Part I

Foundations

Part II

Bargaining

Part III

Auctions and Mechanism Design

Part IV

Matching

Part V

Information and Dynamic Games

Chapter 9

Perfect Bayesian Equilibrium

9.1 From SPE to PBE

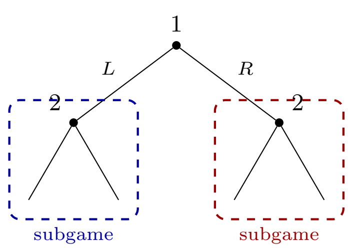
Subgame-perfect equilibrium (Chapter 2) refines Nash equilibrium by demanding sequentially rational play in every *subgame*. The definition is powerful in extensive-form games of **perfect information**, where every singleton history initiates a subgame and SPE is verified by backward induction at each node. It is also adequate for extensive-form games of *imperfect* information whose information sets have a tree-like structure that respects subgame boundaries (e.g., simultaneous-move stages embedded in a larger sequential game).

It loses bite as soon as informational nontrivialities cut across the would-be subgames. Consider a game in which player 1 moves first, player 2 then chooses without observing player 1's move, and the resulting branches are linked by a single information set spanning multiple post-move-1 nodes. The candidate "subgame" rooted at any single post-move-1 node is not a subgame in the formal sense, because it severs an information set. Backward induction has nowhere to start. SPE then either coincides with NE (when there is only one trivially-defined subgame, namely the whole game) or rules out very little.

The contrast between the two cases is best seen visually:

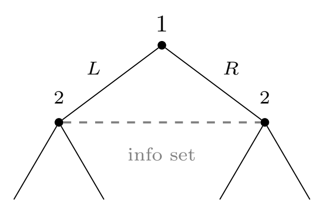
Perfect information

Two proper subgames; SPE applies



Imperfect information

No proper subgame; SPE has no bite



In the left tree, each player-2 node initiates its own proper subgame, and SPE pins down player 2's optimal action separately in each—backward induction works. In the right tree, the information set spanning both player-2 nodes makes the candidate "subgame" rooted at either one ill-defined: severing it would cut the information set in half. To evaluate optimality at the information set, player 2 needs to know how likely she is to be at the left versus the right node—a probability distribution over the nodes inside the information set. SPE has no language for that; PBE does.

The remedy is to add *beliefs* to the equilibrium concept. At every information set the relevant player should hold a probability distribution over the nodes inside that information set, evaluate continuation play under that distribution, and respond optimally. The resulting solution concept is **perfect Bayesian equilibrium (PBE)**: an extension of SPE to extensive-form games *without* well-behaved subgame structure.

Remark (Beliefs as "Why Is This Player Doing This?").

> All of the equilibrium concepts up to this chapter—NE, SPE, even sequential rationality in finite games of perfect information—make no formal use of beliefs. They speak only of strategies and best replies. But in extensive-form games, asking *why* a player chooses a particular action is often more illuminating than verifying that the action is a best reply: "because she thinks the state of the world is X with probability p." She still plays a best response—but a best response to her belief, not to the actual (unobserved) state. PBE makes that intuitive "thinks" a primitive of the equilibrium. Strategies tell you what each player does; beliefs tell you what each player conjectures about the unobserved nodes of the tree. The first novelty of PBE is to put both objects at the same level of formality.

Remark (Why Beliefs Are the Right Object to Add).

> In an SPE, the absence of off-path play is what makes "backward induction at every subgame" coherent: the subgame partition cleanly separates "what happens here" from "what happens elsewhere." When information sets cross those boundaries, evaluating "what happens here" from the inside requires knowing *which node inside the information set is in fact the current one*—and the only thing capable of summarizing that knowledge is a probability distribution over the nodes. PBE is what you get when you treat that distribution as a primitive of the equilibrium and impose two requirements: (i) actions are best replies given the beliefs, (ii) the beliefs are consistent with the strategies via Bayes' rule wherever Bayes' rule applies.

9.2 Definition: Strategies and Beliefs

A PBE is a pair (β, μ) in which $\beta = (\beta_i)_i$ is a profile of behavioral strategies and $\mu = (\mu_i)_i$ is a profile of belief systems, one per player.

Definition 9.1: Belief System

Fix a player i and an information set I_i of player i. A **belief** of player i at I_i is a probability distribution $\mu_i(I_i) \in \Delta(I_i)$ over the decision nodes in I_i. A **belief system** for player i is the collection $\mu_i = \{\mu_i(I_i)\}_{I_i \in \mathcal{I}_i}$ of such distributions, one for every information set of player i.

Definition 9.2: Perfect Bayesian Equilibrium

A pair (β, μ) is a **perfect Bayesian equilibrium** if

1. *Sequential rationality.* For every player i and every information set I_i, the action $\beta_i(I_i)$ that the strategy prescribes at I_i maximizes i's conditional expected payoff at I_i, with the conditional expectation taken under the belief $\mu_i(I_i)$:

$$\beta_i(I_i) \in \arg\max_{a \in A_i(I_i)} \sum_{n \in I_i} \mu_i(I_i)(n) \cdot U_i\big(a, \beta_{-i};\, n\big),$$

   where $U_i(a, \beta_{-i}; n)$ denotes i's continuation payoff from node n if i plays a at I_i and the others play β_{-i} thereafter.

2. *Bayesian consistency.* Given β, the belief $\mu_i(I_i)$ is derived from Bayes' rule whenever Bayes' rule applies—that is, whenever the strategy profile reaches I_i with positive probability.

Remark (Conditional, Not Unconditional).

The expectation in (1) is the *conditional* expected payoff at I_i, not the unconditional ex-ante one. The conditioning on "having reached I_i" is encoded in the support of $\mu_i(I_i)$: by definition $\mu_i(I_i) \in \Delta(I_i)$ is a posterior *over* I_i, so summing $\mu_i(I_i)(n) \cdot U_i(a, \beta_{-i}; n)$ over $n \in I_i$ already restricts attention to the event "I_i has been reached." Two practical consequences. First, multiplying by the (positive) probability of reaching I_i would only rescale the objective by a constant and not change the argmax—so writing the conditioning explicitly outside, e.g. $\mathbb{E}[\cdot \mid I_i]$, is redundant if the belief is already a posterior. Second, this is exactly what makes sequential rationality *local*: each information set is evaluated on its own terms, with continuation payoffs $U_i(\cdot;\, n)$ that depend only on what comes after n and not on how the play got to n.

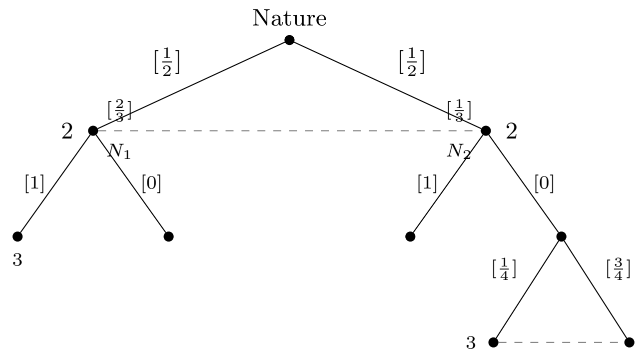
The first condition is sequential rationality at every information set; it is what replaces "optimality on every subgame" from the SPE definition. The second condition is the new piece: it prevents free choice of beliefs at on-path information sets but—deliberately—leaves beliefs at zero-probability information sets unrestricted by Bayes' rule alone. This loophole is what generates the multiplicity of PBE one observes in signaling games (Section 12.3) and motivates further refinements (intuitive criterion, divinity, sequential equilibrium).

Remark ("Whenever Possible"—the Source of Multiplicity).

The phrase *whenever possible* in (2) is doing real work. Bayes' rule pins down the posterior $P(\text{node} \mid I_i)$ only when the prior probability of reaching I_i under β is positive. At an information set that the strategy profile reaches with probability zero (a so-called *off-path* information set), Bayes' rule is silent—$0/0$—and the equilibrium concept allows *any* belief there. Equilibria are then often distinguished only by what beliefs are postulated off the equilibrium path, and these off-path beliefs are exactly what determines whether a candidate deviation is profitable. The discipline of refinements like the intuitive criterion is essentially a proposal for how to tame this freedom.

A worked tree

Consider the extensive-form fragment below. Nature first selects one of two states with equal probability (the initial $\left[\frac{1}{2}\right]$ labels). Player 2's information set spans both post-Nature nodes, so 2 does not directly observe the state; the labels $\left[\frac{2}{3}\right]$ and $\left[\frac{1}{3}\right]$ are 2's belief at that information set. After 2 moves, on one branch a payoff is reached; on the other, player 3 moves at an information set spanned by two nodes with belief $\left[\frac{1}{4}\right]$, $\left[\frac{3}{4}\right]$.



The numbers in brackets along edges are conditional probabilities induced by β; the numbers in brackets next to the dashed information sets are the player's beliefs μ. Bayesian consistency forces an arithmetic relationship between the two. For instance, write N_1 and N_2 for the two nodes in player 2's information set (the left- and right-hand children of Nature, as marked on the diagram). Then 2's belief at that information set must satisfy

$$\mu_2(N_1) \;=\; \frac{\Pr(\text{Nature picks left})\cdot 1}{\Pr(\text{Nature picks left})\cdot 1 + \Pr(\text{Nature picks right})\cdot 1} \;=\; \tfrac{1}{2}$$

under the prior $\left(\frac{1}{2}, \frac{1}{2}\right)$. Any other belief there (such as $\left[\frac{2}{3}, \frac{1}{3}\right]$) would be inconsistent with the prior unless something earlier in β *biased* the path—and that something would itself need to be computed and reconciled with Bayes' rule. The board's specific numbers were illustrative of how to read the diagram, not a uniquely-solved equilibrium.

9.3 Spence Signaling

The leading workhorse of PBE is the **Spence (1973) signaling model** of education and labor markets. It crystallizes the idea that a costly action with no direct productive value can nevertheless be informative about hidden type, and—more cautiously—that the resulting equilibria come in continua, none of which is selected by PBE alone.

9.3.1 Setup

A worker has private **type** $\theta \in \{L, H\}$ with $0 < L < H$ (so θ doubles as the worker's productivity: a type-θ worker produces output worth exactly θ, and L, H are two numerical productivity levels). The prior is

$$\Pr(\theta = H) = \pi, \qquad \Pr(\theta = L) = 1 - \pi, \qquad \pi \in (0,1).$$

The worker first chooses an **education level** $e \in [0, \infty)$. Two competitive firms then observe e and simultaneously offer wages; Bertrand competition forces both wage offers up to the firm's expected productivity given e, so the realized wage is

$$w(e) \;=\; \mathbb{E}[\theta \mid e] \;=\; \mu(H \mid e) \cdot H + \big(1 - \mu(H \mid e)\big) \cdot L,$$

where $\mu(H \mid e)$ is the firms' posterior that the worker is high-type after observing the signal. The worker's payoff is

$$u(\theta, e, w) \;=\; w \;-\; \frac{e}{\theta}.$$

The cost of education is e/θ: high types find education cheaper. Critically, education *does not affect productivity*: θ enters firm payoffs directly, e only through firm beliefs. Education is a pure signal.

Remark (Why Not Just Say You're High-Type?).

> The setup carefully forbids the worker from declaring her type. If declarations were allowed and free, every worker—low or high—would announce "I am high-type," the announcement would carry no information, and firms would ignore it: this is the standard *cheap talk* non-result. The whole point of the model is that a verbal claim is unverifiable, so the worker must instead choose a costly action whose marginal cost depends on the unobserved type. The single-crossing structure e/θ guarantees that *some* levels of e are too painful for the low type to mimic but bearable for the high type, opening a channel through which type can be communicated.

Single-crossing. The cost function e/θ implements the **single-crossing property** (a.k.a. Spence-Mirrlees): for any two education levels $e' > e$ and any wage gap Δw that makes the low type just willing to switch from e to e', the high type strictly prefers to switch. This monotone comparative statics in type is what makes signaling possible at all—without it, no signal could separate types in equilibrium.

9.3.2 An Order Constraint in Every Equilibrium

Before classifying equilibria, we record a useful observation that holds at every PBE.

Proposition 9.3: Monotone Education in Type

In any PBE of the Spence signaling game, $e^H \geq e^L$.

Proof for Proposition.

> Let $w^\theta = \mathbb{E}[\theta \mid e^\theta]$ be the equilibrium wage offered to a worker who chooses e^θ. The two

incentive-compatibility constraints are

$$\begin{aligned} &(\text{IC for } H) \quad w^H - \tfrac{e^H}{H} \;\ge\; w^L - \tfrac{e^L}{H}, \\ &(\text{IC for } L) \quad w^L - \tfrac{e^L}{L} \;\ge\; w^H - \tfrac{e^H}{L}. \end{aligned}$$

Adding the two inequalities, the wage terms cancel and one obtains

$$\tfrac{e^L - e^H}{H} + \tfrac{e^H - e^L}{L} \;\ge\; 0 \quad \Longleftrightarrow \quad (e^H - e^L)\big(\tfrac{1}{L} - \tfrac{1}{H}\big) \;\ge\; 0.$$

Since $\frac{1}{L} - \frac{1}{H} > 0$, we conclude $e^H \ge e^L$.

The proof is general: it nowhere uses the specific values of w^H, w^L beyond the IC inequalities. It illustrates a recurring trick in the analysis of separating-type games—add the two ICs and watch the wage terms drop out, leaving an inequality among the actions and the type-sensitivity of cost.

9.3.3 Separating Equilibria

A **separating equilibrium** is one in which $e^H \neq e^L$, so that θ is fully revealed by the signal. The proposition forces $e^H > e^L$ in such equilibria.

Two structural features. First, $e^L = 0$ in any separating equilibrium: if the low type is identified by his signal, his wage is L regardless of e, and any $e^L > 0$ is strictly dominated by $e = 0$. Second, the high type's signal $e^H = e^*$ must satisfy both ICs simultaneously; since separation forces $w^H = H, w^L = L$, the ICs become

$$\begin{aligned} &(\text{IC for } H) \quad H - \tfrac{e^*}{H} \;\ge\; L - 0 \;=\; L, && \Leftrightarrow && e^* \;\le\; H(H-L), \\ &(\text{IC for } L) \quad L - 0 \;\ge\; H - \tfrac{e^*}{L}, && \Leftrightarrow && e^* \;\ge\; L(H-L). \end{aligned}$$

A note on why the low-type IC must be checked at all, given that we just argued $e^L = 0$. The argument $e^L = 0$ said the low type, *conditional on being identified as low* (so that wage is L regardless of e), optimally chooses no education. But the deviation IC for L rules out is a different one: the low type *pretending to be high* by paying e^* and pocketing wage H. Whether this masquerade is profitable depends on the cost-of-education-to-low-type, e^*/L, against the wage gain $H - L$. The lower bound $e^* \ge L(H-L)$ is precisely the condition that makes the masquerade not pay.

Combining, $e^* \in [L(H-L),\ H(H-L)]$. The lower bound says the signal must be costly enough that the low type would not pay it for the wage gain $H - L$; the upper bound says it must not be *so* costly that even the high type would prefer to mimic the low type.

Proposition 9.4: Separating PBE: A Continuum

For every $e^* \in \big[L(H-L),\ H(H-L)\big]$ there exists a PBE in which $e^H = e^*$, $e^L = 0$, for some choice of off-path beliefs (constructed in the proof below). Hence the Spence signaling game admits a *continuum* of separating equilibria.

Proof for Proposition.

The strategy profile is $e^L = 0,\ e^H = e^*$. The belief system specifies, on-path, $\mu(H \mid 0) = 0$ and $\mu(H \mid e^*) = 1$, generating wages L and H respectively. Off-path beliefs can be chosen to discourage any deviation; the simplest choice is the **pessimistic** specification

$$\mu(H \mid e) \ = \ \begin{cases} 1 & e = e^*, \\ 0 & e \neq e^*, \end{cases}$$

i.e., any out-of-equilibrium signal is read as "low type."

- *Low type's IC.* If L stays at 0, payoff is $L - 0 = L$. If L deviates to any $e' \neq 0$ with $e' \neq e^*$, payoff is $L - e'/L < L$, strictly worse. The only deviation worth examining is $e' = e^*$ (which would give wage H); this is unprofitable iff $L \geq H - e^*/L$, i.e. $e^* \geq L(H - L)$.
- *High type's IC.* If H stays at e^*, payoff is $H - e^*/H$. The best deviation under the pessimistic belief gives wage L, so the most profitable deviation is to $e' = 0$ (cheapest), with payoff L. No-deviation requires $H - e^*/H \geq L$, i.e. $e^* \leq H(H - L)$.

Both ICs hold for $e^* \in [L(H - L),\ H(H - L)]$. Sequential rationality of the firms is automatic given Bertrand competition and the posited beliefs.

Remark (Why a Continuum?).

The interval $[L(H-L), H(H-L)]$ has positive length whenever $L < H$, so there is always more than one separating e^*. The reason is that the on-path beliefs $\mu(H \mid e^*) = 1, \mu(H \mid 0) = 0$ pin down only *two* of the firm's posteriors; the remaining beliefs at every other education level are free, and the equilibrium concept lets us use that freedom to deter deviations. Each e^* in the interval is supported by its own pessimistic specification: the recipe is uniform ("any $e \notin \{0, e^*\}$ is read as low type"), but the location of the on-path spike at e^* shifts across equilibria. So all members of the family share a single rule for how to interpret deviations; they differ only in which e^* the rule targets as the high-type signal.

9.3.4 Pooling Equilibria

A **pooling equilibrium** is one in which both types choose the same signal: $e^L = e^H = \bar{e}$ for some common $\bar{e} \geq 0$. The firm cannot extract any information from $\bar{e}$ alone, so on-path beliefs equal the prior:

$$\mu(H \mid \bar{e}) \ = \ \pi, \qquad w(\bar{e}) \ = \ \pi H + (1 - \pi)L.$$

Off-path beliefs are not pinned down by Bayes' rule. We adopt the standard **pessimistic specification**

$$\mu(H \mid e) \ = \ 0 \quad \text{for every } e \neq \bar{e},$$

under which any deviation is read as "low type" and rewarded with wage L. We show below that this is the harshest off-path belief, hence supports the largest set of pooling equilibria;

weaker off-path specifications collapse most of them.

Identifying the binding deviation. A worker contemplating a deviation from $\bar{e}$ to some $e' \neq \bar{e}$ obtains wage L regardless of e' (under pessimistic beliefs), so her deviation payoff is $L - e'/\theta$. Since this is decreasing in e', the most attractive deviation is the cheapest off-path signal: $e' = 0$. Both types' IC therefore reduces to a single inequality each, comparing the equilibrium payoff at $\bar{e}$ to the deviation payoff at $e' = 0$.

Low-type IC. The low type must (weakly) prefer $\bar{e}$ to deviating to $e' = 0$:

$$\underbrace{\pi H + (1-\pi)L - \tfrac{\bar{e}}{L}}_{\text{equilibrium payoff } \bar{U}_L} \geq \underbrace{L - 0}_{\text{deviation payoff}} = L.$$

Rearranging,

$$\pi H + (1-\pi)L - L \geq \tfrac{\bar{e}}{L} \iff \pi(H-L) \geq \tfrac{\bar{e}}{L} \iff \bar{e} \leq \pi L(H-L).$$

High-type IC. The high type must (weakly) prefer $\bar{e}$ to deviating to $e' = 0$:

$$\pi H + (1-\pi)L - \tfrac{\bar{e}}{H} \geq L.$$

Rearranging,

$$\bar{e} \leq \pi H(H-L).$$

This bound is strictly weaker than the low-type bound, since $\pi H(H-L) > \pi L(H-L)$. So the high-type IC is automatically satisfied whenever the low-type IC is. Geometrically: education is cheaper per unit for the high type ($1/H < 1/L$), so the same on-path payoff requires a smaller utility cost from $\bar{e}$ for H, giving more slack.

Sequential rationality of the firm. Two competitive firms observe e and bid up to expected productivity. On path, expected productivity given the prior π is exactly $w(\bar{e}) = \pi H + (1-\pi)L$. Off path under the pessimistic belief, expected productivity is L. Both wages are best responses, so the firm's strategy is sequentially rational.

Combining the three conditions:

Proposition 9.5: Pooling PBE: A Continuum

For every $\bar{e} \in [0,\ \pi L(H-L)]$ there exists a pure-strategy PBE in which both types choose $e = \bar{e}$, the on-path belief is $\mu(H \mid \bar{e}) = \pi$, the off-path belief is $\mu(H \mid e) = 0$ for $e \neq \bar{e}$, and firms offer the wage $\pi H + (1-\pi)L$ at $\bar{e}$ and the wage L at every other signal.

Proof for Proposition.

Sequential rationality of the firm follows from Bertrand competition together with the on-path posterior π and the off-path pessimistic posterior 0. Both types' IC against the cheapest deviation $e' = 0$ reduce to $\bar{e} \leq \pi L(H-L)$ (low type) and $\bar{e} \leq \pi H(H-L)$ (high type); the first implies the second. ■

The boundary case $\bar{e} = 0$ is the trivial pooling equilibrium in which neither type acquires education and the firm offers the prior wage. The upper boundary $\bar{e} = \pi L(H - L)$ is the most education-intensive pooling equilibrium consistent with the low type's IC; any larger $\bar{e}$ would induce L to drop down to $e' = 0$, breaking the equilibrium.

Why pessimistic off-path beliefs? If we instead specified the off-path belief as $\mu(H \mid e) = 1$ for some $e \neq \bar{e}$ (say $e' < \bar{e}$), then a low type weighing $\bar{e}$ vs. e' would compare $\pi H + (1 - \pi)L - \bar{e}/L$ to $H - e'/L$. Since $H > \pi H + (1 - \pi)L$ and $e' < \bar{e}$, deviation strictly dominates whenever the off-path optimistic wage is offered cheaply enough—so the pooling equilibrium unravels. The pessimistic specification is the unique off-path belief structure under which no deviation looks more attractive than the equilibrium signal. This is the same observation that motivates the intuitive-criterion analysis in §9.3.6 below: the IC objects precisely to the pessimistic belief whenever the deviation is unattractive for L but attractive for H.

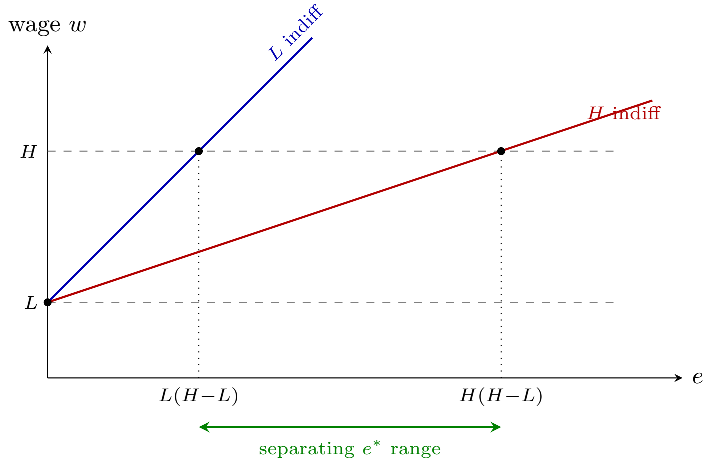
Remark (The Two Roles of Incentive Compatibility).

PBE asks two things: that strategies be optimal given beliefs (sequential rationality), and that beliefs be Bayes-consistent given strategies—wherever Bayes applies. The qualifier is what gives PBE its slack: Bayes pins down on-path beliefs, but off-path beliefs are unconstrained by Bayes alone. As a result, IC ends up doing two distinct jobs.

- *On-path discipline.* IC rules out deviations from one on-path action to another on-path action. At on-path information sets, posteriors are determined by Bayes, and IC must hold against those posteriors. There is no freedom here—this is a necessary condition.
- *Off-path discipline as a design choice.* IC also rules out deviations from on-path actions to off-path actions. But here whether a deviation pays depends on what the firm *thinks* when it sees the off-path signal, and that thought is not pinned down by Bayes. The analyst must select off-path beliefs strategically: pick beliefs harsh enough that every candidate off-path deviation looks unprofitable. The pessimistic specification ("any unexpected signal is a low type") is the harshest possible choice and therefore sustains the largest set of equilibria.

Under more permissive off-path beliefs (e.g., "any unexpected signal is a high type"), most of these equilibria collapse: a low type would mimic the high type's signal, the firm would interpret it as high, and the equilibrium would unravel. The off-path belief is exactly where the multiplicity lives—and where refinements like the intuitive criterion or divinity intervene to discipline it.

The trade-off underlying separating and pooling equilibria can be visualized in (e, w)-space using each type's indifference curves through the no-education / low-wage anchor $(0, L)$:



Reading the figure. Each type's indifference curve from $(0, L)$ has slope $1/\theta$: the low type's curve (slope $1/L$, blue) is steeper than the high type's (slope $1/H$, red) because education is more costly per unit for the low type. A separating equilibrium with high-type signal e^* requires the firm-offered wage H at $e = e^*$ to lie *below* the low type's indifference curve through $(0, L)$ (so the low type does not mimic) and *above* the high type's indifference curve through $(0, L)$ (so the high type prefers signaling to staying at $e = 0$). The wage H touches the low-type curve at $e = L(H - L)$ and the high-type curve at $e = H(H - L)$, giving the interval $[L(H - L), H(H - L)]$ as the range of e^* that supports separation. A pooling equilibrium at $\bar{e} \in (0, \pi L(H - L)]$ analogously sits where the off-path wage L at $e = 0$ and the on-path wage $\pi H + (1 - \pi)L$ at $\bar{e}$ leave the low type weakly content—geometrically the same kind of inequality, with π replacing the 1 that enters the separating bound.

9.3.5 Partial Pooling and the Multiplicity Problem

Beyond pure separation and pure pooling, the model also admits **semi-separating** (partial-pooling) equilibria, in which one type randomizes across two signals. For example, H randomizes between e' and e'' while L plays e'' for sure; firms then assign different posteriors to e' (pure H) and e'' (mixture). Each such construction generates an additional family of PBE indexed by the mixing probability and the off-path beliefs.

The cumulative picture, then, is uncomfortable: the Spence model has *tons* of perfect Bayesian equilibria. Separating and pooling alone yield two continua; partial pooling adds further families. PBE, the natural extension of SPE to incomplete-information games, is too permissive to deliver sharp predictions in this canonical example.

Remark (Why So Many Equilibria, and What to Do About It).

> The proliferation comes from two sources. First, off-path beliefs are unrestricted by Bayes' rule, so different choices generate different equilibria. Second, even on-path the IC constraints in the separating case define a closed interval, not a single point. The literature responds with *refinements*: the intuitive criterion (Cho-Kreps 1987), divinity (Banks-Sobel 1987), undefeated equilibrium (Mailath-Okuno-Fudenberg-Postlewaite 1993), and others. Each restricts off-path beliefs by an additional logical principle ("a deviation that only one type would ever consider should be attributed to that type") and

selects subsets of the PBE set. In the next subsection we develop the intuitive criterion in full, showing that it selects the unique *Riley outcome*—the least-cost separating equilibrium, $e^L = 0,\ e^H = L(H - L)$. The lesson is general: PBE is the right concept for incomplete-information dynamics, but it usually needs a refinement to deliver predictive bite.

9.3.6 Refinement: The Intuitive Criterion

The intuitive criterion of Cho and Kreps (1987) is a forward-induction restriction on off-path beliefs. Its conclusion in the Spence model is sharp: among the continua of separating, pooling, and partial-pooling PBE, exactly one outcome survives—the least-cost separating equilibrium $(e^L, e^H) = (0, L(H - L))$. This subsection states the criterion and proves the selection result.

Motivation: a forward-induction speech. Fix any PBE and consider an off-path signal e'. Suppose a high-type worker contemplates choosing e' and accompanies the choice with the following speech to the firms:

> "I am the high type. Here is why you should believe me. If you grant me the most charitable belief and pay me wage H at e', my payoff $H - e'/H$ exceeds my equilibrium payoff. So I have a reason to send e'. The same most charitable belief, paying H at e', gives a low type only $H - e'/L$, which is strictly below his equilibrium payoff. Even under the most generous interpretation, the low type would never want to send e'. So your belief at e' should put zero weight on L."

The intuitive criterion is precisely this argument formalized. It rules out off-path beliefs that attribute a deviation to a type for whom the deviation is unprofitable under *every* possible belief, when some other type could plausibly benefit.

Definition 9.6: Equilibrium-Dominated Signal

Fix a PBE of the signaling game and let $\bar{U}_\theta$ denote type θ's equilibrium payoff. A non-equilibrium signal e' is **equilibrium-dominated for type** θ if

$$\max_{w \in [L,H]} \big[\, w - e'/\theta \,\big] \;<\; \bar{U}_\theta,$$

i.e., even granting the most favorable wage $w = H$, type θ's payoff at e' falls strictly short of her equilibrium payoff. Equivalently, $H - e'/\theta < \bar{U}_\theta$.

Intuitive Criterion (Cho-Kreps, 1987)

A PBE satisfies the **intuitive criterion** if, for every non-equilibrium signal e' and every pair of types θ, θ',

$$e' \text{ is equilibrium-dominated for } \theta \text{ but not for } \theta' \quad\Longrightarrow\quad \mu(\theta \mid e') \;=\; 0.$$

When e' is dominated for all types or for none, the criterion places no restriction.

The criterion bites only when at least one type strictly cannot benefit from the deviation under any belief while at least one other type might. The off-path belief is then required to put zero weight on the unambiguous loser.

We now apply the criterion to the Spence model and obtain the uniqueness result through three claims.

Proposition 9.7: No Pooling Equilibrium Satisfies the Intuitive Criterion

Every pooling equilibrium of the Spence signaling game (i.e., $\bar{e} \in [0, \pi L(H-L)]$) violates the intuitive criterion.

Proof for Proposition.

In a pooling equilibrium at $\bar{e}$ the wage on path is $w(\bar{e}) = \pi H + (1-\pi)L$, and the equilibrium payoffs are

$$\bar{U}_L \;=\; \pi H + (1-\pi)L - \bar{e}/L, \qquad \bar{U}_H \;=\; \pi H + (1-\pi)L - \bar{e}/H.$$

We show that there exists an off-path signal $e' > \bar{e}$ that is equilibrium-dominated for L but not for H.

Such an e' exists iff $H - e'/L < \bar{U}_L$ and $H - e'/H > \bar{U}_H$ both hold, i.e., iff

$$L\,(H - \bar{U}_L) \;<\; e' \;<\; H\,(H - \bar{U}_H).$$

Substituting the equilibrium payoffs and simplifying,

$$L\,(H - \bar{U}_L) = L\big[(1-\pi)(H-L)\big] + \bar{e},$$
$$H\,(H - \bar{U}_H) = H\big[(1-\pi)(H-L)\big] + \bar{e}.$$

The interval is therefore $\big(L(1-\pi)(H-L)+\bar{e},\; H(1-\pi)(H-L)+\bar{e}\big)$, which is non-empty since $L < H$. Pick any e' in this interval.

Then e' is equilibrium-dominated for L (the LHS inequality) but not for H (the RHS). The intuitive criterion demands $\mu(H \mid e') = 1$, which contradicts the off-path belief $\mu(H \mid e') = 0$ that supports the pooling equilibrium. The pooling equilibrium fails the criterion. ∎

Proposition 9.8: No Separating Equilibrium with $e^H > L(H-L)$ Satisfies the Intuitive Criterion

Every separating equilibrium with $e^L = 0$ and $e^H \in \big(L(H-L),\; H(H-L)\big]$ violates the intuitive criterion.

Proof for Proposition.

In such an equilibrium the wage at e^H is H and the equilibrium payoffs are

$$\bar{U}_L \;=\; L, \qquad \bar{U}_H \;=\; H - e^H/H.$$

Pick any $e' \in \big(L(H-L),\ e^H\big)$. The interval is non-empty by hypothesis ($e^H > L(H-L)$).

- For L: $H - e'/L < H - L(H-L)/L = L = \bar{U}_L$. So e' is equilibrium-dominated for L.
- For H: $H - e'/H > H - e^H/H = \bar{U}_H$ (since $e' < e^H$). So e' is *not* equilibrium-dominated for H.

The intuitive criterion demands $\mu(H \mid e') = 1$. But the equilibrium supporting e^H relied on $\mu(H \mid e') = 0$ for $e' \neq 0, e^H$ to deter the high type from undercutting her own signal. Contradiction.

Proposition 9.9: The Riley Separating Equilibrium Satisfies the Intuitive Criterion

The separating equilibrium with $(e^L, e^H) = (0, L(H-L))$ satisfies the intuitive criterion.

Proof for Proposition.

In the Riley equilibrium $\bar{U}_L = L$ and $\bar{U}_H = H - L(H-L)/H$. Take any non-equilibrium signal $e' > 0$ with $e' \neq L(H-L)$; we check both cases.

Case $e' < L(H-L)$. Then

$$H - e'/L \ > \ H - L(H-L)/L \ = \ L \ = \ \bar{U}_L,$$

so e' is not equilibrium-dominated for L. Similarly, $H - e'/H > H - L(H-L)/H = \bar{U}_H$, so e' is not equilibrium-dominated for H. The criterion places no restriction.

Case $e' > L(H-L)$. Then

$$H - e'/L \ < \ L \ = \ \bar{U}_L \qquad \text{and} \qquad H - e'/H \ < \ H - L(H-L)/H \ = \ \bar{U}_H,$$

so e' is equilibrium-dominated for both types. The criterion places no restriction.

In neither case does the criterion demand a non-trivial belief, so the Riley equilibrium satisfies it (with any off-path belief, including the standard pessimistic $\mu(H \mid e') = 0$).

Theorem 9.10: Uniqueness Under the Intuitive Criterion

In the Spence signaling game, exactly one PBE outcome survives the intuitive criterion: the least-cost separating equilibrium with $e^L = 0$ and $e^H = L(H-L)$.

Remark (Existence and Wider Applicability).

Two practical points worth flagging.

Existence is not automatic. For a general signaling game, the intuitive criterion can in principle eliminate *all* PBE, leaving an empty set of refined equilibria. This does not happen in the Spence model: the proof of the third claim above is itself an existence proof—the Riley equilibrium is constructed and shown to satisfy the criterion. But in some games one needs to check existence separately, sometimes by passing to weaker refinements (divinity, $D1$, $D2$, ...).

Compatibility with sequential equilibrium. Cho and Kreps designed the intuitive criterion as an economic re-statement of more abstract stability concepts in the Kohlberg-Mertens (1986) "stable equilibrium" tradition. In generic signaling games, the intuitive criterion is consistent with sequential equilibrium: every IC-surviving PBE is a sequential equilibrium. The Mas-Colell-Whinston-Green textbook (Chapter 13) develops these ideas in greater detail.

What about partial pooling? The same argument that knocks out pooling also knocks out partial-pooling (semi-separating) equilibria. If the high type randomizes between e' and e'' while the low type concentrates on e'', then any signal slightly above the low-type's equilibrium can be shown to be equilibrium-dominated for L but not for H, producing the same contradiction. The intuitive criterion thus selects a single outcome out of a three-dimensional family of PBE.

Remark (Costly Signals Beyond Education: The Peacock's Tail).

An evolutionary biology analogue of the Spence model arose around the same time. A peacock's elaborate tail is metabolically expensive to grow and makes the bird more conspicuous to predators. It does not improve foraging or fighting ability. Why grow it? The Zahavi (1975) "handicap principle" answers: the tail is a costly, unfakeable signal of underlying genetic fitness. Only a peacock with a strong constitution can afford the energy expenditure and the predation risk; weak peacocks cannot mimic the display, so peahens that prefer big tails systematically end up with high-fitness mates. The structural parallel to Spence is exact: a productive but unobservable trait, a costly action whose marginal cost is decreasing in the trait, and a separating equilibrium in which the high type incurs the cost to communicate her type. Costly signaling is one of those rare unifying ideas that crossed independently into economics and biology in the 1970s and now turns up wherever an agent has private information she would like to credibly convey.

9.4 Reputation in the Finitely Repeated PD

The chain store paradox of Section 11.4 showed that a vanishingly small probability of a "crazy" commitment type can sustain reputational behavior in a sequential game between a long-lived player and short-lived opponents. A natural follow-up question, addressed by Kreps, Milgrom, Roberts, and Wilson (1982) in a companion paper, is whether the same resolution rescues cooperation in the **finitely repeated Prisoner's Dilemma** between two long-lived players—another canonical setting where backward induction predicts unraveling.

9.4.1 Setup

Two players play the Prisoner's Dilemma

	C	D
C	2, 2	−1, 3
D	3, −1	0, 0

for T periods, with periods numbered *backwards*: $t = 1$ is the last period, $t = T$ is the first. Each player is independently *rational* with probability $1 - \varepsilon$ and a *commitment type* (always plays grim trigger: cooperate while the opponent cooperates; switch to permanent defection after any defection) with probability ε. Players observe past actions but not the opponent's type. Let μ_t denote the probability a player ascribes to the *opponent* being a commitment type at the start of period t, conditional on the equilibrium history.

Remark (Why Backward Induction Fails Here).

> Recall from Section 11.3 that under common knowledge of payoffs the only SPE of the finitely repeated PD is (D, D) in every period. The reason is structural: the unique stage NE in the last period pins down (D, D) at $t = 1$, which makes the period-2 subgame strategically equivalent to the stage game, and so on. The reputation model breaks this by giving the rational player something at stake in maintaining ambiguity about her type: defecting reveals rationality (a commitment type would never defect against a cooperator), and revelation collapses the subsequent subgame to (D, D) forever. That continuation loss can outweigh the one-period gain from defecting, sustaining cooperation off the unraveling path.

9.4.2 Constructing the Equilibrium

We solve for a symmetric PBE in which both rational players cooperate with positive probability while both retain reputation, and randomize as the endgame approaches. Let

$$q_t = \Pr(\text{player cooperates in period } t \mid \text{both cooperated through } t+1)\,,$$

let μ_t be the belief about the opponent being commitment, and let p_t be the probability that a *rational* player cooperates in period t. Conditional on histories in which both players have cooperated through period $t + 1$, Bayesian aggregation gives

$$q_t \;=\; (1 - \mu_t)p_t + \mu_t,$$

since commitment types always cooperate so far. Inverting,

$$p_t \;=\; \frac{q_t - \mu_t}{1 - \mu_t}.$$

After any defection, the commitment type switches to permanent D, so observing D from a player who was on the cooperation path reveals rationality with certainty—beliefs collapse to $\mu = 0$ and the continuation reverts to mutual defection.

The construction proceeds in three phases:

- **Phase I** ($t > K + 1$): full cooperation, $q_t = 1$, $p_t = 1$.
- **Phase II** ($1 < t \le K + 1$): rational players randomize.

- **Phase III** ($t = 1$): defect outright.

The cutoff K is chosen as a function of ε. Specifically, fix the integer k satisfying

$$\left[\tfrac{1}{3}\right]^{k+1} \le \varepsilon < \left[\tfrac{1}{3}\right]^{k}, \quad \Leftrightarrow \quad \tfrac{1}{3} \le 3^k \varepsilon < 1,$$

and set $K = 2k$, so that $\frac{1}{3} \le 3^{K/2}\varepsilon < 1$.

9.4.3 Equilibrium Mixing Probabilities and Beliefs

The equilibrium specifies the on-path mixing weights

$$q_1 = q_3 = \cdots = q_{K+1} = 3^{K/2}\varepsilon, \qquad q_2 = q_4 = \cdots = q_K = \left[\tfrac{1}{3}\right]^{K/2+1} \frac{1}{\varepsilon},$$

and $q_t = 1$ for all $t > K + 1$. Notice that for $t \le K + 1$,

$$q_t \, q_{t+1} \; = \; \tfrac{1}{3},$$

which is the key recursion that makes randomization an equilibrium (derived below). On-path beliefs follow the Bayesian update

$$\mu_t \; = \; \begin{cases} \varepsilon & \text{if } t > K, \\ \dfrac{\varepsilon}{q_{K+1}\, q_K \, \cdots \, q_{t+1}} & \text{if } t \le K, \end{cases}$$

which simplifies to $\mu_t = \mu_{t+1}/q_{t+1}$ for any history of mutual cooperation.

9.4.4 Verification at the Last Period

Plug in $t = 1$:

$$\mu_1 \; = \; \frac{\varepsilon}{q_{K+1}\, q_K \cdots q_2} \; = \; 3^{K/2}\varepsilon \; = \; q_1.$$

At $t = 1$ the probability of cooperation by the opponent equals exactly the probability that the opponent is a commitment type. This means rational players defect with probability 1 in the last period (only commitment types cooperate), as one would expect from the unraveling logic.

9.4.5 Indifference in the Mixing Phase

Consider any period $1 < t \le K + 1$ on the cooperation path. We verify that a rational player is indifferent between cooperating and defecting.

Defecting today. Playing D when the opponent cooperates with probability q_t yields

$$V_t(D) \; = \; 3q_t + 0 \cdot (1 - q_t) \; = \; 3q_t,$$

since the opponent's commitment type, after observing D, switches to permanent D, so all continuation payoffs are zero.

Cooperating today. Playing C when the opponent cooperates with probability q_t yields a complicated continuation, but the structure of the equilibrium ensures the simple expression

$$V_t(C) \;=\; (2+3q_{t-1})q_t - 1\cdot(1-q_t).$$

The decomposition: with probability q_t the opponent cooperates today (instantaneous payoff 2), and tomorrow the rational player will defect against a still-cooperating opponent (yielding $3q_{t-1}$ in $t-1$); with probability $1-q_t$ the opponent defects today (instantaneous loss -1), revealing rationality, and all subsequent payoffs are zero.

Setting $V_t(C) = V_t(D)$:

$$3q_t \;=\; (2+3q_{t-1})q_t - (1-q_t) \quad\Longleftrightarrow\quad q_{t-1}\,q_t \;=\; \tfrac{1}{3},$$

which is exactly the recursion built into the proposed q_t schedule.

Remark (Inductive Verification).

> The indifference equation $q_{t-1}q_t = 1/3$ can be verified inductively. *Base case* ($t=2$): defecting yields $3q_2$; cooperating yields $1+3q_1$ if the opponent cooperates (one will defect tomorrow, gaining $3q_1$ when the opponent's commitment type still cooperates) and -1 if the opponent defects. Setting these equal: $3q_2 = q_2(1+3q_1) - (1-q_2)$, which rearranges to $q_1q_2 = 1/3$. *Inductive step*: assume the relation holds for $t-1$, so $V_{t-1}(C) = V_{t-1}(D) = 3q_{t-1}$. Then in period t, defecting earns $3q_t$ today and 0 thereafter (the opponent's commitment type retaliates); cooperating earns 2 today and continuation $3q_{t-1}$ if the opponent cooperated, or -1 today and 0 thereafter if the opponent defected. The same algebra delivers $q_{t-1}q_t = 1/3$.

9.4.6 What the Equilibrium Achieves

The construction yields a striking result: with arbitrarily small $\varepsilon > 0$, sufficiently long horizons T admit a PBE in which rational players cooperate for almost the entire game, randomize during a final endgame phase of length $\Theta(\log(1/\varepsilon))$, and defect outright in the very last period. The fraction of periods featuring full cooperation tends to 1 as $T\to\infty$ with ε fixed.

The mechanism is structurally identical to the chain store paradox: each rational player is willing to cooperate today because doing so preserves the opponent's uncertainty about her type (since a defection would reveal rationality and trigger permanent retaliation). The opponent, in turn, willingly cooperates because the probability that the player is a commitment type is high enough to make cooperation marginally profitable. The reputational equilibrium is held together by the small ε, which provides just enough "cover" for rational players to mimic cooperation without immediate detection.

Remark (Two-Sided vs. One-Sided Reputation).

> The PD reputation model differs from the chain store in one important respect: *both* players are long-lived and *both* have private types. In the chain store, only the store has a type and the entrants are short-lived; reputation flows in one direction. Here, each rational player both maintains and reads reputation simultaneously, and the equilibrium balances the two sides via the $q_{t-1}q_t = 1/3$ recursion. A consequence: even if only *one* side has commitment types (asymmetric uncertainty), the result still goes through; what

matters is that *some* doubt about rationality exists somewhere in the game. The lesson generalizes well beyond PD: cooperative behavior in any finitely repeated game with a unique stage NE can be rescued by a tiny mass of behavioral types, regardless of whether the perturbation is one-sided or two-sided.

Remark (Why a Vanishingly Small ε Suffices).

A natural worry: shouldn't a tiny ε produce only a tiny effect? The answer is no, and the reason is the same logarithmic blow-up that drove the chain store: the length of the cooperative phase scales as $\log(1/\varepsilon)/\log 3$. Halving ε adds roughly $\log 2/\log 3 \approx 0.63$ extra periods of cooperation. So even at $\varepsilon = 10^{-6}$, the cooperative phase lasts about 13 periods—enough that in any moderately long horizon, almost all play looks cooperative. This logarithmic sensitivity is what makes "a small grain of doubt" a structural feature of repeated games rather than a minor perturbation.

Remark (Chapter Summary).

Perfect Bayesian equilibrium (PBE, Definition 9.2) extends subgame perfection to games of incomplete information. The concept has two ingredients: a strategy profile that is sequentially rational at every information set, and a belief system that is consistent with Bayes' rule wherever Bayes' rule applies. PBE has bite in two canonical settings. *Spence's job-market signaling model*: the rich type set produces a continuum of separating PBE (each indexed by an off-path belief restriction), a continuum of pooling PBE, and various hybrids; the *intuitive criterion* of Cho-Kreps selects the least-cost separating equilibrium, which has clean welfare implications and matches the empirical pattern of overinvestment in education. *Reputation in finitely repeated PD*: a tiny mass of "tit-for-tat" commitment types can sustain near-fully-cooperative play in long but finite horizons, with the cooperative phase scaling as $\log(1/\varepsilon)$. PBE is the right concept for incomplete-information dynamics, but its multiplicity (especially of pooling and partial-pooling equilibria) means it usually requires a refinement to deliver predictive bite.

Part VI

Problem Sets and Solutions

Part VII

Exams and Solutions