is of the form

$$I(\theta_i, h) = \{((\theta_i, \theta'_{-i}), h) : \theta'_{-i} \in \Theta_{-i}\}$$

To solve the game, we will be looking for a profile of behavioral strategies in Γ (for each player and each type) $((\sigma_i(\theta_i))$ and a belief system $\mu_i(h)$ that specifies a common belief, after the history h, held by all players other than i about player i's type.

Let s be a profile of behavioral strategies in Γ. Define $O_h(s)$ to be the probability measure on terminal histories of Γ generated by s, given the history h.

Define $O(\sigma_{-i}, s_i, \mu_{-i} \mid h)$ to be the probability measure on terminal histories of Γ, given i uses s_i, other players use type-dependent behavioral strategies σ_{-i}, the history reached is h, and beliefs are given by μ_{-i}.

Definition 232.1: Let $\langle \Gamma, (\Theta_i), (p_i), (u_i) \rangle$ be a Bayesian extensive game with observable actions, with $\Gamma = \langle N, H, P \rangle$. A pair $((\sigma_i), (\mu_i)) = ((\sigma_i(\theta_i))_{i \in N, \theta_i \in \Theta_i}, (\mu_i(h))_{i \in N, h \in H \setminus Z})$ is a **perfect Bayesian equilibrium** if the following conditions are satisfied

1. *Sequential Rationality*: For every non-terminal history $h \in H \setminus Z$, every player $i \in P(h)$, and every type $\theta_i \in \Theta_i$, the probability measure $O(\sigma_{-i}, \sigma_i(\theta_i), \mu_{-i} \mid h)$ is weakly preferred by type θ_i to $O(\sigma_{-i}, s_i, \mu_{-i} \mid h)$ for any strategy s_i of player i in Γ.

2. *Correct initial beliefs*: $\mu_i(\varnothing) = p_i$ for each $i \in N$.

3. *Action-determined beliefs*: If $i \notin P(h)$ and $a \in A(h)$, then $\mu_i(h, a) = \mu_i(h)$. If $i \in P(h)$, $a \in A(h)$, $a' \in A(h)$, and $a_i = a'_i$, then $\mu_i(h, a) = \mu_i(h, a')$.

4. *Bayesian updating*: If $i \in P(h)$ and a_i is in the support of $\sigma_i(\theta_i)(h)$ for some θ_i in the support of $\mu_i(h)$, then for any $\theta_i' \in \Theta_i$, we have

$$\mu_i(h,a)(\theta_i') = \frac{\sigma_i(\theta_i')(h)(a_i)[\mu_i(h)(\theta_i')]}{\sum_{\theta_i \in \Theta_i} \sigma_i(\theta_i)(h)(a_i)[\mu_i(h)(\theta_i)]}.$$

Interpreting conditions 3 and 4:

Action-determined beliefs–since types are independent, beliefs about player i's type cannot be influenced by the actions of the other players. You can't signal what you don't know. If player i does not make a move after the history h, then the action profile after history h does not change μ_i. If player i does move after the history h, then the actions of players other than i do not change μ_i.

Bayesian updating–If player i's action after history h is consistent with beliefs others have about player i at h, given σ_i, then the new belief is derived from Bayes' rule.

If player i's action after history h is **inconsistent** with beliefs others have about player i at h, given σ_i, then the denominator in the Bayes' rule expression is zero. At this point, the other players form a new belief about player i, which is the basis for future Bayesian updating until σ_i is contradicted again, and so on.

Notice that perfect Bayesian equilibrium (PBE) is simply not defined for extensive games with unobservable actions. Beliefs are about nature's initial choice of types, not about which history we are at within an information set.

Proposition 234.1: Let (β, μ) be a sequential equilibrium of the extensive game associated with the finite Bayesian extensive game with observable actions, $\langle \Gamma, (\Theta_i), (p_i), (u_i) \rangle$. For every $h \in H$, $i \in P(h)$, and $\theta_i \in \Theta_i$, let $\sigma_i(\theta_i)(h) = \beta_i(I(\theta_i, h))$. Then there is a collection, $(\mu_i(h))_{i \in N, h \in H}$ where $\mu_i(h)$ is a probability measure on Θ_i, such that

$$\mu(I(\theta_i, h))(\theta, h) = \prod_{j \in N \setminus \{i\}} \mu_j(h)(\theta_j)$$

holds for all $\theta \in \Theta$ and $h \in H$ and $((\sigma_i), (\mu_i))$ is a Perfect Bayesian equilibrium of the Bayesian extensive game.

If we are in the framework of Bayesian games with observable actions, any sequential equilibrium is a PBE, although there may be PBE that are not sequential equilibria. The key to the proof is showing that the beliefs in the sequential equilibrium (about histories, conditional on an information set) can be used to construct beliefs in the PBE (about types, which are independent and commonly held).

Signaling Games

A signaling game is a special case of a Bayesian extensive game with observable actions. One player, the sender, is informed of an uncertain parameter, θ_1, and chooses an action, m. The action or message may directly affect payoffs. The other player is the receiver, who observes the message and takes an action, a.

Spence's model of education: Worker (sender) knows her productivity, θ, while the employer (receiver) does not. We think of competition between employers causing the wage to equal the expected productivity of the worker, but we simplify by assuming that there is one employer with utility $-(w- \theta)^2$.

The message is the level of education, e, and the worker's utility is $w - e/\theta$. The more productive, the easier it is to acquire education.

$$\begin{aligned} \theta &= \theta^L \text{ with probability } p^L \\ \theta &= \theta^H \text{ with probability } p^H \end{aligned}$$

Restrict attention to pure strategy equilibria: e^L and e^H.

Pooling Equilibrium: $e^L = e^H = e^*$.

When the worker chooses e^*, the employer chooses $w^* = p^L\theta^L + p^H\theta^H$. To find the values of e^* consistent with PBE, we specify beliefs off the equilibrium path that make deviations least likely. The employer believes that any deviation is from a low-productivity type, so (by sequential rationality) the wage would be

$$w(e) = \theta^L \text{ for } e \neq e^*.$$

The best deviation is to choose $e = 0$, and the worker most tempted to deviate is a low-productivity type, so sequential rationality is satisfied if and only if

$$\theta^L \leq w^* - e^*/\theta^L$$

holds, which implies

$$e^* \leq \theta^L p^H(\theta^H - \theta^L). \quad (1)$$

Thus, for any e^* satisfying (1), we have a PBE. The worker chooses e^*. The employer chooses

$$\begin{aligned} w(e) &= w^* \text{ if } e = e^* \\ w(e) &= \theta^L \text{ if } e \neq e^*. \end{aligned}$$

Beliefs are

$$\begin{aligned} \mu(e)(\theta^L) &= p^L \text{ if } e = e^* \\ \mu(e)(\theta^L) &= 1 \text{ if } e \neq e^*. \end{aligned}$$

Notice that, in these pooling equilibria, the worker gets education to avoid being labelled a loser, not to distinguish herself as a high-productivity type. (Don't be a high school dropout.) One could modify beliefs so that sufficiently high education signals a high-productivity type, but where the signaling is too costly to be worthwhile.

Notice also that PBE accomodates games with a continuum of actions, while the definition of sequential equilibrium would have to be even more complicated to accomodate these games.

Separating Equilibrium: $e^L \neq e^H$.

First, we must have $e^L = 0$ in any separating equilibrium, or else a low-productivity type should deviate to $e = 0$. To find the range of e^H consistent with a separating equilibrium, we can restrict attention to beliefs $\mu(e)(\theta^L) = 1$ if $e \neq e^H$.

Beliefs are

$$\begin{aligned} \mu(e)(\theta^L) &= 1 \text{ if } e \neq e^H \\ \mu(e)(\theta^L) &= 0 \text{ if } e = e^H. \end{aligned}$$

Sequential rationality by the employer implies

$$\begin{aligned} w(e) &= \theta^L \text{ if } e \neq e^H \\ w(e) &= \theta^H \text{ if } e = e^H. \end{aligned}$$

Clearly we have a separating PBE if and only if neither type benefits from deviating to the other type's education:

$$\theta^L \geq \theta^H - e^H/\theta^L \tag{2}$$

$$\theta^L \leq \theta^H - e^H/\theta^H. \tag{3}$$

From (2) and (3), we have a separating PBE whenever $\theta^L(\theta^H - \theta^L) \leq e^H \leq \theta^H(\theta^H - \theta^L)$ holds.

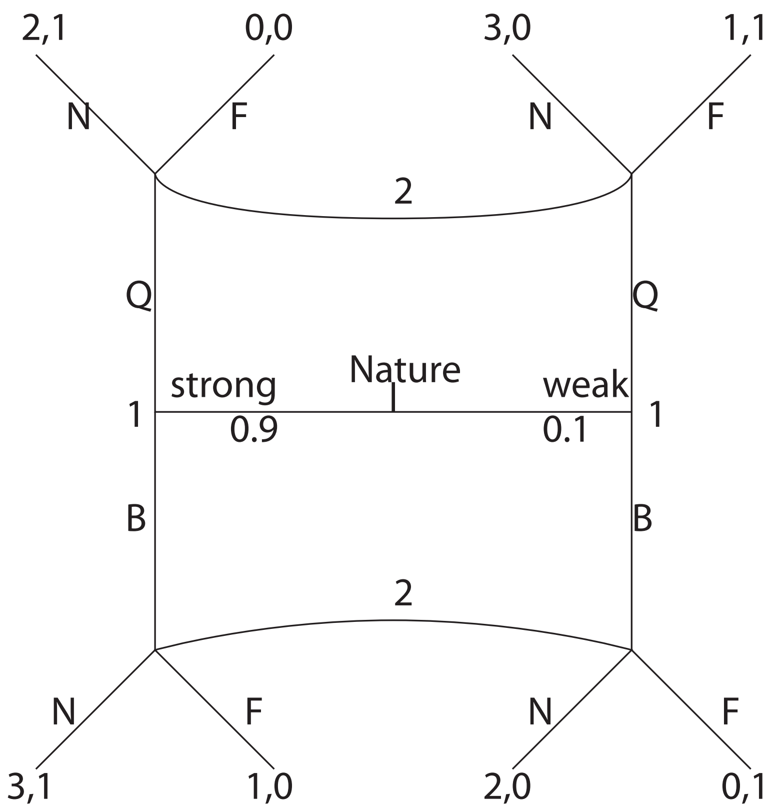
Because higher productivity is less costly to signal, a separating equilibrium always exists.

Notice that education is not really productive in the model. (Get an MBA to signal that, since you can handle the work, you will be productive.) However, if you measure the productivity of workers, you might think that education is productive.

If the equilibrium is such that $e^H = \theta^H(\theta^H - \theta^L)$, what should an employer think about a worker that chooses a slightly lower education? Who is more likely to make such a deviation?

Refinements of Sequential Equilibrium

Sequential equilibrium places almost no restrictions on beliefs off the equilibrium path. However, there are many games in which certain beliefs off the equilibrium path should be rejected. For example, no one should believe that a player is choosing a strictly dominated action. The following example cannot make an argument based on dominated strategies, but one of the sequential equilibria is less plausible than the other.



Beer and Quiche

The title of this game is based on the book in the 1980's, "Real Men Don't Eat Quiche."

There are two classes of sequential equilibrium.

Eq. 1: Strong and weak types of player 1 choose B. Player 2 fights if he observes Q, but not if he observes B. Beliefs: Player 2 uses Bayes' rule if he observes B, and believes that player 1 is at least as likely to be weak as strong if he observes Q.

Given the beliefs, player 2's strategy is sequentially rational. The beliefs are consistent. Given player 2's strategy, player 1's strategy is seqentially rational. A weak player 1 receives a payoff of 2, but deviating to Q would give him a payoff of 1, due to the fact that player 2 is prepared to fight.

Eq. 2: Strong and weak types of player 1 choose Q. Player 2 fights if he observes B, but not if he observes Q. Beliefs: Player 2 uses Bayes' rule if he observes Q, and believes that player 1 is at least as likely to be weak as strong if he observes B.

Given the beliefs, player 2's strategy is sequentially rational. The beliefs are consistent. Given player 2's strategy, player 1's strategy is seqentially rational. A strong player 1 receives a payoff of 2, but deviating to B would give him a payoff of 1, due to the fact that player 2 is prepared to fight.

Here is the problem: consistency does not take into account the fact that a weak player 1 is receiving a payoff of 3, and would never consider a deviation to B. A strong player 1 could consider B, hoping that player 2 chooses not to fight. Thus, player 2 should infer that anyone deviating to B must be strong, and not fight. This logic invalidates this sequential equilibrium.

The **Intuitive Criterion** is a refinement of sequential equilibrium that formalizes this logic. See Cho and Kreps (QJE 1987).

Given an equilibrium, let $S(m)$ denote the set of types who prefer the equilibrium to the out-of-equilibrium message m, for any beliefs of player 2 and any best response by player 2. [Types in $S(m)$ are ruled out.]

If, for some out-of-equilibrium message m, there is a type who prefers to send the message m, given that player 2 rules out types in $S(m)$ and chooses a best response accordingly, then the equilibrium outcome fails the intuitive criterion.

Trembling Hand Perfect Equilibrium

The notion of trembling hand perfection is that players think that other players make mistakes with small probability. Actions must be optimal given the equilibrium and also given the perturbed belief that allows for mistakes. With sequential equilibrium, the epsilon perturbations are only used in defining consistent beliefs, not sequential rationality.

The concept of perfection can be applied to strategic games as well as extensive games, and the results are slightly different.

Definition 248.1: A **trembling hand perfect equilibrium** of a finite strategic game is a mixed strategy profile σ with the property that there exists a sequence $(\sigma^k)_{k=0}^{\infty}$ of completely mixed strategy profiles that converges to σ, such that for each player i, σ_i is a best response to σ^k_{-i} for all k.

It follows by continuity that any trembling hand perfect equilibrium is a Nash equilibrium.

$0,0$	$0,0$	$0,0$
$0,0$	$1,1$	$2,0$
$0,0$	$0,2$	$2,2$

For this game, (T,L) and (B,R) are Nash but not perfect.

For two player strategic games, a profile is trembling hand perfect if and only if it is a mixed strategy NE, and neither player's strategy is weakly dominated.

Directly extending the definition of trembling hand perfection to extensive games has the problem that players do not consider the possibility that they will tremble in the future, which allows for trembling hand perfect equilibria that are not subgame perfect.

	L	R
A, a	1, 1	1, 1
A, b	1, 1	1, 1
B, a	0, 2	2, 0
B, b	0, 2	3, 3

$((A, a), L)$ is trembling hand perfect, because player 2 can think that player 1's most likely mistake is (B, a).

However, the unique subgame perfect equilibrium of the extensive game is $((B, b), R)$.

Definition 251.1: A trembling hand perfect equilibrium of a finite extensive game is a behavioral strategy profile that corresponds to a trembling hand perfect equilibrium of the agent strategic form of the game.

Note: In the agent strategic form, there is a separate player for each information set, with all agents of a given player having the same payoffs.

Proposition 251.2: For every trembling hand perfect equilibrium β of a finite extensive game with perfect recall, there is a belief system μ such that (β, μ) is a sequential equilibrium of the game.

Note: The converse is not true, since sequential equilibrium has no bite for games with a single, simultaneous move. However, the converse is true generically. Also, since a trembling hand perfect equilibrium exists, a sequential equilibrium exists as well.

The previous example showed a strategy profile that is perfect according to the strategic form, but not according to the extensive form. Here is an example of a strategy profile that is perfect according to the extensive form, but not according to the strategic form.

	L	R
L, a	$0, 0$	$1, 1$
L, b	$0, 0$	$1, 1$
R, a	$0, 0$	$0, 0$
R, b	$1, 1$	$1, 1$

$((L, b), R)$ is perfect according to the extensive form, because the initial player 1 can believe that he is more likely to tremble later than player 2. This cannot be perfect according to the strategic form, because player 1 is using a weakly dominated strategy.

Perturbations of the Game

Rather than perturbing strategies, a different robustness test on equilibrium is to perturb the game itself.

In the chain store game, if we introduce a small probability that the chain store is a "crazy" type that will always fight, then there are perfect Bayesian equilibria in which even the normal chain store is willing to fight for most of the game.

This is about perturbing the game, not about introducing irrationality. The crazy type might be completely rational, but simply receive a higher utility from fighting.

The following example, from Bagwell (GEB 1995), shows that the first mover advantage is completely eliminated when we introduce a small amount of noise in player 2's observation of player 1's move. The following matrix shows the payoffs as a function of the actions of the two players. Player 2 has two available actions, but four strategies, since her action can be a function of her obserrvation of player 1's action.

	S	C
S	5, 2	3, 1
C	6, 3	4, 4

There is a unique subgame perfect equilibrium, in which player 1 chooses S, and player 2 chooses S if S is observed, C if C is observed. The outcome is (S,S) [for Stackelberg], and payoffs are (5,2).

On the other hand, if the game is simultaneous move, the unique Nash equilibrium is (C,C) [for Cournot], and the payoffs are (4,4). Thus, there is a first mover advantage.

Now consider the noisy-leader game, where player 2 receives a signal, $\phi \in \{S, C\}$, such that

$$pr(\phi = S \mid S) = 1 - \varepsilon = pr(\phi = C \mid C).$$

Player 1 chooses an action, $a_1 \in \{S, C\}$

Player 2 chooses a function, $\omega(\phi)$, which can be written as $(\omega(S), \omega(C))$.

Consider a pure strategy Nash equilibrium of the noisy-leader game, $(a_1, (\omega(S), \omega(C))$. No matter what the signal is, $\omega(\phi)$ must be a best response to a_1. Thus, player 2 always plays the same action $a_2 = \omega(S) = \omega(C)$. Anticipating that player 2 will always play the same action, a_1 must be a best response to a_2. Therefore, the actions chosen must form a Nash equilibrium of the simultaneous move game, and all of the first mover advantage is lost.

This argument is completely general, as long as no off-equilibrium-path information sets can exist (common support assumption).

For fixed ε, no matter how small, this is trembling hand perfect.

What is going on? For positive ε, no matter how small, player 2 chooses a best response to player 1's strategy, and ignores her signal. If we suppose there can be an equilibrium with player 1 choosing S, player 2 must be playing S, but then player 1 can deviate to C; player 2 will think it is a "tremble" by nature and not a deviation. There is a discontinuity in the limit, where a deviation is recognized as such, allowing the Stackelberg outcome.

But when ε is small, the assumption of common knowledge of rationality is really being tested!

With mixed strategies, the first-mover advantage comes back. Player 1 chooses S with probability $1 - \varepsilon$ and C with probability ε. When C is observed, the small chance of an incorrect signal is balanced by the small chance that player 1 chose an unlikely action. Player 2 responds to S with S, but mixes in response to C.