Game Theory with Applications to Finance and Marketing

Lecture 4: Dynamic Incomplete-Information Games and Perfect Bayesian Equilibrium

Chyi-Mei Chen, Room 1102, Management Building 2
(TEL) 3366-1086
(EMAIL) cchen@ccms.ntu.edu.tw

1. This note consists of two parts. Part I considers the three dynamic games with incomplete information that we may encounter most frequently (namely, screening games, signalling games, and reputation games) and defines the perfect Bayesian equilibrium. In part II, we consider other useful equilibrium concepts, such as sequential equilibrium, Grossman-Perry equilibrium, and divine equilibrium.

2. (**Part I**.)

3. **Screening Games**. Let us start with a special screening game, in which a seller must design the optimal *non-linear pricing scheme*. Consider a risk-neutral monopolistic manufacturer that intends to sell a product to n segments of consumers. The unit cost of production is $c \geq 0$. A segment-i consumer receives gross utility $\theta_i V(q)$ if q units of the product are consumed. Assume that

$$0 < \theta_1 < \theta_2 < \cdots < \theta_n.$$

Let $\pi_i > 0$ be the population of consumer i. Each consumer seeks to maximize consumer surplus $u(q, T, \theta) = \theta V(q) - T$, where T is the monetary payment required to obtain the q units of the product. The manufacturer seeks to maximize expected profits, which is

$$\max_{\{T_i, q_i; i=1,2,\cdots,n\}} \sum_{i=1}^{n} \pi_i (T_i - c q_i),$$

where T_i is the seller's expected revenue from consumers of segment i, and q_i is the sales volume to consumers of segment i. Note that mathematically, the seller's problem is the same as the seller's problem in a different scenario where the seller is faced with only one buyer with n possible, privately known, taste parameters $\theta_1 < \theta_2 < \cdots < \theta_n$, and with π_i being the probability that the buyer's taste parameter is θ_i. In the following, we shall take the latter interpretation unless explicitly stated otherwise.[1]

We shall assume the following Spence-Mirrlees (or sorting, or single-crossing) condition holds:

$$V' > 0 = V(0) > V''.$$

The condition $V' > 0$ implies that at any pair (q, T),

$$\frac{\partial}{\partial \theta}[-\frac{\frac{\partial u}{\partial q}}{\frac{\partial u}{\partial T}}] > 0.$$

That is, at (q, T), if we want to increase q to $q+dq$, and ask by how much we should raise T in order to keep the consumer's surplus unchanged, then the answer depends on θ, and the higher the consumer's θ is, the higher the required increase in T must be. This fact holds for any pair (q, T), and it ensures that a menu of *separating contracts* is possible.[2]

4. **Revelation Principle**. The seller (the manufacturer) must design a selling mechanism, which determines the amount q of the product sold to the buyer, and the associated monetary payment T.

   Any conceivable selling mechanism (or scheme, or contract) can be represented by a *game form*[3] $(S, \{(T(s), q(s)),\ \forall s \in S\})$, where S is an arbitrary pure strategy space specified for the buyer, with the only requirement that elements in S are verifiable in the court of law.

[1] When resale among consumers is allowed, these two problems are no longer mathematically equivalent. Apparently, resale never occurs when there is only one buyer.

[2] This condition is also necessary to ensure a separating equilibrium in a signaling game, which will be defined in a later section.

[3] A game form differs from a normal-form game in that the buyer's payoff, which depends on θ, is unknown to the mechanism designer (the seller), and hence is left unspecified. The seller can specify a game form, but he cannot specify a normal form game for the buyer: without knowing θ, the seller cannot specify the buyer's payoff as a function of $s \in S$ in the strategic game.

Given the game form chosen by the seller, a buyer of type θ can decide to or not to play this game form (giving rise to the IR conditions in program (P) below), and if the buyer chooses to play this game form, he can choose a best response $s^*(\theta) \in S$, and with this choice he understands that he will consume the amount $q(s^*(\theta))$ and is required to pay the seller $T(s^*(\theta))$. Essentially all selling schemes the seller can possibly think of take this form. However, given any such game form $(S, \{(T(s), q(s)),\ \forall s \in S\})$ and the corresponding Bayesian equilibrium $s^*(\theta)$ of the game form, let us consider the following *equivalent direct game form*: The seller asks the buyer to report his $\theta \in \Theta \equiv \{\theta_i; i = 1, 2, \cdots, n\}$, and given the buyer's report θ (the buyer can lie about his true θ if he likes), the buyer will consume $f(\theta)$ and is required to pay the seller $W(\theta)$, where $f(\cdot) = q(s^*(\cdot))$ and $W(\cdot) = T(s^*(\cdot))$. In playing the former game form, the buyer chooses $s^*(\theta)$ as his best response when his type is θ, which implies that the buyer in playing the latter *direct* game form, one best response for him is to report his true θ! To see this, suppose that the buyer of type θ strictly preferred to report another type θ'. However, this would mean that this type of buyer should have strictly preferred $s^*(\theta') \in S$ to $s^*(\theta) \in S$ in the original game form, yielding a contradiction. Our conclusion is therefore this: in searching for an optimal (meaning expected profit maximizing) mechanism or game form, the seller can without loss of generality confine his attention to those game forms that make q and T contingent on the buyer's report about his θ (certainly this requires that θ be describable; we shall come back to this point later), and that ensure that truthfully reporting his type is always one optimal strategy for the buyer (hence giving rise to the IC conditions in program (P) below). This has become a well-known theorem in *contract theory*, referred to as the *revelation principle*.

5. Because of the revelation principle, the seller's problem can be stated as

$$\text{(P)} \quad \max_{(q_i, T_i),\ i=1,2,\cdots,n} \sum_{i=1}^{n} \pi_i [T_i - c q_i]$$

subject to

$$\begin{cases} \text{(IC)} & \forall i, j \quad \theta_i V(q_i) - T_i \geq \theta_i V(q_j) - T_j; \\ \text{(IR)} & \forall i \quad\ \ \theta_i V(q_i) - T_i \geq 0. \end{cases}$$

In the above, we have written $q_i = q(\theta_i)$ and $T_i = T(\theta_i)$ to ease notation.

6. Because of the Spence-Mirrlees condition, program (P) can be replaced by the following simpler program (P′), in the sense that the two programs yield the same optimal selling mechanism.
Theorem AS-1 Problem (P) is equivalent to problem (P′):

$$\max_{(q_i,T_i),\ i=1,2,\cdots,n} \sum_{i=1}^{n} \pi_i [T_i - cq_i]$$

subject to[4]

$$\begin{cases} \text{(LDIC)} & \forall i \geq 2 \quad \theta_i V(q_i) - T_i \geq \theta_i V(q_{i-1}) - T_{i-1} \\ \text{(IR1)} & \qquad\qquad \theta_1 V(q_1) - T_1 \geq 0. \\ \text{(monotonicity)} & \forall i \geq j \quad q_i \geq q_j \end{cases}$$

Proof. This theorem can be proved in 3 steps. First, it can be shown that monotonicity is implied by IC constraints in (P), so that (P′) has a strictly smaller set of feasible solutions than (P) does. Second, it can be shown that at the solution of (P′), (IR1) will be binding, and all the LDIC constraints will be binding. Finally, it can be shown that a feasible solution to (P′) that makes (IR1) and all LDIC constraints binding is also a feasible solution to (P). ‖

7. It can be shown that, if

$$V'(0) > \frac{c}{\theta_i} > V'(+\infty) \equiv \lim_{y \uparrow +\infty} V'(y),$$

then the *socially efficient* level of consumption q_i^* for the type-θ_i buyer, which maximizes $\theta_i V(q) - cq$, is such that

$$\theta_i V'(q_i^*) = c.$$

If the seller has full information about the buyer's parameter θ_i, then the seller's optimal (q_i^*, T_i^*), which is termed *the first-best scheme*, is such that $T_i^* = \theta_i V(q_i^*)$, so that the seller extracts all the consumer surplus from the buyer. In this case, q_i^* must maximize the *social*

[4] LDIC stands for "Local Downward Incentive Compatibility" conditions.

benefit from serving the type-θ_i buyer, which is $\theta_i V(q) - cq$. It follows that $\theta_i V'(q_i^*) = c$ if

$$V'(0) > \frac{c}{\theta_i} > V'(+\infty),$$

and $q_i^* = T_i^* = 0$ if $V'(0) \leq \frac{c}{\theta_i}$.

Now applying theorem AS-1 it is easy to show that under information asymmetry, the seller's optimal scheme, which is termed *the second-best scheme*, is such that the type-n consumers's quantity q_n^{**} attains social efficiency. This is called the property of *efficiency at the top*. On the other hand, if

$$V'(0) > \frac{c}{\theta_1} > \frac{c}{\theta_n} > V'(+\infty),$$

then for all $i < n$, $\theta_i V'(q_i^{**}) > c = \theta_n V'(q_n^{**})$, implying the inefficiency problem of *under-consumption* for each type $i = 1, 2, \cdots, n-1$. Also, the type-1 consumers have no surplus in equilibrium.

8. **Example 1**. Suppose that $n = 2$. The full-information optimal scheme, called the *first best* scheme, $\{(q_i^*, T_i^*); i = 1, 2\ \}$, is such that, assuming an interior solution,[5]

$$\theta_i V'(q_i^*) = c, \quad \theta_i V(q_i^*) = T_i^*,$$

where the two equations require respectively (i) type-θ_i buyer's consumption efficiency and (ii) no consumer's surplus for either type. Intuitively, with full information, the seller can employ his full bargaining power to obtain a producer surplus that equals the entire social benefit, and recognizing that each bit of efficiency gain will ultimately be enjoyed by himself, the seller chooses the first-best output levels to fulfill the social efficiency.

[5] A sufficient condition for an interior first-best solution is the following Inada condition $\lim_{x\downarrow 0} V'(x) = +\infty$ and $\lim_{x\uparrow +\infty} V'(x) = 0$. Another sufficient condition is

$$V'(+\infty) < \frac{c}{\theta_n} < \frac{c}{\theta_1} < V'(0).$$

To give a counter-example, note that if $\frac{c}{\theta_1} > V'(0) > \frac{c}{\theta_2}$, then with the first-best scheme the seller chooses to serve only the type-θ_2 consumer.

With information asymmetry, the first best scheme violates θ_2-type buyer's IC if $q_1^* > 0$:

$$\theta_2 V(q_2^*) - T_2^* = 0 = \theta_1 V(q_1^*) - T_1^* < \theta_2 V(q_1^*) - T_1^*;$$

that is, both types would claim to be type 1! Thus the first-best scheme is never an incentive-feasible scheme in the presence of information asymmetry as long as $q_1^* > 0$. (A scheme is incentive-feasible if it satisfies (IR_1), (IR_2), (IC_1), and (IC_2).) As Theorem AS-1 shows, the optimal scheme $\{(q_i^{**}, T_i^{**});\ i = 1, 2\}$ in this case, called the *second-best* scheme for obvious reasons, requires IR1 and IC2 be binding and that $q_2^{**} \geq q_1^{**}$. It follows that

$$\theta_2 V'(q_2^{**}) = c, \quad \theta_1 V'(q_1^{**}) = \frac{c}{1 - \frac{\pi_2}{\pi_1}\frac{\theta_2 - \theta_1}{\theta_1}} > c,$$

if there is an interior solution for q_1^{**}, or else $q_1^{**} = 0$. We conclude that the θ_2-type buyer maintains his consumption efficiency and may enjoy some consumer surplus, whereas the θ_1-type buyer suffers from underconsumption and has no consumer surplus.

Note that with information asymmetry, the θ_2-type buyer's output level is still the first-best level (simply because no other types would like to claim that they are type θ_2), and given this fact, the seller would like to minimize the θ_2-type buyer's surplus (the seller's surplus from serving the θ_2-type buyer plus the θ_2-type buyer's surplus becomes a constant sum–the efficiency gain of producing q_2^*). The θ_2-type buyer's surplus is

$$(\theta_2 - \theta_1)V(q_1^{**}),$$

which increases with the difference between θ_2 and θ_1 and with the quantity rendered to the θ_1-type buyer. This creates an incentive for the seller to reduce q_1 to below q_1^*. If $q_1^{**} = 0$, then the θ_2-type buyer's surplus is also zero, but the latter remains positive whenever $q_1^{**} > 0$.

The seller's deviation from the first-best selling scheme is motivated by the information asymmetry between him and the buyer. Recall that the seller chooses to fulfill social efficiency in the full information case. Here, with information asymmetry, the seller in choosing $q_1^{**} > 0$ cannot exhaust the θ_2-type buyer's surplus, and this implies that the seller's producer surplus is less than the social benefit. From this perspective, it is not surprising that the seller wants to distort the selling scheme in the presence of information asymmetry.

There is another implication from the above analysis. Note that the type-θ_2 buyer has rent exactly because he has private information. Thus the buyer has an incentive to over-invest in activities that help maintain his information advantage.

9. **Example 2**. A monopolistic firm intends to sell one product to three segments of consumers. For $j = 1, 2, 3$, the population of segment j is denoted by π_j, and a segment-j consumer obtains a surplus $\theta_j v(q) - t$ if she pays t to the firm in order to consume q units of the product. The firm's unit production cost is c. Suppose that $v(q) = \sqrt{1+q}$, and for all $j = 1, 2, 3$, $\theta_j = j$.
(i) Suppose that $\pi_2 = 0$. Suppose that $c = \frac{2}{3}$. Find the optimal contract for the firm.[6]
(ii) Now, ignore part (i). Suppose instead that $c = \frac{1}{4}$, $\pi_1 = \frac{3}{4}$, $\pi_2 = \pi_3 = \frac{1}{8}$. Prove or disprove that $q_1^{SB} = \frac{7}{9}$, $q_2^{SB} = 3$, and $q_3^{SB} = 35$.

10. **Remarks**. The screening games that we have considered so far are a special class of contracting games. We shall cover contracting games in the sequel of this course, in the next semester. In a contracting game, some players have the authority of designing a contract or a mechanism that specifies game forms (a game form is a normal form game leaving the players unspecified) for other players. The following facts are important in modelling and solving a contracting game:
(i) A contract can only enforce events which are observable to all contracting parties and verifiable in the court of law (or other equivalent contract enforcers).
(ii) That LDIC must be binding at optimum is not always true if we allow risk averse principals (note that we have assumed in theorem AS-1 that the seller is risk neutral, and we have done the same in examples 1 and 2.) An example is the model in Hart (1983, Optimal Labor

[6] Compute $\theta_1 v'(0)$, and compare it to c. Conclude that the first-best consumption q_1^* for segment-1 consumers is zero. Now, use the fact that $q_1^{SB} \leq q_1^*$ to determine the optimal q_1^{SB}. Then, by the principle of *efficiency at the top*, derive the optimal q_3^{SB} from $\theta_3 v'(q_3^{SB}) = c$. Show that $q_3^{SB} = \frac{65}{16}$.

Contracts under Asymmetric Information: An Introduction, *Review of Economic Studies*). There, both the workers and the firm (owner) are risk averse. If the firm' profitability were common knowledge, the first best labor contract would take care of both productive efficiency (instead of consumption efficiency as in theorem AS-1) and optimal risk sharing. This first-best contract violates the firm's IC under information asymmetry if the firm is more risk averse than the workers. To get the idea, consider the case where the workers are risk neutral and the firm is risk averse. In this case, the first-best contract will require workers bear all the risk and hence the firm ends up with a constant profit. But this first-best contract becomes infeasible when the firm's profitability is the firm's private information: every type of the firm would like to to claim to be the type that is assigned with the highest non-random profit! At the other extreme, i.e. only the workers are risk averse, the first-best labor contract remains optimal under information asymmetry. In all cases in between, the first best contract is not second best and LDIC will not necessarily be binding at optimum.[7]

11. Now consider the case where $\Theta = [\underline{\theta}, \overline{\theta}]$ is an interval. We shall assume that the density $f(\theta) > 0$ at each point $\theta \in \Theta$. Let $F(\theta)$ be the corresponding distribution function. The optimal contract problem facing the monopoly is

$$(\text{Q}) \quad \max_{q(\cdot),T(\cdot)} \int_{\Theta} [T(\theta) - cq(\theta)] f(\theta) d\theta,$$

$$\text{s.t.} \quad (\text{IR}) \quad \theta V(q(\theta)) - T(\theta) \geq 0, \forall \theta \in \Theta,$$

$$(\text{IC}) \quad \theta V(q(\theta)) - T(\theta) \geq \theta V(q(\theta')) - T(\theta'), \forall \theta, \theta' \in \Theta.$$

Observe that, for $\theta > \theta'$, by IC,

$$\theta V(q(\theta)) - T(\theta) \geq \theta V(q(\theta')) - T(\theta') \geq \theta' V(q(\theta')) - T(\theta'),$$

[7]That information asymmetry leads to various inefficiency problems should be the first lesson to be learned here. We have seen in the above non-linear pricing examples that all consumers except for the highest type suffer from under-consumption under the seller's second best non-linear pricing scheme. Here, Hart shows that when the full-information Walrasian models in labor economics fail to explain the high and sticky unemployment observed in the real world, risk aversion on the part of firms plus information asymmetry can easily justify this phenomenon.

and thus IR1 (i.e., $\underline{\theta} V(q(\underline{\theta})) - T(\underline{\theta}) \geq 0$) and IC together imply IR. Next, as in the proof for theorem AS-1, IC plus the Spence-Mirrlees condition implies that $q(\theta)$ is weakly increasing (monotonicity), and hence almost everywhere differentiable. Now, for true type θ and the agent's report θ', define

$$W(\theta, \theta') \equiv \theta V(q(\theta')) - T(\theta').$$

Let $\hat{\theta}^*(\theta)$ be one optimal report made by a type-θ agent. By IC, one solution is that $\hat{\theta}^*(\theta) = \theta$ for all $\theta \in \Theta$. As a necessary condition, the local downward and upward IC (LIC, from now on) requires that, for all $\theta \in (\underline{\theta}, \overline{\theta})$,

$$w(\theta) \equiv W(\theta, \hat{\theta}^*(\theta)) = W(\theta, \theta) \geq W(\theta, \theta + d\theta).$$

By envelope theorem,[8] we have

$$(A) \qquad w'(\theta) \equiv \frac{dW(\theta, \hat{\theta}^*(\theta))}{d\theta} = V(q(\theta)).$$

Just like in the case of n types, here LIC, monotonicity and the Spence-Mirrlees condition together imply IC.

Thus, a maximization program with the same objective function as in (Q) but imposing only IR1, monotonicity and LIC, must yield the same optimal solution. In particular, IR1 means that

$$(B) \qquad w(\underline{\theta}) = 0.$$

[8] Consider the maximization program

$$\max_y g(y, x)$$

with g being twice continuously differentiable, and suppose that given each x, the optimal $y^*(x)$ is differentiable and it solves the following first-order condition (and hence is an interior solution)

$$\frac{\partial g}{\partial y}(y^*(x), x) = 0.$$

Define $G(x) \equiv g(y^*(x), x)$. Then the envelope theorem says that

$$\frac{dG}{dx} = \frac{\partial g}{\partial x}(y^*(x), x).$$

This theorem can be verified by direct computations.

By the fundamental theorem of calculus, (A) and (B), or (LIC) and (IR1), yield

$$w(\theta) = \int_{\underline{\theta}}^{\theta} V(q(x))dx,$$

implying that

$$T(\theta) = \theta V(q(\theta)) - \int_{\underline{\theta}}^{\theta} V(q(x))dx.$$

Hence instead of solving the above maximization program (Q), we can solve the following program:

$$(\text{P}'') \quad \max_{q(\cdot)} \int_{\Theta} [\theta V(q(\theta)) - \int_{\underline{\theta}}^{\theta} V(q(x))dx - cq(\theta)] f(\theta) d\theta$$

$$\text{s.t.} \quad q(\theta) \quad \text{is increasing in} \quad \theta.$$

The objective function in (P″) can be further rewritten as

$$\max_{q(\cdot)} \int_{\Theta} ([\theta V(q(\theta)) - cq(\theta)] f(\theta) - V(q(\theta))(1 - F(\theta))) d\theta$$

using integration by parts.[9] Here comes our major result for the current continuum case:

Theorem AS-2 If

$$g(\theta) = \theta - [\frac{f(\theta)}{1 - F(\theta)}]^{-1}$$

[9]Note that

$$\int_{\Theta} [\int_{\underline{\theta}}^{\theta} V(q(x))dx] f(\theta) d\theta = \int_{\Theta} [\int_{\underline{\theta}}^{\theta} V(q(x))dx] dF(\theta)$$

$$= F(\bar{\theta}) \int_{\Theta} V(q(x))dx - \int_{\Theta} F(\theta) d[\int_{\underline{\theta}}^{\theta} V(q(x))dx]$$

$$= \int_{\Theta} V(q(x))dx - \int_{\Theta} F(\theta) V(q(\theta)) d\theta]$$

$$= \int_{\Theta} [1 - F(\theta)] V(q(\theta)) d\theta.$$

is strictly increasing in θ, then the optimal solution $q^{SB}(\cdot)$ to (P$''$) is characterized by (i) $q^{SB}(\theta) > 0$ if and only if θ exceeds some $\hat{\theta} \in [\underline{\theta}, \overline{\theta}]$; (ii) complete sorting (i.e., $q^{SB}(\theta) > q^{SB}(\theta')$ for all $\theta > \theta'$ such that $\hat{\theta} < \theta' < \theta \leq \overline{\theta}$); and (iii) quantity discounts (i.e., $T(q^{SB}(\theta))$ is strictly concave in $q^{SB}(\theta)$).

Proof. The objective function in (P$''$) is

$$\int_{\Theta} [g(\theta)V(q(\theta)) - cq(\theta)]f(\theta)d\theta.$$

Now, if we perform point-wise optimization to the integrand, and refer to the point-wise optimal solution as $q^*(\cdot)$, then we obtain

$$\theta V'(q^*(\theta)) = c + \frac{1 - F(\theta)}{f(\theta)} V'(q^*(\theta)), \tag{1}$$

whenever $q^*(\theta) > 0$. When

$$g(\theta) = \theta - [\frac{f(\theta)}{1 - F(\theta)}]^{-1}$$

is strictly increasing in θ, it is easy to see that $q^*(\cdot)$ satisfies the monotonicity, and is hence almost everywhere differentiable. It follows that

$$\frac{d}{d\theta}[g(\theta)V(q^*(\theta)) - cq^*(\theta)] = g'(\theta)V(q^*(\theta)) \geq 0,$$

and the inequality is strict whenever $q^*(\theta) > 0$. Thus there exists $\hat{\theta} \in \Theta$ such that

$$[g(\theta)V(q^*(\theta)) - cq^*(\theta)]f(\theta) \leq 0$$

if and only if $\theta \leq \hat{\theta}$. Thus we have

$$q^{SB}(\theta) = \begin{cases} 0, & \text{if } \theta \leq \hat{\theta}; \\ \\ q^*(\theta), & \text{if } \theta > \hat{\theta}. \end{cases}$$

This proves parts (i) and (ii). Note that, except for type $\overline{\theta}$, all other types of consumers suffer from under-consumption.

Finally, we consider part (iii). Note that we can write $T^{SB}(\theta)$ as $T(q^{SB}(\theta))$; i.e., if $q^{SB}(\theta) = q^{SB}(\theta')$ then we must have $T^{SB}(\theta) = T^{SB}(\theta')$ in order to satisfy the IC's for types θ and θ'. Note that the type-θ consumer's first-order condition is

$$T'(q^{SB}(\theta)) = \theta V'(q^{SB}(\theta)).$$

Now we can re-write (1) as

$$\frac{p-c}{p} = \frac{1-F(\theta)}{\theta f(\theta)},$$

where $p = p(q^{SB}(\theta)) \equiv T'(q^{SB}(\theta))$ is the marginal price for an additional unit purchased. Note that the above right-hand side is decreasing in θ by the monotone hazard rate property, and that the above left-hand side is strictly increasing in p. Thus we have $\frac{dp}{d\theta} < 0$. From this, using the chain rule, we deduce that

$$\frac{dp(q^{SB}(\theta))}{dq^{SB}(\theta)} = \frac{dp/d\theta}{dq^{SB}/d\theta} < 0.$$

This proves that $T(q^{SB}(\theta))$ is strictly concave in $q^{SB}(\theta)$. $\|$

12. **Remarks**. Uniform, normal, and exponential distributions all satisfy the strictly monotone hazard rate property; i.e. $\frac{f(\theta)}{1-F(\theta)}$ is strictly increasing. Note that as long as the hazard rate is weakly increasing, $g(\cdot)$ is strictly increasing, and by Theorem AS-2, the second best scheme involves complete sorting.

13. Now we consider a generalized version of the above screening game. A risk neutral (monopolistic) seller sells a single product to a consumer (the buyer). The buyer has n possible types, $\theta_1 < \theta_2 < \cdots < \theta_n$. The buyer, given his type θ_i, seeks to maximize his consumer surplus

$$u(q, R, \theta_i) = N(q, \theta_i) - R,$$

where q is the purchased quantity, R the seller's profit (the producer surplus), and N the social surplus. More specifically,

$$N(q, \theta_i) = \int_0^q P(x, \theta_i)dx - cq,$$

where $P(\cdot, \theta_i)$ is θ_i-type buyer's inverse demand function and c the seller's constant cost of production; and

$$R = T - cq$$

where T is the total payment the buyer makes to the seller for this transaction. We shall assume that $N(\cdot, \cdot)$ is a twice continuously differentiable function. Note that

Total surplus= consumer's surplus + producer's surplus,

and hence, equivalently, we have

$$u^i = \int_0^q P(x, \theta_i)dx - T.$$

Note that the preceding model is a special version of the current one, under the following restriction:

$$\int_0^q P(x, \theta_i)dx \equiv \theta_i V(q),$$

where $V' > 0 = V(0) > V''$. We shall assume that $\forall i$,

$$N_{12} = \frac{\partial^2 N}{\partial q \partial \theta_i} > 0,$$

which implies the Spence-Mirrlees (or sorting, single-crossing) condition: $\forall i$,

$$\frac{\partial}{\partial \theta_i}[-\frac{\frac{\partial u^i}{\partial q}}{\frac{\partial u^i}{\partial T}}] > 0.$$

The seller's problem is, by the revelation principle,

$$\text{(P)} \quad \max_{(q_i, R_i),\ i=1,2,\cdots,n} \sum_{i=1}^n p_i R_i$$

subject to

$$\begin{cases} \text{(IC)} & \forall i, j \quad N(q_i, \theta_i) - R_i \geq N(q_j, \theta_i) - R_j \\ \text{(IR)} & \forall i \quad\ \ \ N(q_i, \theta_i) - R_i \geq 0. \end{cases}$$

Here note that p_i is the prob. for θ_i.

To conduct the analysis, we first give four lemmas.

Lemma 1. $(R,q) \succeq_\theta (R',q'),\ q>q',\ \theta'>\theta \Rightarrow (R,q) \succ_{\theta'} (R',q')$.
Proof.

$$0 < \int_\theta^{\theta'} \int_{q'}^{q} N_{12}(t,s)dtds$$
$$= [N(q,\theta') - N(q',\theta')] - [N(q,\theta) - N(q',\theta)]$$
$$\leq [N(q,\theta') - N(q',\theta')] - [R - R'].\|$$

Lemma 2. $(R,q) \sim_\theta (R',q'),\ q>q',\ \theta>\theta'' \Rightarrow (R',q') \succ_{\theta''} (R,q)$.
Proof.

$$0 < \int_{\theta''}^{\theta} \int_{q'}^{q} N_{12}(t,s)dtds$$
$$= [N(q,\theta) - N(q',\theta)] - [N(q,\theta'') - N(q',\theta'')]$$
$$= [R - R'] - [N(q,\theta'') - N(q',\theta'')].\|$$

Lemma 3. Suppose that with $\{(R_i,q_i); i=1,2,\cdots,n\}$, where $\{q_i; i=1,2,\cdots,n\}$ satisfies the monotonicity, LDIC's are all binding. Then for all $i=1,2,\cdots,n-1$, $(R_i,q_i) \succeq_{\theta_i} (R_{i+1},q_{i+1})$.
Proof.

$$\int_{q_i}^{q_{i+1}} \int_{\theta_i}^{\theta_{i+1}} N_{12}(t,s)dsdt \geq 0$$
$$\Rightarrow N(q_{i+1},\theta_{i+1}) - N(q_{i+1},\theta_i) - N(q_i,\theta_{i+1}) + N(q_i,\theta_i) \geq 0$$
$$\Rightarrow [N(q_{i+1},\theta_{i+1}) - R_{i+1}] - [N(q_{i+1},\theta_i) - R_{i+1}] - [N(q_i,\theta_{i+1}) - R_i] + [N(q_i,\theta_i) - R_i] \geq 0$$
$$\Rightarrow N(q_i,\theta_i) - R_i \geq N(q_{i+1},\theta_i) - R_{i+1},$$

where the last "⇒" follows from the fact that LDIC_{i+1} is binding. ‖

Lemma 4. Suppose that with $\{(R_i, q_i); i = 1, 2, \cdots, n\}$, where $\{q_i; i = 1, 2, \cdots, n\}$ satisfies the monotonicity, LDIC's are all satisfied. Then for all $i = 1, 2, \cdots, n-1$, there exists $R_i' \geq R_i$ such that $(R_i', q_i) \sim_{\theta_i} (R_{i+1}, q_{i+1})$.
Proof. This follows from lemma 3 and the fact that $u(q, R, \theta) = N(q, \theta) - R$. $\|$

Theoerm AS-3.
Problem (P) is equivalent to problem (P′) below:

$$(\mathrm{P}') \quad \max_{(q_i, R_i),\ i=1,2,\cdots,n} \sum_{i=1}^{n} p_i R_i$$

subject to

$$\begin{cases} \text{(LDIC)} & \forall i \geq 2 \quad N(q_i, \theta_i) - R_i \geq N(q_{i-1}, \theta_i) - R_{i-1} \\ \text{(IR1)} & \qquad N(q_1, \theta_1) - R_1 \geq 0. \\ \text{(monotonicity)} & \forall \theta_i \geq \theta_j \quad q_i \geq q_j \end{cases}$$

Proof. We shall prove the following three statements: (i) The constraints of (P) imply those of (P′); (ii) At the solution of (P′), $\forall i \geq 2$, LDIC (Local Downward Incentive Compatibility conditions) must be binding; and (iii) The solution of (P′) satisfies the constraints of (P).
proof of (i): Everything is obvious except for the monotonicity, for which note that by IC, for $\theta_i \geq \theta_j$,

$$N(q_i, \theta_i) - R_i \geq N(q_j, \theta_i) - R_j$$

and

$$N(q_j, \theta_j) - R_j \geq N(q_i, \theta_j) - R_i.$$

Rearranging, we have

$$N(q_i, \theta_i) - N(q_i, \theta_j) \geq N(q_j, \theta_i) - N(q_j, \theta_j)$$

or, equivalently,

$$\int_{\theta_j}^{\theta_i} N_2(q_i, \theta) d\theta \geq \int_{\theta_j}^{\theta_i} N_2(q_j, \theta) d\theta.$$

But this means that

$$\int_{q_j}^{q_i}[\int_{\theta_j}^{\theta_i} N_{12}(q,\theta)d\theta]dq \geq 0.$$

Since the expression in the square bracket is positive (by Spence-Mirrlees condition), it is necessary that $q_i \geq q_j$.
<u>proof of (ii)</u>: Assume not. Then, there is some $i \geq 2$ such that

$$N(q_i,\theta_i) - R_i > N(q_{i-1},\theta_i) - R_{i-1}.$$

But, we can raise R_j by some positive ϵ for all $j \geq i$ without affecting LDIC's. This certainly raises the seller's payoff, contradicting to optimality.
<u>proof of (iii)</u>: Fix θ_i. We want to show that

$$(R_i,q_i) \succeq_{\theta_i} (R_j,q_j),\ \forall j \neq i.$$

This certainly is true for $j = i-1$. Since for all $j = 1,2,\cdots,i-2$, $q_j \leq q_{j+1}$ and

$$(R_{j+1},q_{j+1}) \sim_{\theta_{j+1}} (R_j,q_j),$$

by monotonicity, lemma 1, and LDIC$_i$, we have

$$(R_i,q_i) \sim_{\theta_i} (R_{i-1},q_{i-1}) \succ_{\theta_i} (R_{i-2},q_{i-2}) \succ_{\theta_i} \cdots \succ_{\theta_i} (R_1,q_1).$$

On the other hand, by monotonicity, Lemmas 2,3, and 4, we have

$$(R_i,q_i) \succeq_{\theta_i} (R_{i+1},q_{i+1}) \succeq_{\theta_i} (R'_{i+1},q_{i+1}) \succ_{\theta_i} (R_{i+2},q_{i+2}) \succ_{\theta_i} \cdots \succ_{\theta_i} (R_n,q_n).$$

Thus all IC's are satisfied at an optimum of (P$'$). It remains to verify that all the IR's are also satisfied. Suppose that $q_1 = 0$. Then $R_1 = 0$. As for all θ_i, $N(0,\theta_i) = 0$, the IR's follow from IC's automatically. Suppose instead that $q_1 > 0$. Define $(R'_1,q'_1) = (0,0)$. Since $(R_1,q_1) \succeq_{\theta_1} (R'_1,q'_1)$, it follows from the IC's and Lemma 1 that for all θ_i, $(R_i,q_i) \succ_{\theta_i} (R_1,q_1) \succ_{\theta_i} (R'_1,q'_1)$, proving that IR's are all satisfied.‖

A direct consequence of Theorem AS-3 is the following corollary.
Corollary AS-4. At optimum of (P), $N_1(q_i,\theta_i) > 0$, $\forall i < n$, and $N_1(q_n,\theta_n) = 0$. That is, over-consumption (that $q_i > q_i^{FB}$ for some i)

never takes place at optimum, and except type n, all other types of the buyer suffer from under-consumption.[10]

Proof. Suppose that $N_1(q_i, \theta_i) < 0$ for some i. Then let j be the smallest such i. Demand Law implies that N is strictly concave in q so that q_i^{FB} is such that

$$N_1(q_i^{FB}, \theta_i) = 0, \ \forall i.$$

By the sorting condition, this implies that $q_i^{FB} < q_{i+1}^{FB}$. Thus we have (if $j \geq 2$; the same argument goes for $j = 1$ in the absence of type θ_{j-1} below)

$$q_j > q_j^{FB} > q_{j-1}^{FB} \geq q_{j-1}.$$

Now reducing q_j and raising R_j slightly to keep type θ_j indifferent implies that LDIC$_{j+1}$ is relaxed but monotonicity and all other LDIC's are left unaffected. This raises the seller's expected profit from type θ_j and it further allows the seller to extract rents from types $\theta_{j+1}, \cdots, \theta_n$ by raising $R_i, \ i \geq j+1$. This is a contradiction.

Next, suppose $N_1(q_n, \theta_n) > 0$. It is obvious that the seller can benefit from replacing (q_n, R_n) by (q_n^{FB}, R'), where $R' > R_n$ makes type θ_n indifferent about this change. This affects neither LDIC nor monotonicity. This is another contradiction.

Finally, let k be the smallest i such that $q_i = q_i^{FB}$. Suppose $k < n$. Consider changing (q_k^{FB}, R_k) into $(q_k^{FB} - e, R')$ where $R' < R_k$ makes type θ_k indifferent about this change. This move does not affect monotonicity (because $q_i = q_i^{FB} > q_{i-1}^{FB} \geq q_{i-1}$) by the definition of k if e is small enough. This move reduces the seller's profit from type θ_k by

$$\begin{aligned} R_k - R' &= N(q_k^{FB}, \theta_k) - N(q_k^{FB} - e, \theta_k) \\ &= N_1(q_k^{FB}, \theta_k)e - N_{11}(\xi, \theta_k)\frac{e^2}{2}, \\ &= -N_{11}(\xi, \theta_k)\frac{e^2}{2}, \end{aligned}$$

where $\xi \in [q_k^{FB} - e, q_k^{FB}]$ and the second equality follows from Taylor's expansion theorem with remainders.

[10]See the note *Corollary AS-2* for another proof that does not rely on the differentiability of N.

On the other hand, type θ_{k+1}'s LDIC is changed into

$$N(q_{k+1},\theta_{k+1}) - R_{k+1} \geq N(q_k^{FB} - e, \theta_{k+1}) - R'$$

$$= N(q_k^{FB},\theta_{k+1}) - N_1(q_k^{FB},\theta_{k+1})e + \frac{e^2}{2}N_{11}(\zeta,\theta_{k+1}) - R_k - N_{11}(\xi,\theta_k)\frac{e^2}{2},$$

where $\zeta \in [q_k^{FB} - e, q_k^{FB}]$. Thus after lowering q_k^{FB} by $e > 0$, the seller can replace $R_i,\ i \geq k+1$ by

$$R_i + N_1(q_k^{FB},\theta_{k+1})e - \frac{e^2}{2}(N_{11}(\zeta,\theta_{k+1}) - N_{11}(\xi,\theta_k)).$$

The net change in the seller's expected profit is therefore

$$[\sum_{i=k+1}^{n} p_i][N_1(q_k^{FB},\theta_{k+1})e - \frac{e^2}{2}(N_{11}(\zeta,\theta_{k+1}) - N_{11}(\xi,\theta_k))] - p_k N_{11}(\xi,\theta_k)\frac{e^2}{2},$$

which is strictly positive as $e \downarrow 0$, because (i) $q_k^{FB} < q_{k+1}^{FB}$; (ii) $N_{11} = \frac{\partial^2 N}{(\partial q)^2} < 0$ so that $N_1(q_k^{FB},\theta_{k+1}) > N_1(q_{k+1}^{FB},\theta_{k+1}) = 0$; and (iii) $N(\cdot,\cdot)$ is a C^2 function of which the first- and second-order partial derivatives are bounded uniformly on the set

$$\{(q,\theta) : q \in [\frac{1}{2}(q_{k-1}^{FB} + q_k^{FB}), q_k^{FB}],\ \theta \in [\theta_k,\theta_{k+1}]\}.$$

This contradicts to optimality. The proof is complete. $\|$

14. **Signalling Games**. Now we introduce *signaling games.* A game is called a signaling game if it is featured by (i) two players (called 1 and 2); (ii) only player 1 has types $\theta \in \Theta$; (iii) player 1 (also referred to as the *informed player*) given her type sends a signal $a_1 \in A_1$; (iv) player 2 (also referred to as the *uninformed player*) chooses action $a_2 \in A_2$ upon seeing a_1; and (v) the game ends after player 2 chooses a_2.

15. Let Θ be a finite set; i.e., $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$. Let $\mathcal{A}_i$ be the set of probability distributions over A_i. Let Σ_1 be the set of player 1's mixed strategies $\sigma_1 : \Theta \to \mathcal{A}_1$, where given a particular σ_1, let $\sigma_1(a_1|\theta)$ be the probability that player 1 chooses $a_1 \in A_1$ given her type θ. Let Σ_2 be the set of player 2's mixed strategies $\sigma_2 : A_1 \to \mathcal{A}_2$, where given

a particular σ_2, $\sigma_2(a_2|a_1)$ is the prob. that player 2 chooses a_2 upon seeing the signal a_1 sent by player 1. Let $p(\theta)$ be player 2's prior beliefs about player 1's type θ. Let $\mathcal{P}$ be the set of prob. distributions over Θ. Let $\mathcal{M}$ be the set of mappings $\mu : A_1 \to \mathcal{P}$.

Definition 1. A *perfect Bayesian equilibrium* (PBE) for the signaling game specified above is a triple $(\sigma_1, \sigma_2, \mu) \in \Sigma_1 \times \Sigma_2 \times \mathcal{M}$ such that

$$\text{(P1) } \forall \theta \in \Theta,\ \forall \alpha_1 \in \mathcal{A}_1,\ \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \sigma_1(a_1|\theta)\sigma_2(a_2|a_1)u_1(a_1, a_2, \theta)$$

$$\geq \sum_{a_1 \in A_1} \sum_{a_2 \in A_2} \alpha_1(a_1)\sigma_2(a_2|a_1)u_1(a_1, a_2, \theta);$$

$$\text{(P2) } \forall a_1 \in A_1,\ \forall \alpha_2 \in \mathcal{A}_2,\ \sum_{\theta \in \Theta} \mu(\theta|a_1) \sum_{a_2 \in A_2} \sigma_2(a_2|a_1)u_2(a_1, a_2, \theta)$$

$$\geq \sum_{\theta \in \Theta} \mu(\theta|a_1) \sum_{a_2 \in A_2} \alpha_2(a_2)u_2(a_1, a_2, \theta);$$

$$\text{(B) } a_1 \in A_1,\ \sum_{\theta' \in \Theta} p(\theta')\sigma_1(a_1|\theta') > 0,$$

$$\Rightarrow \mu(\theta|a_1) = \frac{p(\theta)\sigma_1(a_1|\theta)}{\sum_{\theta' \in \Theta} p(\theta')\sigma_1(a_1|\theta')}.$$

Thus a PBE for the signaling game is a BE plus a system of posterior beliefs such that the posterior beliefs are obtained by applying Bayes' law to the prior $p(\cdot)$ and player 1's equilibrium strategy σ_1 whenever possible. Unlike in the static games with incomplete information where the definition of BE is adequate, here we cannot define player 2's optimal strategies unless we specify explicitly what player 2's beliefs are when she chooses her actions. Since these beliefs are endogenously dependent upon a_1, we need in the above definition a "system" of posterior beliefs; namely, one probability distribution over Θ for each $a_1 \in A_1$, whether or not a_1 is reached on the equilibrium path.

16. As an example, recall the Spence signaling model, where $\Theta = \{H, L\}$, $H > L \geq 0$, $A_1 = [0, +\infty)$, $p(H) = \alpha$, $A_2 = \mathcal{R}$, $u_1(a_1, a_2, \theta) = a_2 - a_1 c_\theta$ with $0 < c_H < c_L$, and $u_2(a_1, a_2, \theta) = \theta - a_2$. The interpretation is that player 1 would become an employee capable of producing output worth θ (which is player 1's private information) for the firm, must first decide to receive a number a_1 of years of education before going

to work for player 2, and a_1 is later regarded as a signal about θ by the Bertrand-competitive employer player 2 during the job interveiw. Player 2, who then sets the wage a_2 for player 1 upon seeing his diploma. The assumption that player 2 is Bertrand competitive in recruiting player 1 implies player 2's equilibrium payoff (expected value of u_2) has to be zero, and hence $a_2 = \mu(H|a_1)H + [1 - \mu(H|a_1)]L$.

Note that by assumption receiving education is wasteful for player 1, because it does not change player 1's productivity H or L. The sorting condition $0 < c_H < c_L$ says that, when required to spend more time on education, the type-H player 1 is willing to accept a smaller amount of raise in his pay than the type-L player 1 to compensate his loss from having to stay longer at school. This condition ensures that this game has a separating equilibrium.[11]

17. We usually classify the PBE's in a signaling game into two categories: the pure-strategy PBE's and the mixed-strategy PBE's. A PBE is a pure-strategy PBE if both σ_1 and σ_2 are pure strategies. Otherwise, the PBE must involve some player using a mixed strategy, and is hence termed a mixed strategy PBE. Two important pure-strategy PBE's are separating PBE's and pooling PBE's. (If Θ has exactly two elements, these are the only possible kinds of pure-strategy PBE's.) Given a PBE, let us call $a_1 \in A_1$ an *equilibrium signal* if there exists $\theta \in \Theta$ such that $\sigma_1(a_1|\theta) > 0$; that is, a_1 may appear on the equilibrium path of the PBE with a strictly positive probability.[12]

    - In a separating PBE, different types of player 1 use different pure strategies, so that in equilibrium, either a_1 occurs with zero prob. or seeing a_1 player 2 knows θ for sure. In this case, $\mu(\cdot|a_1)$ is either degenerate at some θ or we can impose no restrictions on it. More precisely, if we let a_1^k be such that $\sigma_1(a_1^k|\theta_k) = 1$ so that $\sigma_1(a_1^k|\theta) = 0$ for all $\theta \neq \theta_k$, then Bayes' Law imposes no

[11] One can check that no such sorting condition holds in the signaling game *beer and quiche*, which will be studied in a later section, and hence there does not exist a separating equilibrium for that game.

[12] By this definition, a separating PBE is one where all equilibrium signals fully reveal the informed player's types, and a pooling PBE is one where no equilibrium signals ever fully reveal the informed player's types.

restrictions on $\mu(\cdot|a)$ for all $a \in A_1 \setminus \{a_1^1, a_1^2, \cdots, a_1^n\}$. On the other hand, for $k = 1, 2, \cdots, n$,

$$\mu(\theta|a_1^k) = \frac{p(\theta)\sigma_1(a_1^k|\theta)}{\sum_{m=1}^n p(\theta_m)\sigma_1(a_1^k|\theta_m)} = \frac{p(\theta)\sigma_1(a_1^k|\theta)}{p(\theta_k)\sigma_1(a_1^k|\theta_k)} = \begin{cases} 1, & \theta = \theta_k; \\ 0, & \theta \neq \theta_k. \end{cases}$$

- In a pooling PBE, all types of player 1 choose the same $a_1^* \in A_1$ in equilibrium (that is, $\sigma_1(a_1^*|\theta_k) = 1$ for all $k = 1, 2, \cdots, n$), so that by Bayes' Law

$$\mu(\theta|a_1^*) = \frac{p(\theta)\sigma_1(a_1^*|\theta)}{\sum_{k=1}^n p(\theta_k)\sigma_1(a_1^*|\theta_k)} = \frac{p(\theta)}{\sum_{k=1}^n p(\theta_k)} = p(\theta), \quad \forall \theta \in \Theta.$$

  On the other hand, Bayes' Law imposes no restrictions on $\mu(\cdot|a_1')$ if $a_i' \in A_1 \setminus \{a_1^*\}$.

- More generally, a pure-strategy PBE can be always described as a *partition PBE*, where for each equilibrium signal there is one partition set of Θ that contains those types of player 1 that send that equilibrium signal, and there is a partition set containing all the off-the-equilibrium signals. Note that a pooling equilibrium is simply a partition equilibrium with Θ being the unique partition set containing the single equilibrium signal, and a separating equilibrium is a partition equilibrium where to each type of the informed player there correspondingly exists a singleton partition set containing the equilibrium signal sent by that type. From here we see that for a separating PBE to exist it is necessary that the number of feasible signals (the cardinality of A_1) must be greater than or equal to the number of player 1's types (the cardinality of Θ).

- If Θ and A_1 both have two elements, then a PBE where one type uses a pure strategy and the other type uses a mixed-strategy is a semi-pooling PBE. In general, a semi-pooling PBE is a PBE where some equilibrium signals fully reveal the informed player's type, but other equilibrium signals do not. In other words, in a semi-pooling PBE, there must exist $a_1, a_1' \in A_1$ and $\theta', \theta'', \theta''' \in \Theta$ with $\sigma_1(a_1|\theta) = 0 < \sigma_1(a_1|\theta')$ for all $\theta \neq \theta'$ and $\sigma_1(a_1'|\theta''), \sigma_1(a_1'|\theta''') > 0$ such that

$$\mu(\theta'|a_1) = 1 > \mu(\theta''|a_1'), \mu(\theta'''|a_1') > 0.$$

That is, player 2 knows player 1's type for sure upon seeing a_1, but the uncertainty remains if she sees a_1', where a_1 and a_1' are both equilibrium signals.

18. To ease notation, replace a_1 and a_2 by e and w respectively, and denote $\mu(H|e)$ by $f(e)$. By definition, a pure strategy PBE for Spencian signaling game is a triple $(e_H, e_L, f(e) = prob.(H|e), w(e))$, such that (i) if $e_H = e_L = e^*$ (a *pooling equilibrium*), then

$$f(e^*) = \alpha, \text{ (prior beliefs are also posterior)}$$

$$f(e) \in [0,1], \ \forall e \neq e^*, \text{ (Bayes law has no bite)}$$

$$w(e) = f(e)H + [1 - f(e)]L, \ \forall e,$$

$$w(e^*) - c_H \cdot e^* \geq w(e) - c_H \cdot e, \ \forall e,$$

$$w(e^*) - c_L \cdot e^* \geq w(e) - c_L \cdot e, \ \forall e;$$

and such that (ii) if $e_H \neq e_L$ (called a *separating equilibrium*), then

$$f(e_H) = 1, \ f(e_L) = 0 \text{ (uncertainty is completely resolved)}$$

$$f(e) \in [0,1], \ \forall e \neq e_H, e_L, \text{ (Bayes law has no bite)}$$

$$w(e) = f(e)H + [1 - f(e)]L, \ \forall e,$$

$$w(e_H) - c_H \cdot e_H \geq w(e) - c_H \cdot e, \ \forall e,$$

$$w(e_L) - c_L \cdot e_L \geq w(e) - c_L \cdot e, \ \forall e.$$

19. Now we solve for the pooling equilibria for the Spence signaling game. Suppose both H and L choose e^* and $f(e) = 0, \forall e \neq e^*$. (Note that we have arbitrarily selected a set of supporting beliefs $f(e)$ for off-equilibrium signals e, subject to the only requirement that the system $f(e)$ of posterior beliefs supports e^* as the best response of both type H and type L.) Then, $w_H = w_L = w(e^*) = \alpha H + (1-\alpha)L \equiv \overline{w}$. We need the following *incentive compatibility conditions* (hereafter IC conditions) to hold for respectively type-H and type-L workers:

$$\overline{w} - c_H e^* \geq L - c_H \cdot 0,$$

$$\overline{w} - c_L e^* \geq L - c_L \cdot 0.$$

Thus, we have a continuum (an uncountably infinite number) of pooling equilibria: Each $e^* \leq \frac{\overline{w}-L}{c_L}$ corresponds to one pooling PBE.

20. Next consider the separating equilibria. Suppose H and L chooses respectively E and e, $E \neq e$, and $f(e') = 0, \forall e' \neq E$ (again this system of posterior beliefs best relaxes the informed player's IC; it is unnecessarily strong). In this case, of course, $w(E) = H$, $w(e) = L$ (why?). We need the following IC conditions to hold:

$$H - c_H E \geq L - c_H \cdot 0,$$

$$L - c_L e \geq H - c_L \cdot E.$$

Immediately, we have $e = 0$ (why?). Again, we have a continuum of separating equilibria: $\frac{H-L}{c_H} \geq E \geq \frac{H-L}{c_L}$.

21. Now consider the equilibrium where one type randomizes over two signal levels and the other concentrates on one signal level.[13]

At first, consider the semi-pooling PBE where H randomizes over E (prob. p) and e (prob. $1-p$) and L plays e with probability one, with $E > e$. (One can verify that it would violate L's IC if we assumed that L plays E with probability one; this results from the sorting condition $c_H < c_L$.) For simplicity, assume that $f(e') = 0$, $e' \neq E, e$. To be consistent with equilibrium, it must be that, for H,

$$H - c_H E = \frac{\alpha(1-p)H + (1-\alpha)L}{(1-\alpha) + \alpha(1-p)} - c_H e,\ H - c_H E \geq L,$$

and for L,

$$\frac{\alpha(1-p)H + (1-\alpha)L}{(1-\alpha) + \alpha(1-p)} - c_L e \geq \max(L, H - c_L E).$$

[13] Are these the only possible type of semi-pooling equilibria? In equilibrium type $i \in \{H, L\}$ randomizes over at most two different education levels. For example, suppose that in equilibrium L randomizes over $e_1 > e_2 > e_3$. Then L feels indifferent about $(w(e_j), e_j)$ for all $j = 1, 2, 3$. In this case, H must strictly prefer e_1 to the other two (why?). But then both e_2 and e_3 are separating outcomes, and L should not have randomized over e_2! Conclude that there can be a third type of semi-pooling PBE where both types of the worker randomize over two education levels. Assume therefore H randomizes over E and m and L over m and e, with $E > m > e$. It follows that $e = 0$. Let H and L assign respectively prob. a and prob. b to m. Then $w(m) = \frac{\alpha a H + (1-\alpha) b L}{a\alpha + (1-\alpha) b}$ and $w(E) = H$ with $L = w(m) - c_L m$ and $H - c_H E = w(m) - c_H m$. An obvious supporting belief is that $f(e') = 0$ for all $e' \neq E, m$. Conclude that every $E \in (\frac{H-L}{c_L}, \frac{H-L}{c_H})$ corresponds to one such semi-pooling PBE where $w(m) = \frac{c_L H - c_H L - c_H c_L E}{c_L - c_H} \in (L, H)$, and in fact, there is $q \in (0, 1)$ such that $w(m) = qH + (1-q)L$ and a, b are such that $\frac{b}{a} = \frac{\alpha - \frac{\alpha}{q}}{1-\alpha}$.

Thus, such an equilibrium exists if and only if

$$E \leq \frac{H-L}{c_H},\ e \leq \frac{1}{c_L}[\frac{\alpha(1-p)H+(1-\alpha)L}{(1-\alpha)+\alpha(1-p)}],$$

and

$$\frac{(1-\alpha)(H-L)}{c_H} \leq E-e \leq \frac{(H-L)}{c_H}.$$

Next, consider the semi-pooling PBE where H plays some E with prob. 1 and L randomizes over E (prob. q) and e (prob. $1-q$) with $E > e \geq 0$. Again, assume $f(e') = 0,\ \forall e' \neq E, e$. It follows that $e = 0$! Such an equilibrium exists if and only if

$$\frac{\alpha(H-L)}{c_L} \leq E \leq \frac{H-L}{c_L}.$$

22. Another well-known signaling game is the following "game of beer and quiche," where A and B meet in a bar, and A may be weak (w) or strong (s), which is A's private information. The game proceeds as follows. A first decides to order either a beer (b) or a quiche (q), and upon observing A's order, B decides to or not to fight A. We assume that in the absence of B, A prefers beer (b) to quiche (q) if he is (s), otherwise he prefers (q) to (b). The prior beliefs of B are such that A is (s) with probability 0.9. Now the payoffs: if A orders and eats something he dislikes, he gets 0, or else he gets 1, and if B does not fight A, A gets an additional payoff of 2. On the other hand, B gets 1 if he has no chance to fight, gets 2 if he fights A and A is of the weak type, and gets zero if he fights A and A is of the strong type.

    This game has two pooling PBE's:
    (1) Equilibrium (B): Both types of A order a beer and B's strategy is to fight A if and only if he sees A order a quiche. What are the supporting beliefs? Let $f(s)$ =pro.(A is strong| A orders s), for all $s \in \{b, q\}$. Then of course $f(b) = 0.9$. Note that $s = q$ is a zero probability event. Recall that Bayes Law says

$$P(E|F)P(F) = P(E \bigcap F),$$

where E and F are two random events. From the probability theory, we know that for any two events C and D,

$$C \subset D \Rightarrow P(C) \leq P(D).$$

Thus we have

$$P(F) = 0 \Rightarrow P(E \bigcap F) = 0,$$

since $E \bigcap F \subset F$. Thus given $P(F) = 0$, Bayes Law requires

$$P(E|F) \cdot 0 = 0,$$

and hence $P(E|F)$ can be anything contained in $[0,1]$. Let E be the event that A is of the strong type, and F the event that B has observed that A ordered a quiche. We conclude that any $f(q) \in [0,1]$ will be consistent with Bayes Law in this case. We must find at least one $f(q) \in [0,1]$ so that the above strategy profile does constitute the two players' best responses against each other. Note that for B to fight A after seeing A order a quiche, it is necessary that

$$1 \leq f(q) \cdot 0 + [1 - f(q)] \cdot 2 \Rightarrow f(q) \leq \frac{1}{2}.$$

Now we show that given the beliefs $f(b) = 0.9$ and $f(q)$ being anything in $[0, \frac{1}{2}]$, the aforementioned A's and B's strategies are respectively the two players best responses. For A, if his type is (s), he gets $1 + 2 = 3$ if he orders a beer, and if he deviates and orders a quiche, then not only he eats something he hates but he also must fight B, yielding a payoff of $0 + 0 = 0$. Thus A will not deviate if he is of type (s). What if A is of type (w)? If he orders a beer, then he must eat something he hates, but the good news is that he can avoid fighting B, so that his payoff is $0 + 2 = 2$; and if he deviates and orders a quiche, then he will have to fight B, so that his payoff is $1 + 0 = 1$. We conclude that the weak type of A does not want to deviate either. What about B? We have shown that given $f(q) \leq \frac{1}{2}$, fighting A if A dares to order the quiche is really optimal for B. On the other hand, if A orders a beer, then since B expects both types of A to do so, ordering the beer really does not tell B anything new, and B's posterior beliefs are identical to his prior beliefs (A is of the strong type with prob. 0.9), and so not to fight A

is optimal for B.
To sum up, we have shown that the following is a PBE (check if it corresponds to our definition of a PBE!):
(i) The strong type of A orders a beer;
(ii) The weak type of A also orders a beer;
(iii) B's strategy must describe what he will do in every possible contingency: B will fight A if A ordered a quiche, and B will not fight A if A ordered a beer;
(iv) The supporting beliefs fully describe what B thinks of A *in every possible contingency*: B thinks that A is of the strong type with prob. 0.9 if he sees A order a beer; and B thinks that A is of the strong type with prob. $f(q)$ if he sees A order a quiche, where $f(q)$ is any real number contained in $[0, \frac{1}{2}]$.
(2) Equilibrium (Q): Both types of A order a quiche and B's strategy is to fight A if and only if he sees A order a beer. What are the supporting beliefs? Let $f(s)$ =pro.(A is strong| A orders s), for all $s \in \{b, q\}$. Then of course $f(q) = 0.9$. Now for $f(b)$ to induce B to fight A after seeing A order a beer, it is necessary that

$$1 \leq f(b) \cdot 0 + [1 - f(b)] \cdot 2 \Rightarrow f(b) \leq \frac{1}{2}.$$

Now we show that given the beliefs $f(q) = 0.9$ and $f(b)$ being anything in $[0, \frac{1}{2}]$, the aforementioned A's and B's strategies are respectively the two players best responses. For A, if his type is (w), he gets $1 + 2 = 3$ if he orders a quiche, and if he deviates and orders a beer, then not only he eats something he hates but he also must fight B, yielding a payoff of $0 + 0 = 0$. Thus A will not deviate if he is of type (w). What if A is of type (s)? If he orders a beer, then he must eat something he hates, but the good news is that he can avoid fighting with B, so that his payoff is $0+2 = 2$; and if he deviates and orders a beer, then he will have to fight B, so that his payoff is $1 + 0 = 1$. We conclude that the strong type of A does not want to deviate either. What about B? We have shown that given $f(b) \leq \frac{1}{2}$, fighting A if A dares to order the beer is really optimal for B. On the other hand, if A orders a quiche, then since B expects both types of A to do so in equilibrium, ordering the quiche really does not tell B anything new about A, and B's posterior beliefs are identical to his prior beliefs (A is of the strong type with

prob. 0.9), and so not to fight A is optimal for B.
To sum up, we have shown that the following is a PBE (check if it corresponds to our definition of a PBE!):
(i) The strong type of A orders a quiche;
(ii) The weak type of A also orders a quiche;
(iii) B's strategy must describe what he will do in every possible contingency: B will fight A if A ordered a beer, and B will not fight A if A ordered a quiche;
(iv) The supporting beliefs fully describe what B thinks of A *in every possible contingency*: B thinks that A is of the strong type with prob. 0.9 if he sees A order a quiche; and B thinks that A is of the strong type with prob. $f(b)$ if he sees A order a beer, where $f(b)$ is any real number contained in $[0, \frac{1}{2}]$.

23. In addition to the two pooling equilibria, there are no other PBE's for the game of beer and quiche. To see this, note first that in a separating PBE B must use only pure strategies, and chooses not to fight upon seeing the signal sent by the strong type of A. However, the weak type of A would prefer to send that signal as well, upsetting the supposed separating equilibrium.

    Next we explain why there can be no semi-pooling PBE for this game. First we claim that if a semi-pooling PBE exists, then at least one type of A must take a pure strategy in equilibrium. To see this, suppose that both types of A randomize in equilibrium with the strong type and the weak type of A choosing beer with probability p_s and p_w respectively, and we shall demonstrate a contradiction. If both types of A randomize over beer and quiche this way in equilibrium, then B will use a pure strategy either upon seeing beer or upon seeing quiche, for otherwise, B must have $f(b) = f(q) = 0.5$, or equivalently,

$$9p_s = p_w, \ 1 - p_w = 9 - 9p_s,$$

    which gives a contradiction. Thus B must use a pure strategy sometimes. However, if B chooses with probability one not to fight upon seeing either quiche or beer, then it is impossible that both types of A feel indifferent about quiche and beer: one type of A can get payoff 3 by sending that signal with probability one. Can it be possible that B chooses with probability one to fight upon seeing either quiche or beer? If this is the case, then one type of A will assign zero probability to such a signal: by sending the other signal, the probability that B fights is no greater than one, but that type of A gets payoff 1 by ordering something he genuinely likes. Hence again, it cannot happen that both types of A randomize in equilibrium.

Having established that in a semi-pooling PBE one type of A must use a pure strategy, now we continue to show that semi-pooling PBE cannot exist at all.

Let $x, y \in \{s, w\}$ be the types of A that use respectively a pure strategy and a mixed strategy in the semi-pooling PBE. Let $m, n \in$ {beer, quiche} be the signal sent by the type-x A and the signal that is sent by the type-y A exclusively. If such a semi-pooling PBE exists, then $f(n)$ equals either zero or one (depending on whether y is strong or weak), so that upon seeing signal n, B will use a pure strategy. Hence there are two possible cases.

First, suppose that B chooses to fight upon seeing signal n. Since the type-y A feels indifferent about signals m and n, and since B fights with a probability no greater than 1 after seeing signal m, the type-y A must genuinely like to order n rather than m; that is, $n =$ beer if y is strong, and $n =$ quiche if y is weak. This implies that if such a PBE exists, the type-y A's equilibrium payoff is 1. It follows that upon seeing signal m, B should fight with a probability less than 1 (but greater than zero), which in turn implies that $f(m) = 0.5$. The latter is possible, only if y is strong, with the probability of choosing signal m being $\frac{1}{9}$! However, this implies that upon seeing n, B knows that he is facing the strong-type A, and B's best response is not to fight, which contradicts the assumption that B chooses to fight upon seeing n.

Next, suppose that B chooses not to fight upon seeing signal n. This implies that the type-y A genuinely likes to order m better than n: $m =$ beer if y is strong, and $m =$ quiche if y is weak. This also implies that if such a PBE exists, the type-y A's equilibrium payoff must be equal to 2. For this type of A to feel indifferent about sending m and sending n, it is necessary that B fights with probability 0.5 upon seeing signal m. This implies that $f(m) = 0.5$, showing that y is strong, and it chooses

m with probability $\frac{1}{9}$. But then $n =$ quiche, and the type-x (where x is actually weak) would be better off sending n with probability one than sending the supposed equilibrium signal m! This is a contradiction.

Thus we have shown that no semi-pooling PBE can exist either.

24. **Cho-Kreps Refinement**. According to I.-K. Cho and David Kreps, some PBEs may involve *implausible* supporting beliefs and should be disregarded.

    To demonstrate Cho and Kreps' idea, consider Equilibrium (Q) in the previous section. There, B knows that ex-ante A may be strong with prob. 0.9 and A hates the quiche if he is strong, and yet B still thinks that A is more likely to be the weak type when he sees A deviate by ordering the beer. Consider the speech that the strong-type of A would have made to B if he were allowed to: I am having beer, so I am the strong type. To see this, note that if I were the weak type, I would have got 3 by having the quiche, and a weak type could never get a payoff of 3 by having beer, which is so *no matter how you may respond after the beer is ordered*! Moreover, if this speech can convince you that I am strong, then I expect you to not fight me, so that, as a strong type, I have the beer that I like and I do not have to fight you. In fact, I expect to get 3 if this communication works, and that is why I am having beer.....

    These two suppositions
    (i) the weak type of A is absolutely better off by not deviating; and
    (ii) if supposition (i) is accepted then the strong type of A is expected to be treated in a better way by B that justifies the strong type's deviation in the first place,
    comprises the so-called 'intuitive criterion.'

25. **Definition 2**. Those PBE's survive the intuitive criterion are called *intuitive equilibria.*

26. We now show that in the game of beer and quiche, only Equilibrium (B) is intuitive. By definition, a PBE is intuitive, if either (i) we cannot find a deviation which certain types of the informed player would never make; or (ii) we can find a deviation which certain types of the informed player would never make, but by restricting to the supporting

beliefs that assign zero prob. to these types, we still cannot find a type of the informed player that strictly prefers to deviate.
Take equilibrium (Q) for example. The equilibrium, by assumption, involves both types of A playing the strategy (q), but as we stated above, by having beer, the weak type of A can at best get the payoff 2 (which occurs if B decides not to fight A following an order of beer). Thus the weak type of A strictly prefer his equilibrium action to strategy (b). Now all reasonable beliefs should assign zero prob.'s to the weak type of A following an order of beer from A. That means that there is only one reasonable belief, the belief that assigns prob. 1 to the strong type of A after beer is ordered. Given this belief, B is expected to behave optimally, which is not to fight A. But then the strong type of A can get the payoff 3 by deviating from (q), while he gets 2 by ordering quiche. Thus the strong type of A would strictly prefer to deviate from (q), proving that equilibrium (Q) is not intuitive. This shows that both of the suppositions defined above hold for this equilibrium, so that this PBE fails the intuitive criterion, and it is not an intuitive equilibrium.

27. Next let us ask if equilibrium (B) is intuitive. Having observed the deviation (q), can we conclude that at least one type of A would never have done this? Apparently, the strong type has obtained a payoff of 3 on the equilibrium path, and by deviating and ordering (q), he could get no more than 2. Thus the strong type of A strictly prefers his equilibrium payoff to what he could get by deviating and ordering quiche. In this case, any reasonable beliefs after quiche is ordered should assign prob. 1 to the weak type of A. Now what is the best response of B given this reasonable belief? Of course B should fight A! But then, even the weak type of A could not gain by deviating from (b) to order (q)! To sum up, Cho-Kreps' first supposition holds but the second supposition fails for this PBE, and hence this PBE survives from Cho-Kreps' intuitive criterion. Thus this PBE is an intuitive equilibrium.

28. Now we show that there is also one single intuitive equilibrium in Spence signalling game. First we claim that all but one separating equilibria derived are unreasonable, because they are supported by some off-the-equilibrium beliefs that require the uninformed employer

to believe that the type-L worker has played weakly dominated strategies. More specifically, consider a separating equilibrium $(E, 0)$ where $E \in (\frac{H-L}{c_L}, \frac{H-L}{c_H}]$. Now consider what happens to the employer's posterior beliefs if a deviation to $E - \epsilon > \frac{H-L}{c_L}$ occurs. Note that the type-L worker can guarantee himself a payoff weakly higher than L by receiving no education: By receiving no education the worst thing can happen to the type-L worker is that the employer can recognize his identity for sure and pay him L accordingly. By receiving $(E - \epsilon)$ units of education, on the other hand, he gets at most

$$H - c_L(E - \epsilon) < H - c_L \cdot \frac{H - L}{c_L} = L,$$

and hence receiving $E - \epsilon$ units of education is a weakly dominated strategy for the type-L worker. Thus the only reasonable beliefs for the employer upon seeing the signal $E - \epsilon$ must be that the deviator is the type-H worker and pays the deviator the wage H! But then, the type-H worker would strictly prefer the signal of deviation $E - \epsilon$ to the equilibrium signal E, since signals E and $E - \epsilon$ yield the same wage, but receiving education is costly. Thus we have shown that all separating equilibria involving $E > \frac{H-L}{c_L}$ are unreasonable.

Next, we show that all the pooling equilibria of the Spence signalling game are vulnerable to the intuitive criterion. Suppose $e^* \leq \frac{\overline{w}-L}{c_L}$ is a pooling equilibrium. Suppose that the deviation

$$e' = \frac{H - \overline{w}}{c_L} + e^* + \epsilon$$

has occurred, where $\epsilon > 0$. First we verify the first supposition of the intuitive criterion. By sending e^*, the type-L worker gets

$$\overline{w} - c_L \cdot e^* > H - c_L \cdot e',$$

where the right-hand side is what he can get by sending e' if the employer is convinced that the deviator is the type-H worker. We conclude that only H could have possibly sent e'. But then, given this belief, we can verify the second supposition of the criterion; that is, the type-H worker would strictly prefer e' to e^* as long as $\epsilon > 0$ is small enough: In equilibrium, the type-H worker gets

$$\overline{w} - c_H e^*,$$

and by sending e' and receiving $w(e') = H$, his payoff becomes

$$w(e') - c_H e' = H - c_H[\frac{H - \overline{w}}{c_L} + e^*] - c_H\epsilon,$$

which, for $\epsilon > 0$ sufficiently small, is strictly greater than

$$H - c_L[\frac{H - \overline{w}}{c_L}] - c_H e^* = \overline{w} - c_H e^*.$$

Thus no pooling equilibria derived in class are robust against the intuitive criterion.

29. Finally, we consider semi-pooling equilibria. Fix any such equilibrium, there must be e which both types of worker send with a strictly positive probability in equilibrium. Note that given the type of the worker, her randomization must take place over those $(w(e), e)$ lying on the same indifference curve. We can show that all semi-pooling PBE's of the Spence signalling game are not intuitive. Suppose that e^* is the highest education level to which both types assign strictly positive prob.'s. Then L never randomizes over beyond e^* and $w(e^*) < H$. Consider the deviation $e' = \frac{H - w(e^*) + c_L e^*}{c_L} + \epsilon$. Show that L always gets a higher payoff in equilibrium than sending e'. Show that if treated fairly after sending e', H will send e' when $\epsilon > 0$ is small enough, as $w(e^*) - c_H e^* < H - c_H e'$ in this case.

30. **Example 3**. In the following we have a sequence of signalling games. In each game, player 1 has two equally likely types, denoted by t_1 and t_2, and given his type, player 1 must send a signal. There are three possible signals that player 1 can choose, which are m_1, m_2, and m_3. Upon seeing the signal selected by player 1, player 2 must then form a posterior belief about player 1's type, and given her belief, player 2 must take an action. There are three feasible actions for player 2, which are a_1, a_2, or a_3. The game ends after player 2 chooses her action.

    Each signalling game below is depicted by three tables. The k-th table gives the two players' payoffs in the event that player 1 chooses to send signal m_k; $k = 1, 2, 3$. As you can see, player 1's payoff not only depends on the two players' actions, it also depends on player 1's type.

For example, in the first table appearing in Problem 1 below, player 1 gets 2 and player 2 gets 1 if player 1 is of type t_1 and he sends signal m_1, and player 2 responds by taking action a_1; and in the second table, player 1 gets 0 and player gets 6 if player 1 is of type t_2 and he sends signal m_2, and player 2 responds by taking action a_3.[14]

(a) Find the PBEs:

m_1	a_1	a_2	a_3
t_1	$(2,1)$	$(2,0)$	$(0,2)$
t_2	$(1,3)$	$(2,0)$	$(2,1)$

m_2	a_1	a_2	a_3
t_1	$(3,1)$	$(1,0)$	$(0,0)$
t_2	$(2,1)$	$(0,0)$	$(0,6)$

m_3	a_1	a_2	a_3
t_1	$(1,2)$	$(1,1)$	$(3,0)$
t_2	$(0,2)$	$(3,1)$	$(1,1)$

Solution. There is a unique pooling equilibrium for this game. Let x, y, z be respectively type-1 informed player's equilibrium message, type-2 informed player's equilibrium message, and the uninformed player's equilibrium action upon seeing *the* message appearing in the pooling equilibrium.[15] Let μ_i be the uninformed

[14] Recall that a PBE is defined as:
(A) *a strategy for player 1*, which specifies one m_k for each type t_i;
(B) *a strategy for player 2*, which specifies one a_j for each m_k; and
(C) *a posterior belief for player 2*, which specifies one probability distribution on the set $\{t_1, t_2\}$ for each m_k, and moreover, (A), (B) and (C) must also satisfy:
(1) the a_j specified in (B) after player 2 sees m_k must be expected-utility-maximizing for player 2 given player 2's posterior belief specified in (C); and
(2) the m_k specified in (A) for type t_i must be expected-utility-maximizing for player 1 of type t_i, given that player 2's strategy is specified in (B).
This is a complicated definition. However, a definition is a definition. So, when you report that a PBE is found, you must make sure that you report (A),(B) and (C), and you must also verify that (A), (B) and (C) satisfy (1) and (2)!

[15] Thus z alone does not fully describe the uninformed player's *strategy.* The latter must specify the uninformed player's action following each m_i, whether or not m_i may appear in equilibrium with a positive probability. See requirement (B) in the preceding footnote!

player's posterior prob. for type-1 informed player upon seeing message m_i. Then a pure-strategy PBE is $\{(x, y, z), \mu_i,\ i = 1, 2, 3\}$. One can show that the unique PBE here is such that

$$x = y = m_1,\ z = a_1,\ \mu_1 = \frac{1}{2},\ \mu_2 \in [0, \frac{5}{6}],\ \mu_3 \in [0, 1].$$

(b) Find the PBEs:

m_1	a_1	a_2	a_3
t_1	$(1, 2)$	$(2, 2)$	$(0, 3)$
t_2	$(2, 2)$	$(1, 4)$	$(3, 2)$

m_2	a_1	a_2	a_3
t_1	$(1, 2)$	$(1, 1)$	$(2, 1)$
t_2	$(2, 2)$	$(0, 4)$	$(3, 1)$

m_3	a_1	a_2	a_3
t_1	$(3, 1)$	$(0, 0)$	$(2, 1)$
t_2	$(2, 2)$	$(0, 0)$	$(2, 1)$

Solution. This game has a unique pooling PBE:

$$x = y = m_3,\ z = a_1,\ \mu_2 \in [0, 1], \mu_1 \in [0, \frac{2}{3}],\ \mu_3 = \frac{1}{2}.$$

As ϵ^n goes to zero, the following sequence of totally mixed strategies for both types of informed player, $(\epsilon^n, \epsilon^n, 1 - 2\epsilon^n)$ produces a sequence of posterior beliefs converging to the uninformed's posterior beliefs in the pooling PBE.

(c) Find intuitive equilibria:

m_1	a_1	a_2	a_3
t_1	$(0, 3)$	$(2, 2)$	$(2, 1)$
t_2	$(1, 0)$	$(3, 2)$	$(2, 1)$

m_2	a_1	a_2	a_3
t_1	$(1, 2)$	$(2, 1)$	$(3, 0)$
t_2	$(0, 1)$	$(3, 1)$	$(2, 6)$

m_3	a_1	a_2	a_3
t_1	$(1,6)$	$(4,1)$	$(2,0)$
t_2	$(0,0)$	$(4,1)$	$(0,6)$

Solution. There are two pooling PBE's: (i) $x = y = m_1,\ z = a_2,\ \mu_1 = \frac{1}{2},\ \mu_2 \in [\frac{5}{7}, 1],\ \mu_3 \in [0,1]$; (ii) $x = y = m_2,\ z = a_3,\ \mu_1 \in [\frac{2}{3}, 1],\ \mu_2 = \frac{1}{2},\ \mu_3 \in [0,1]$. Only equilibrium (i) is intuitive. To see that (ii) is not. Note that t_2 can deviate and send m_1. The first supposition of Cho-Kreps is satisfied: t_1 would do strictly better by staying on the equilibrium (3) than deviating (≤ 2). Hence, the deviator is treated as t_2. But, given this, the uninformed player would choose a_2 that gives 3 to t_2, strictly higher than what the latter could get by staying on the equilibrium. Hence the second supposition is also satisfied.

(d) Find intuitive equilibria:

m_1	a_1	a_2	a_3
t_1	$(4,0)$	$(0,3)$	$(0,4)$
t_2	$(3,4)$	$(3,3)$	$(1,0)$

m_2	a_1	a_2	a_3
t_1	$(2,0)$	$(0,3)$	$(3,2)$
t_2	$(0,3)$	$(0,0)$	$(2,2)$

m_3	a_1	a_2	a_3
t_1	$(2,3)$	$(1,0)$	$(1,2)$
t_2	$(4,3)$	$(0,4)$	$(3,0)$

Solution. Once again, this game has two pooling PBE's: (i) $x = y = m_2,\ z = a_3,\ \mu_1 \in [\frac{3}{4}, 1],\ \mu_2 = \frac{1}{2},\ \mu_3 \in [0, \frac{1}{4}]$; (ii) $x = y = m_3,\ z = a_1,\ \mu_1 \in [\frac{1}{4}, 1],\ \mu_2 \in [0, \frac{1}{3}] \bigcup [\frac{2}{3}, 1],\ \mu_3 = \frac{1}{2}$. Both satisfy intuitive criterion. To see this, take (i) for example. Consider the deviation with m_1. But this fails the first supposition. Next consider the deviation with m_3. The first supposition holds as t_1 could do strictly better by staying in equilibrium. But, the second supposition fails because, given that the deviator in this case must be t_2, the uninformed player would choose a_2 that makes t_2 worse off than in equilibrium. The reasoning that (ii) is also intuitive is similar.

(e) Find the intuitive equilibria:

m_1	a_1	a_2	a_3
t_1	$(4,1)$	$(2,4)$	$(1,5)$
t_2	$(5,6)$	$(2,5)$	$(2,2)$

m_2	a_1	a_2	a_3
t_1	$(1,3)$	$(3,1)$	$(4,2)$
t_2	$(1,3)$	$(1,4)$	$(3,3)$

m_3	a_1	a_2	a_3
t_1	$(3,3)$	$(2,0)$	$(1,4)$
t_2	$(3,4)$	$(1,5)$	$(0,1)$

Solution. Once again, this game has two pooling PBE's: (i) $x = y = m_1,\ z = a_2,\ \mu_1 = \frac{1}{2},\ \mu_2 \in [\frac{1}{3}, 1],\ \mu_3 \in [0, \frac{1}{4}] \bigcup [\frac{3}{4}, 1]$; (ii) $x = y = m_3,\ z = a_1,\ \mu_1 \in [\frac{1}{4}, 1],\ \mu_2 \in [0, 1],\ \mu_3 = \frac{1}{2}$. Both satisfy intuitive criterion.

31. We shall now make some modifications of the game of beer and quiche, and look for the PBE's of the new game. Assume that there are three types of A, the strong (s), the weak (w), and the crazy (c), where the crazy A fights just like the strong A (and so B would rather not fight the crazy A), and he enjoys fighting: the crazy A gets 2 if he can fight B, and zero if otherwise. Suppose that the prior beliefs of B are such that the three types of A are equally likely. Recall that all A can do in this game is ordering food, and so the crazy A cannot ask B for a fight. (You should refer to the original game of beer and quiche for the payoffs of respectively the strong A, the weak A, and B.)
(i) First suppose that the crazy A and the strong A have the same preferences for beer and quiche. Is there a PBE where B adopts a pure strategy in equilibrium? If your answer is yes, is the equilibrium unique?
(ii) Suppose instead that the crazy A and the weak A have the same preferences for beer and quiche. Is there a PBE where B adopts a pure strategy in equilibrium? If your answer is yes, is the equilibrium

unique?

Solution. Consider part (i). Let F, N, b, q stand for the strategies "fight," "not fight," "beer," and "quiche." Let s, w, and c represent the three types of A. Note that B has 4 possible pure strategies:

$$\begin{pmatrix} b \to F \\ q \to F \end{pmatrix}, \begin{pmatrix} b \to N \\ q \to N \end{pmatrix}, \begin{pmatrix} b \to F \\ q \to N \end{pmatrix}, \begin{pmatrix} b \to N \\ q \to N \end{pmatrix}.$$

We claim that none of these strategies can be consistent with a PBE.

First suppose that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to F \\ q \to F \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to b \\ w \to q \\ c \to b \end{pmatrix},$$

but then seeing b player B would be better off responding by strategy N, a contradiction. Suppose instead that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to N \\ q \to N \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to b \\ w \to q \\ c \to b \end{pmatrix},$$

but then seeing q player B would be better off responding by strategy F, another contradiction. Now suppose that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to F \\ q \to N \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to q \\ w \to q \\ c \to b \end{pmatrix},$$

but then seeing b player B would be better off responding by strategy N, again a contradiction. Finally, suppose that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to N \\ q \to F \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to b \\ w \to b \\ c \to q \end{pmatrix},$$

but then seeing q player B would be better off responding by strategy N, also a contradiction. Hence we conclude that there is no PBE where B adopts a pure strategy in equilibrium. This finishes part (i).

Now consider part (ii). We claim that the three pure strategies

$$\begin{pmatrix} b \to F \\ q \to F \end{pmatrix}, \begin{pmatrix} b \to F \\ q \to N \end{pmatrix}, \begin{pmatrix} b \to N \\ q \to N \end{pmatrix},$$

cannot be consistent with a PBE. The reasoning is similar to that in part (i). First suppose that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to F \\ q \to F \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to b \\ w \to q \\ c \to q \end{pmatrix},$$

but then seeing b player B would be better off responding by strategy N, a contradiction. Now suppose that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to F \\ q \to N \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to q \\ w \to q \\ c \to b \end{pmatrix},$$

but then seeing b player B would be better off responding by strategy N, another a contradiction. Finally, suppose that in equilibrium B adopts the strategy

$$\begin{pmatrix} b \to N \\ q \to F \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to b \\ w \to b \\ c \to q \end{pmatrix},$$

but then seeing q player B would be better off responding by strategy N, also a contradiction.

Thus we are left with one last option,

$$\begin{pmatrix} b \to N \\ q \to N \end{pmatrix}.$$

Rationally expecting B's strategy, A's best response is

$$\begin{pmatrix} s \to b \\ w \to q \\ c \to q \end{pmatrix},$$

which does justify B's adopting the strategy

$$\begin{pmatrix} b \to N \\ q \to N \end{pmatrix}.$$

Hence there is a unique PBE where the BE part is such that B adopts the strategy

$$\begin{pmatrix} b \to N \\ q \to N \end{pmatrix}$$

and A adopts the strategy

$$\begin{pmatrix} s \to b \\ w \to q \\ c \to q \end{pmatrix}.$$

To complete our descriptions for this PBE, we need to specify the supporting beliefs for B. Note that there are no *off-the-equilibrium signals* in this PBE; that is, both b and q are expected to be seen by B with a positive probability. Thus the supporting beliefs are determined completely by Bayes Law applying to A's equilibrium strategy. The unique set of supporting beliefs is such that

$$\text{pro.}(s|b) = 1, \ \text{pro.}(w|q) = \text{pro.}(c|q) = \frac{1}{2}.$$

32. **Cheap Talks**. Consider a seller and a buyer interested in the transaction of an indivisible good X. The cost to produce good X may be 5 dollars or 1 dollar, with probabilities s and $1-s$ respectively, which is the seller's private information. The buyer's valuation for good X may be 10 dollars or 3 dollars, with probabilities b and $1-b$, which is the buyer's private information. It is common knowledge that the seller's cost and the buyer's valuation are statistically independent. Both the seller and the buyer are risk neutral without time preferences, and they seek to maximize the expected surplus; that is, both parties get zero payoff if trade does not occur, and the the seller's payoff and the buyer's payoff are respectively $p-c$ and $v-p$ when trade occurs at price p with the seller's cost and the buyer's valuation being respectively, c and v. We shall assume that $s > \frac{1}{2}$.

    Consider the following extensive game.

    - At time 0, the seller privately sees her cost, and the buyer privately sees his valuation.
    - At time 1, the seller can say "I will never re-imburse the buyer's transportation cost" or "I might re-imburse the buyer's transportation cost" in front of the public (including the buyer). Note that the seller will not be held responsible for what she says at time 1, and making a statement at time 1 does not incur any cost to the seller.
    - At time 2, the seller and the buyer simultaneously decide whether they should go to a marketplace and trade. Going to the marketplace incurs a transportation cost 0.5 dollars for the buyer but no transportation cost for the seller.

- At time 2, if both parties appear at the marketplace, then the seller's cost and the buyer's valuation become their common knowledge, and they must simultaneously submit bids p_s and p_b. Trade will not occur if $p_s > p_b$, and trade will occur at the price $\frac{p_s+p_b}{2}$ if $p_s \leq p_b$. Here, assume that the seller and the buyer only believe in pure-strategy Nash equilibria, and if more than 1 pure-strategy Nash equilibrium exists, then both the buyer and the seller believe that all pure-strategy Nash equilibria are equally likely to occur.

(i) Show that there exists a perfect Bayesian equilibrium where the seller feels indifferent about what to say at time 1, and the buyer goes to the marketplace at time 2 if and only if his valuation is 10 dollars. In equilibrium the seller never reimburses the buyer's transportation cost. Write down completely the supporting beliefs.

(ii) Show that there exists a perfect Bayesian equilibrium where the seller says "I will never re-imburse the buyer's transportation cost" if her cost is 5 dollars and "I might re-imburse the buyer's transportation cost" if her cost is 1 dollar, and where at time 2 the buyer always vists the marketplace if his valuation is 10 dollars, and he visits the marketplace when his valuation is 3 dollars only after the seller says "I might re-imburse the buyer's transportation cost." In equilibrium the seller never reimburses the buyer's transportation cost. Don't forget to write down completely the supporting beliefs.

(iii) Explain why there exists a separating PBE in part (ii) even though signaling (making a statement) incurs no cost to either type of the seller. In particular, is the Spence-Mirrlees sorting condition satisfied?
Solution. Consider part (i). Since the buyer's equilibrium time-2 strategy does not depend on what the seller says at time 1, the seller feels indifferent about what to say at time 1 also. Thus nobody takes the seller's talk at time 1 seriously. Such an equilibrium is called a "babbling" equilibrium.

We claim that in this babbling equilibrium the buyer should go to the marketplace if and only if his valuation is 10 dollars.

- When the buyer's valuation is 3 dollars, by going to the marketplace, he gets expected payoff

$$-0.5 + s \times 0 + (1-s) \times \int_1^3 (3-t)\frac{1}{3-1}dt = \frac{1}{2} - s < 0,$$

where zero is the buyer's payoff if he chooses not to go to the marketplace, and $(3-t)$ is the buyer's payoff obtained in the t-equilibrium of the subgame where the seller and the buyer submit bids simultaneously to determine the transaction price, and in equilibrium they both submit the same $t \in [1,3]$. We have assumed that these t-equilibria are all equally likely to arise.

- On the other hand, when the buyer's valuation is 10 dollars, by going to the marketplace, he gets expected payoff

$$-0.5 + s \times \int_5^{10} (10-t)\frac{1}{10-5}dt + (1-s) \times \int_1^{10} (10-t)\frac{1}{10-1}dt$$
$$= \frac{1}{2}[5 \times s + 9(1-s) - 1] > 0,$$

and so this type of buyer will go to the marketplace in equilibrium.

Since the seller cannot be held responsible for what she says at time 1, she never re-imburses the buyer's transportation cost in equilibrium. This finishes the proofs for the assertions about the babbling equilibrium.

Next, consider part (ii). We verify this equilibrium in a few steps.

(a) In equilibrium, the buyer believes that the seller says "I might re-imburse the buyer's transportation cost" at time 1 if and only if the seller's cost is 1 dollar. Thus the buyer, regardless of his valuation for the product, should go to the marketplace after the seller makes that statement. To see this, we only need to show that going to the marketplace is optimal for the buyer when his valuation is 3 dollars, for the gain from going to the marketplace is even higher when the buyer's valuation is 10 dollars. Now observe that, when the buyer's valuation is 3 dollars, going to the marketplace generates an expected payoff

$$-0.5 + 1 \times \int_1^3 (3-t)\frac{1}{3-1}dt = \frac{1}{2} > 0,$$

where the “1” in front of the integral is the posterior probability that the buyer assigns to the event that the seller's cost is equal to 1 dollar. This posterior probability equals 1, because there is a separating equilibrium at time 1, where the statement “I might re-imburse the buyer's transportation cost” (which is a lie!) signals that the seller's cost is 1 dollar instead of 5 dollars.

(b) Next, observe that the buyer should go to the marketplace after the seller says “I will never re-imburse the buyer's transportation cost” if and only if the buyer's valuation is 10 dollars. This happens because for the buyer whose valuation is 3 dollars,

$$-0.5 + 1 \times 0 + 0 \times \int_1^3 (3-t)\frac{1}{3-1}dt < 0;$$

and for the buyer whose valuation is 10 dollars,

$$-0.5 + 1 \times \int_5^{10} (10-t)\frac{1}{10-5}dt + 0 \times \int_1^{10} (10-t)\frac{1}{10-1}dt > 0.$$

(c) Now we examine the seller's ICs. For the seller whose cost is 1 dollar, saying “I might re-imburse the buyer's transportation cost” at time 1 is optimal, for by the preceding steps (a) and (b), that will induce both types of the buyer to show up at the marketplace, whereas saying “I will never re-imburse the buyer's transportation cost” will scare the buyer whose valuation is 3 dollars away. Thus the seller will not deviate from making her equilibrium statement if her cost is 1 dollar.

(d) Now, for the seller whose cost is 5 dollars, saying “I might re-imburse the buyer's transportation cost” and saying “I will never re-imburse the buyer's transportation cost” are equally good, because by the preceding steps (a) and (b), the buyer whose valuation is 10 dollars will always go to the marketplace at time 2, and a seller whose cost is 5 dollars only wants to trade with the buyer whose valuation is 10 dollars. Thus the seller will not deviate from making her equilibrium statement if her cost is 5 dollars.

(e) The above steps together prove that the strategies specified for the seller and the buyer do constitute a separating PBE, with the supporting beliefs being such that the buyer believes that the seller's cost is 1 dollar if she says "I might re-imburse the buyer's transportation cost" at time 1, and the buyer believes that the seller's cost is 5 dollars if she says otherwise at time 1.

Finally, consider part (iii). This is an example of a cheap-talk model. In this model, talk is cheap because it does not incur signalling costs, and the signal-sender does not have to be responsible for the content of the talk. Part (i) shows that in such a model, there is always a "babbling" equilibrium in which nobody takes the cheap talk seriously. Part (ii) shows that, however, there may exist an equilibrium which cannot arise if we delete the "talking stage." In our problem, it is a separating equilibrium, where the cheap talk is taken seriously by everyone. Although signalling per se is costless in a cheap-talk game, there is an endogenous Spence-Mirrlees condition that arises from the buyer's equilibrium response to the talk. It is this endogenous sorting condition that sustains the separating PBE in part (ii).[16]

33. **Information Cascade**. A sequence of agents have the same access to an investment opportunity. At date $n \in \mathbf{Z}_+$, agent n after observing the preceding agents' investment decisions, must decide his own. The agent can decide either to invest (I) or not to invest (N), with investment incurring a cost of one dollar. The prior beliefs of all agents are such that the project may succeed with prob. $\frac{1}{2}$ and when it does, it generates 2 dollars. (It generates nothing if it fails.) Note that being able to observe the preceding agents' investment decisions is valuable because when one agent's investment succeeds, all other agents' invest-

[16]The cheap-talk games were first studied by Crawford and Sobel (1982, Strategic Information Transmission,*Econometria*). Battaglini (2002, Multiple Referrals and Multidimensional Cheap Talk, *Econometrica*) shows that with multiple talk-givers (unlike in our problem, where the seller is the only one giving the talk), it is a very general result that cheap talk gives rise to a separating equilibrium. For an application of cheap-talk games to marketing, see Li (2005, Cheap Talk and Bogus Network Externalities in the Emerging Technology Market, *Marketing Science*). There, the author shows that revenue sharing contract instead of linear whole pricing is consistent with a separating PBE where a manufacturer tells its retailer the former's private information about how much the period-1 demand may benefit the period-2 demand (a positive network effect), and in equilibrium the manufacturer tells the truth.

ments will succeed also. In other words, we assume perfect correlations among these agents' investment outcomes.

In addition to the preceding agents' behavior, agent n also receives a private signal $\tilde{s}_n$ regarding the likelihood of success, and the outcome of $\tilde{s}_n$ is either H or L. The collection of signals $\{\tilde{s}_n; n \in \mathbf{Z}_+\}$ are independently and identically distributed. Let

$$\frac{1}{2} \leq p = \text{prob.}(\tilde{s}_n = H|\text{the project will succeed})$$

$$= \text{prob.}(\tilde{s}_n = L|\text{the project will fail}).$$

Note that when $p = \frac{1}{2}$, the signal is completely uninformative about whether the investment will succeed or not. On the other hand, when $p = 1$, the signal unambiguously tells the agent whether the investment will succeed or not. The purpose of this exercise is to show that it is fairly likely that an agent disregards his own private signal and simply follows what the preceding agents have done, and when this happens to some agent i, all the agents following agent i will mimic the preceding agents also. We will show that this happens quite often, but such a phenomenon may be inefficient: the resulting investment decisions can be wrong, not just from an ex-post perspective, but also from an ex-ante perspective, if these agents *could* pool their private signals together instead of observing one another's decisions.

We say that a cascade happens if for some $n \in \mathbf{Z}_+$, it becomes common knowledge at date n that all agents m, $m \geq n$, will make the same investment decision. We call the cascade an "up" cascade if the common decision is to invest; or else, a "down" cascade.
(i) Let U_n be the prob. that an up cascade starts at date n, where n is even. Show that

$$U_n = \frac{1 - (p - p^2)^{\frac{n}{2}}}{2}.$$

(ii) Deduce from (i) that the closer p is to $\frac{1}{2}$, the later the cascade is likely to start.
(iii) Let E_n be the prob. that no cascades start before date n, where n is even. Show that E_n falls exponentially with n.

(iv) A cascade is called "correct," if the true state is "success" and it is an "up" cascade or the true state is "failure" and it is a "down" cascade. Show that

$$\text{prob.}(U_n|\text{success}) = \frac{p(p+1)[1-(p-p^2)^{\frac{n}{2}}]}{2(1-p+p^2)},$$

which is increasing in p and n.

(v) Show by (iv) that the prob. that an incorrect cascade starts at date n, where n is even, is

$$\frac{(p-2)(p-1)[1-(p-p^2)^{\frac{n}{2}}]}{2(1-p+p^2)},$$

which is remarkably high even if p is far from $\frac{1}{2}$.

Solution Conditional on the state being success, agent 2 will invest

- with probability $\frac{1}{2}$ if either his own signal is L and the first agent invested (revealing fully that the first agent's signal is H) or his signal is H and the first agent did not invest;
- with probability one, if his signal is H and the first agent invested;
- with probability zero, if his signal is L and the first agent did not invest.

Thus, conditional on the true state being success, the prob. that cascades have not occurred before the third agent enters is $2 \cdot \frac{p(1-p)}{2} = p(1-p)$, and the prob. that an up cascade has occurred before the third agent enters is $p^2 + \frac{p(1-p)}{2} = \frac{p(p+1)}{2}$, and the prob. that a down cascade has occurred before the third agent enters is $(1-p)^2 + \frac{p(1-p)}{2} = \frac{(2-p)(1-p)}{2}$. Next, observe that the probabilities for respectively up and down cascades to occur before agent 3 enters conditional on the true state being failure are exactly the reverse of the probabilities obtained conditional on the true state being success. Thus the ex-ante probabilities for the events of an up and a down cascades before agent 3 enters are both

$$\frac{1}{2}[\frac{p(p+1)}{2} + \frac{(2-p)(1-p)}{2}] = \frac{p^2-p+1}{2},$$

and the prob. of no cascade before agent 3 enters is $p(1-p)$. Observe that if cascades have not occurred before the $(2n+1)$-th agent enters, then the decision problem facing agent $(2n+1)$ is identical to that of agent 1. It follows that the ex-ante prob. of no cascades occur before agent $(2n+1)$ enters is $p^n(1-p)^n$. By symmetry, the probability that an up cascade occurs before agent $(2n+1)$ enters is then $\frac{1-p^n(1-p)^n}{2}$, and this is also the probability that a down cascade occurs before agent $(2n+1)$ enters.

Observe that the prob. $[p(1-p)]^n$ falls exponentially in n. Note also that given n, $[p(1-p)]^n$ attains its maximum at $p=\frac{1}{2}$ (where the signal is completely uninformative). We conclude that for all n, the closer p is to $\frac{1}{2}$, the probability for the event that no cascades have occurred before date n becomes higher. Equivalently, the more precise the signal is (the closer p is to one), the earlier a cascade may occur. The idea is that in this case the earlier agents' investment behavior has a higher information value for the statistical inference problem faced by a subsequent agent. This finishes parts (i)-(iii). The rest two assertions are left to the reader.

34. **Reputation Games**. Now we introduce *reputation games.* In the game *chain-store paradox*, there are two players E and I, the entrant and the incumbent. E can choose to be "In" or "Out." If E stays out, the game is over and E and I get respectively 0 and $\frac{3}{4}$. If E enters, I can either "prey" (or predate) or "acquiesce" (or accomodate). If I preys, both E and I get -1, or else, I gets 0 and E gets 1.

    (i) Find the unique SPNE.

    (ii) Now we introduce incomplete information by assuming that E thinks that with probability x, I may actually get $\frac{1}{2}$ when he preys rather than -1. This type of I (who enjoys preying) is called "crazy." E's problem is that he cannot tell exactly if I is crazy or not. Find the BE for this game.

    (iii) Now, we introduce dynamics. Suppose the game in (ii) is repeated one more time. In each period the incumbent is facing a new entrant, but all entrants can observe whether or not the incumbent has preyed before. The incumbent seeks to maximize the sum of its temporal payoffs without discounting. Find the PBE for this game.

(iv) Solve for the PBE for the repeated chain-store paradox, assuming that the stage game is repeated for 3 times.[17] [18]

Solution. Let x_i be the probability that the i-th entrant (denoted by E_i from now on) assigns to the event that the incumbent is crazy when E_i is about to decide whether to enter. Note that x_1 is the prior belief held by all entrants, but for $i > 1$, x_i is the posterior belief that is derived in equilibrium using Bayes' Law whenever possible. Let i^* be the smallest i such that the type of the incumbent becomes publicly known

[17] The idea here is that the sane type of incumbent may want to pool with the crazy type in order to convince the subsequent entrants that entry will result in them being preyed, for if this works, then the short-term loss resulting from preying may be compensated by the increase in the long-term gain resulting from the subsequent entrants choosing to stay out. In equilibrium, this incentive problem on the part of the sane type of incumbent is of course correctly expected by the entrants, but just like in a game of signal-jamming, the sane type of incumbent may be unable to commit not to manipulate the entrants' beliefs. Consequently, even the sane type of incumbent may prey (at least in the earlier periods) in equilibrium. Part (iii) shows that this will actually happen, only if there are sufficiently many entrants waiting to enter—the one-time loss from preying is more likely to be compensated by the long-term gain if preying convinces many entrants that staying out is their best option.

[18] While the sane type of incumbent may benefit from pooling with the crazy type, note that the crazy type of incumbent does not want to pool with the sane type. Why? Observe that $\frac{3}{4} > \frac{1}{2} > 0$, which implies that the crazy type of incumbent prefers that the entrants stay out, but if this is not possible, then it prefers preying to accommodating after entry takes place. Since preying fulfills short-term payoff maximization, and it raises the entrants' posterior belief that the incumbent may be crazy (thereby raising the likelihood that these entrants choose to stay out), the crazy type of incumbent always prefers preying to accommodating after entry takes place. Note that the sane type of incumbent's pooling with the crazy type makes preying a noisy signal for the event that the incumbent is actually crazy, and hence it hurts the crazy type of incumbent. Put another way, if the sane type of incumbent were willing to always accommodate after entry takes place, then the crazy type of incumbents would become better off, because preying will perfectly reveal its type, and will ensure that all the subsequent entrants choose to stay out.

in equilibrium at the stage where the incumbent interacts with E_i. We write $i^* = +\infty$ if the incumbent's type remains its private information at the time that the game ends.

Consider part (i). We have a dynamic game with complete information, and hence we shall look for an SPNE. Backward induction implies that in the unique SPNE the sane-type incumbent had better accommodating the entrant after entry takes place (as $0 > -1$), and hence the entrant will enter for sure in equilibrium (because $1 > 0$).

Consider part (ii). The game in part (ii) is a static game with incomplete information because the uninformed entrant must finish its move before seeing the informed incumbent's action. Thus we look for a BE. Backward induction says that upon seeing entry take place, the sane-type incumbent will accommodate (as $0 > -1$), but the crazy-type incumbent will prey (as $\frac{1}{2} > 0$). Thus the entrant would get 0 by staying out, and it would get the expected payoff

$$x_1 \cdot (-1) + (1 - x_1) \cdot 1 = 1 - 2x_1$$

by entering the industry. We conclude that in the unique Bayesian equilibrium, the entrant would choose to enter if and only if $x_1 < \frac{1}{2}$.

Consider part (iii). The game in part (iii) is a dynamic game with incomplete information because E_2 can observe the informed incumbent's interaction with E_1 before deciding whether to enter or to stay out. Thus we look for a PBE.

Consider the subgame where E_1 has just entered and the incumbent must decide whether to prey or to accommodate E_1. First consider the sane-type incumbent's decision. Preying E_1 generates -1 to the sane-type incumbent immediately, and it may at best generate $\frac{3}{4}$ to the sane-type incumbent in the next stage, if E_2 would then choose to stay out. Hence by preying E_1 the sane-type incumbent can get no more than

$$-1 + \frac{3}{4}.$$

By accommodating E_1 instead, the sane incumbent gets 0 immediately, and in the next stage the sane-type incumbent can get $\frac{3}{4}$ if E_2 would stay out or 0 by accommodating E_2 if E_2 would enter. Hence by accommodating E_1 the sane-type incumbent can get no less than

$$0 + 0.$$

We thus conclude that sane-type incumbent would accomodate E_1.

Now consider the crazy-type incumbent's decision. Recall that the crazy-type incumbent would get $\frac{3}{4}$ if an entrant stays out, and $\frac{1}{2}$ if the entrant gets in and is preyed, and zero if the entrant gets in and is not preyed. Thus expecting the sane-type incumbent to accommodate E_1, the crazy-type incumbent's best response is to prey E_1: given that E_1 has entered, preying generates an immediate payoff $\frac{1}{2}$, and it proves to E_2 that the incumbent is crazy, which by part (ii) and by E_2's posterior belief $x_2 = 1$ upon seeing E_1 being preyed, would induce E_2 to stay out, thereby generating $\frac{3}{4}$ for the crazy-type incumbent in the second stage.

Thus we conclude that the subgame where E_1 has just entered has a unique separating equilibrium, where the sane incumbent would accommodate E_1 and the crazy incumbent would prey E_1.

Now, should E_1 enter or not? By staying out, it gets zero, and by entering, it may be preyed with probability x_1. Thus E_1 would stay out if $x_1 \geq \frac{1}{2}$. Thus $i^* = 1$ if $x_1 < \frac{1}{2}$, and $i^* = +\infty$ if $x_1 \geq \frac{1}{2}$. This completes our derivation for the unique (separating) PBE for part (iii).

Finally, consider part (iv). First consider the case where $x_1 \geq \frac{1}{2}$. Consider the subgame where E_1 has just entered and the incumbent must decide whether to prey or to accomodate E_1.

- Can there be an equilibrium where both types of the incumbent choose to prey E_1? In such an equilibrium, following E_1 being preyed, we must have $x_2 = x_1 \geq 0.5$, and hence E_2 and E_3 would both stay out, implying that the sane incumbent and the crazy incumbent get respectively $-1+\frac{3}{4}+\frac{3}{4}$ and $\frac{1}{2}+\frac{3}{4}+\frac{3}{4}$ in equilibrium. Upon seeing E_1 being accommodated, the only intuitive posterior belief is $x_2 = 0$, and E_2 will enter for sure. Thus by accommodating E_1, the sane incumbent and the crazy incumbent would get respectively $0+0+0$ and $0+\frac{1}{2}+\frac{3}{4}$, proving that neither of them would deviate even under the intuitive belief. Thus such a pooling equilibrium does exist, and we have shown that it is a Cho-Kreps intuitive equilibrium.

- Can there be an equilibrium where both types of the incumbent choose to accommodate E_1? In such an equilibrium, following E_1 being accommodated, we must have $x_2 = x_1 \geq 0.5$, and hence E_2 and E_3 would both stay out, implying that the sane incumbent and the crazy incumbent both get $0+\frac{3}{4}+\frac{3}{4}$ in equilibrium. Upon seeing E_1 being preyed, the only intuitive posterior belief is $x_2 = 1$, and E_2 and E_3 would both stay out. Thus by preying E_1, the sane incumbent and the crazy incumbent would get respectively $-1+\frac{3}{4}+\frac{3}{4}$ and $\frac{1}{2}+\frac{3}{4}+\frac{3}{4}$, proving that the crazy incumbent would strictly like to deviate under this intuitive belief! Thus no such intuitive equilibrium can exist.

  Can there still exist such a non-intuitive pooling PBE? Note that by deviating and preying E_1, the crazy incumbent's payoff would be at least $\frac{1}{2}+\frac{1}{2}+\frac{3}{4}$ if $x_2 < \frac{1}{2}$ and $\frac{1}{2}+\frac{3}{4}+\frac{3}{4}$ if $x_2 \geq \frac{1}{2}$. Thus such a non-intuitive pooling PBE cannot exist either.

- Can there be a separating equilibrium where the two types of incumbent choose to treat E_1 differently? In such an equilibrium, the sane incumbent's equilibrium payoff is at most $0+0+0$, but by mimicking the crazy incumbent's behavior towards E_1, the sane incumbent would get at least $-1+\frac{3}{4}+\frac{3}{4}$, proving that no separating equilibrium can exist.

- Can there be an equilibrium where the sane incumbent randomizes over preying and accommodating E_1? Let α and β be respectively the probabilities that E_2 may stay out and that E_3 may stay out upon seeing E_1 being preyed, and let α' and β' be respectively the probabilities that E_2 may stay out and that E_3 may stay out upon seeing E_1 being not preyed. Since the sane incumbent must feel indifferent about preying or accommodating E_1, we must have

$$-1 + \frac{3}{4}(\alpha + \beta) = 0 + \frac{3}{4}(\alpha' + \beta'),$$

$$\Rightarrow \alpha + \beta - \alpha' - \beta' = \frac{4}{3},$$

so that the crazy incumbent must strictly prefer preying to accommodating E_1:

$$\{\frac{1}{2} + \frac{3}{4}(\alpha + \beta) + \frac{1}{2}[(1-\alpha) + (1-\beta)]\}$$
$$-\{0 + \frac{3}{4}(\alpha' + \beta') + \frac{1}{2}[(1-\alpha') + (1-\beta')]\}$$
$$= \frac{1}{2} + \frac{1}{4}[\alpha + \beta - \alpha' - \beta'] = \frac{5}{6} > 0.$$

Thus an equilibrium where the sane incumbent randomizes over preying and accommodating E_1 must be a semi-separating equilibrium, where the sane incumbent may prey E_1 with probability y and the crazy incumbent must prey E_1 with probability one. It follows that E_2's posterior belief is $x_2 = 0$ upon seeing E_1 being accommodated and $x_2 > x_1 \geq \frac{1}{2}$ upon seeing E_1 being preyed. Hence E_2 will enter if and only if seeing E_1 being accommodated. This implies that by preying E_1 the sane incumbent's payoff is $-1 + \frac{3}{4} + \frac{3}{4}$ and by accommodating E_1 the sane incumbent's payoff becomes $0 + 0 + 0$, proving that the sane incumbent does not feel indifferent about preying or accommodating E_1, which is a contradiction. Thus there does not exist an equilibrium where the sane incumbent randomizes over preying and accommodating E_1.

- By a similar argument, one can show that there is no equilibrium where only the sane type adopts a pure strategy upon seeing entry by E_1.

The conclusion here is that when $x_1 \geq \frac{1}{2}$ there is a unique intuitive equilibrium for the subgame where E_1 has just entered. It is a pooling equilibrium, where both types of the incumbent choose to prey E_1. In equilibrium, therefore, E_1 would stay out for sure, which results in $x_2 = x_1 \geq \frac{1}{2}$, implying that E_2 would also stay out for sure. This in turn implies that $x_3 = x_2 = x_1 \geq \frac{1}{2}$, and hence E_3 would also stay out. We conclude that when $x_1 \geq \frac{1}{2}$, $i^* = +\infty$ in equilibrium.

Next consider the case $x_1 < \frac{1}{2}$. Consider the subgame where E_1 has just entered and the incumbent must decide whether to prey or to accommodate E_1.

- It is easy to show that there is no pooling equilibrium for this subgame where both types of the incumbent choose to prey E_1. Indeed, if such an equilibrium exists, then upon seeing E_1 being preyed, the other two entrants' belief would be $x_2 = x_1 < \frac{1}{2}$, and hence E_2 would get in, which implies that, by part (iii) above, the sane-type incumbent would get $-1 + 0 + 0$ by preying E_1. Note that the sane-type incumbent can guarantee itself a payoff of at least $0 + 0 + 0$ by accommodating each and every entrant.

- Can there be an equilibrium where only the sane incumbent preys E_1 for sure? In such an equilibrium, by the discussion in the preceding paragraph the crazy incumbent must accommodate E_1 with a positive probability. Thus seeing E_1 being accommodated, E_2's posterior belief is such that $x_2 = 1$, and hence both E_2 and E_3 would stay out, by part (iii) above. This implies that the sane incumbent would strictly prefer to deviate and to accommodate E_1, a contradiction.
- Can there be an equilibrium where both types of the incumbent accommodate E_1 for sure? In such an equilibrium, by part (iii) E_2 would enter upon seeing the incumbent accommodate E_1 (because $x_2 = x_1 < \frac{1}{2}$), and since by part (iii) the crazy incumbent would prey E_2, E_3 would then stay out. It follows that in such an equilibrium the crazy incumbent's equilibrium payoff must be

$$0 + \frac{1}{2} + \frac{3}{4},$$

following E_1's entry.

What would happen if the crazy incumbent deviates and preys E_1? If this induces E_2 to stay out with a positive probability q, then following E_2's staying out we have $x_3 = x_2 \geq \frac{1}{2}$, and E_3 may then stay out with probability r, implying that the crazy incumbent's deviation payoff would be

$$\frac{1}{2} + qr[\frac{3}{4} + \frac{3}{4}] + q(1-r)[\frac{3}{4} + \frac{1}{2}] + (1-q)[\frac{1}{2} + \frac{3}{4}],$$

which is higher than the crazy incumbent's equilibrium payoff, a contradiction. Thus following the crazy incumbent's deviation, E_2 must get in for sure, implying that the crazy incumbent's deviation payoff would be

$$\frac{1}{2} + \frac{1}{2} + \frac{3}{4},$$

which is again higher than the crazy incumbent's equilibrium payoff, a contradiction. Thus no such equilibrium exists.

- Can there be an equilibrium where only the sane incumbent accommodates E_1 for sure? In such an equilibrium, by the discussion in the preceding paragraph the crazy incumbent must prey E_1 with a positive probability. Thus seeing E_1 being preyed, E_2's posterior belief is such that $x_2 = 1$, and hence both E_2 and E_3 would stay out, by part (iii) above. This implies that the sane incumbent would strictly prefer to deviate and to prey E_1, a contradiction.
- Thus we are left with a single possibility, where the sane incumbent randomizes over preying and accommodating E_1. Let α and β be respectively the probabilities that E_2 may stay out and that E_3 may stay out upon seeing E_1 being preyed, and let α' and β' be respectively the probabilities that E_2 may stay out and that E_3 may stay out upon seeing E_1 being not preyed. Since the sane

incumbent must feel indifferent about preying or accommodating E_1, we must have

$$-1 + \frac{3}{4}(\alpha + \beta) = 0 + \frac{3}{4}(\alpha' + \beta'),$$

$$\Rightarrow \alpha + \beta - \alpha' - \beta' = \frac{4}{3},$$

so that the crazy incumbent must strictly prefer preying to accomodating E_1:

$$\{\frac{1}{2} + \frac{3}{4}(\alpha + \beta) + \frac{1}{2}[(1 - \alpha) + (1 - \beta)]\}$$
$$-\{0 + \frac{3}{4}(\alpha' + \beta') + \frac{1}{2}[(1 - \alpha') + (1 - \beta')]\}$$
$$= \frac{1}{2} + \frac{1}{4}[\alpha + \beta - \alpha' - \beta'] = \frac{5}{6} > 0.$$

Thus an equilibrium where the sane incumbent randomizes over preying and accommodating E_1 must be a semi-separating equilibrium, where the sane incumbent may prey E_1 with probability y and the crazy incumbent must prey E_1 with probability one.

It follows that E_2's posterior belief upon seeing that E_1 being accommodated must be such that $x_2 = 0$, and by part (iii), we must have $\alpha' = \beta' = 0$; that is, E_2 and E_3 would both enter upon seeing the incumbent accommodated E_1. We claim that this implies $1 > \alpha > 0$. Indeed, if instead $\alpha = 1$, then $x_3 = x_2 \geq \frac{1}{2}$ after E_1 is preyed, so that $\beta = 1$ also, which is a contradiction to the fact that

$$\alpha + \beta - \alpha' - \beta' = \frac{4}{3}.$$

On the other hand, if $\alpha = 0$ so that E_2 enters for sure after E_1 is preyed, then we would have $\beta = \frac{4}{3}$, which is an obvious contradiction.

Thus we must have $1 > \alpha > 0$; that is, E_2 must randomize over entering and staying out after E_1 is preyed. By part (iii), this requires that E_2 hold the posterior belief $x_2 = \frac{1}{2}$ upon seeing E_1 being preyed.

The probability y that the sane-type of incumbent may prey E_1 must be such that

$$x_2 = \frac{x_1 \cdot 1}{x_1 \cdot 1 + (1 - x_1) \cdot y} \Rightarrow y = \frac{x_1}{1 - x_1} \in (0, 1).$$

Let E_2 enter with probability a upon seeing E_1 being preyed. Then with probability $(1-a)$, E_2 may stay out following E_1 being preyed, and in that event, we have $x_3 = x_2 = \frac{1}{2}$, and we again assume that E_3 may enter with probability a. (Thus we are assuming that $\alpha = 1 - a$ and $\beta = (1 - a)^2$.) For the sane incumbent to feel indifferent about preying or accommodating E_1, we must have

$$-1 + (1 - a)a(\frac{3}{4} + 0) + (1 - a)^2(\frac{3}{4} + \frac{3}{4}) = 0 + 0 + 0,$$

$$\Rightarrow a = \frac{9 - \sqrt{57}}{6} \sim 0.2417.$$

Now, should E_1 enter after all? Note that E_1 would get 0 by staying out, and by entering it may get -1 with probability $2x_1$ or 1 with probability $1 - 2x_1$. Thus E_1 would get in if and only if $x_1 < \frac{1}{4}$.

So, what happens to i^*? First, if $x_1 \in [\frac{1}{4}, \frac{1}{2})$, then E_1 stays out and E_2 gets in for sure, and E_3 gets in if and only if E_2 was not preyed. In this case, $i^* = 2$. Second, if $x_1 < \frac{1}{4}$, then E_1 gets in with probability 1. If E_1 is not preyed, then $i^* = 1$ and E_2 and E_3 get in with probability 1. If E_1 is preyed, then E_2 gets in with probability a and in this case $i^* = 2$. It occurs with probability $(1-a)$ that E_2 stays out following E_1's being preyed, and following E_2's staying out, E_3 gets in with probability a, implying that $i^* = 3$ in this case. Finally, it can happen that E_1 is the only one who's ever entered on the equilibrium path (then E_1 must be preyed), and in this case $i^* = +\infty$. In sum, when $x_1 \in [0, \frac{1}{4})$, $i^* = 1, 2, 3$ or $+\infty$.

35. Let us modify the game of chain-store paradox by assuming 4 entrants instead of 3. It can be verified that no entry will ever occur if $x_1 > \frac{1}{2}$. Suppose that $x_1 \leq \frac{1}{2}$, and consider the subgame where E_1 has entered. It can be verified that there is no PBE for this subgame where the sane incumbent plays a pure strategy when facing E_1.[19]

    Thus we look for a PBE for this subgame where the sane incumbent preys with probability $y \in (0, 1)$ when facing E_1. This implies that the sane incumbent's equilibrium payoff is zero. Now, consider E_2's belief upon seeing E_1 being preyed by the incumbent. E_2 must believe that the probability that the incumbent is crazy is

$$\frac{x_1 \times 1}{x_1 \times 1 + (1 - x_1) \times y} > x_1.$$

    There are several possibilities.

    - Upon seeing E_1 being preyed, E_2 stays out. This requires that

$$\frac{x_1 \times 1}{x_1 \times 1 + (1 - x_1) \times y} \geq \frac{1}{4}.$$

[19]Suppose that the sane incumbent acquiesces with probability one. Then the sane incumbent's type is revealed immediately, and so the sane incumbent gets zero in equilibrium. But by deviating and preying, the sane incumbent gets -1 immediately, and will be regarded as the crazy type for sure, inducing E_2, E_3, E_4 to stay out, so that the payoff from deviation is $-1 + 3 \times \frac{3}{4} > 0$, a contradiction. Next, suppose that the sane incumbent preys with probability one. We have then a pooling equilibrium after E_1 enters. It follows that $x_2 = x_1 \leq \frac{1}{2}$. Two possibilities exist. First, if $x_1 > \frac{1}{4}$, then E_2 will stay out, according to our analysis in the pervious section. Then, $x_3 = x_2 = x_1 \leq \frac{1}{2}$, and hence E_3 will enter. It follows that the sane incumbent's payoff by preying with probability one after E_1 enters is $-1 + \frac{3}{4} + 0 + 0 < 0$, and hence the sane incumbent would be better off deviating and acquiescing E_1. Thus we have a contradiction. Second, if $x_1 \leq \frac{1}{4}$, then E_2 will get in, and our analysis for the 3-entrant case in the previous section says that the sane incumbent's payoff will be zero following E_2's entry. But then the sane incumbent by preying after E_1 gets in gets only $-1 + 0 < 0$, which is again a contradiction. To sum up, when $x_1 \leq \frac{1}{2}$, it is necessary that in equilibrium the sane incumbent randomizes between preying and acquiescing after E_1 enters.

It follows that

$$x_3 = \frac{x_1 \times 1}{x_1 \times 1 + (1 - x_1) \times y}.$$

In this case, $x_3 = \frac{1}{2}$. To see this, note that if If $x_3 < \frac{1}{2}$ so that E_3 and E_4 will enter, then by preying the sane incumbent's payoff would be negative even if E_2 stays out for sure, which is inconsistent with the assumption that the sane incumbent feels indifferent about preying when facing E_1. If, on the other hand, $x_3 > \frac{1}{2}$ so that E_3 and E_4 will stay out, then by preying the sane incumbent's payoff is strictly positive, which is again inconsistent with the assumption that the sane incumbent feels indifferent about preying when facing E_1. Hence we conclude that $y = \frac{x_1}{1-x_1} \in (0,1)$ in this case. Let the probability that E_3 enters be a, and assume that in case E_3 stays out, E_4 also enters with probability a. Then we need

$$-1 + \frac{3}{4} + a(1-a) \times \frac{3}{4} + (1-a)^2(\frac{3}{4} + \frac{3}{4}) = 0.$$

We obtain

$$a = \frac{9 - \sqrt{21}}{6} \in (0,1).$$

Let us verify that

$$\frac{x_1 \times 1}{x_1 \times 1 + (1 - x_1) \times y} \geq \frac{1}{4}.$$

This inequality holds if and only if

$$\Leftarrow 0 < y \leq \frac{3x_1}{1 - x_1}.$$

Since $y = \frac{x_1}{1-x_1}$, there is no contradiction arising here.

- Upon seeing E_1 being preyed, E_2 gets in. In this case, the sane incumbent gets zero payoff following E_2's entry by playing a mixed strategy against E_2. Apparently, this cannot be consistent with the assumption that the sane incumbent feels indifferent about preying when facing E_1.

- Upon seeing E_1 being preyed, E_2 randomizes between getting in and staying out. This implies that

$$\frac{x_1 \times 1}{x_1 \times 1 + (1 - x_1) \times y} = \frac{1}{4}.$$

Moreover, let λ be the probability that E_2 stays out. Note that following E_2's staying out, $x_3 = \frac{1}{4}$, and so E_3 will enter for sure. To be consistent with the assumption that the sane incumbent feels indifferent about preying when facing E_1, however, we need

$$-1 + \lambda(\frac{3}{4} + 0 + 0) + (1 - \lambda) \times 0 = 0,$$

which is impossible.

To sum up, in the subgame where E_1 gets in, the sane incumbent must prey with probability $\frac{x_1}{1-x_1}$, and all of E_2, E_3 and E_4 enter if E_1 was not preyed; and E_2 stays out upon seeing E_1 being preyed, and after that, E_3 enters with probability $\frac{9-\sqrt{21}}{6}$, and in the event that E_3 stays out, then E_4 enters with probability $\frac{9-\sqrt{21}}{6}$ also.

Back to the first stage of the game, we see that E_1 enters if and only if $x_1 < \frac{1}{4}$, and following that the players act as the preceding summary describes. In case $x_1 \geq \frac{1}{4}$, then E_1 stays out, and so does E_2, but following that, E_3 and E_4 both enter.

36. (**Part 2**.)

37. Recall the definition of signaling games: a game is called a signaling game if it is featured by (i) two players (called 1 and 2); (ii) only player 1 has types $\theta \in \Theta$; (iii) player 1 (also referred to as the *informed player*) given her type sends a signal $a_1 \in A_1$; (iv) player 2 (also referred to as the *uninformed player*) chooses action $a_2 \in A_2$ upon seeing a_1; and (v) the game ends after player 2 chooses a_2.

Let Θ be a finite set; i.e., $\Theta = \{\theta_1, \theta_2, \cdots, \theta_n\}$. Let $\mathcal{A}_i$ be the set of probability distributions over A_i. Let Σ_1 be the set of player 1's mixed strategies $\sigma_1 : \Theta \to \mathcal{A}_1$, where given a particular σ_1, let $\sigma_1(a_1|\theta)$ be the probability that player 1 chooses $a_1 \in A_1$ given her type θ. Let Σ_2 be the set of player 2's mixed strategies $\sigma_2 : A_1 \to \mathcal{A}_2$, where given a particular σ_2, $\sigma_2(a_2|a_1)$ is the prob. that player 2 chooses a_2 upon seeing the signal a_1 sent by player 1. Let $p(\theta)$ be player 2's prior beliefs about player 1's type θ. Let $\mathcal{P}$ be the set of prob. distributions over Θ. Let $\mathcal{M}$ be the set of mappings $\mu : A_1 \to \mathcal{P}$.

38. Given the signaling game defined in the preceding section, given $\mu \in \mathcal{M}$ and $a_1 \in A_1$, let

$$BR(\mu, a_1) = \arg\max_{a_2} \sum_{\theta\in\Theta} \mu(\theta|a_1)u_2(a_1, a_2, \theta),$$

$$BR(T, a_1) = \bigcup_{\mu\in\mathcal{M}:\mu(T|a_1)=1} BR(\mu, a_1),$$

$$MBR(T, a_1) = \bigcup_{\mu\in\mathcal{M}:\mu(T|a_1)=1} MBR(\mu, a_1),$$

where T is a subset of Θ, $\mu(T|a_1) = 1$ means that $\mu(\theta'|a_1) = 0$ for all $\theta' \in \Theta \setminus T$, and BR and MBR stand for respectively the set of pure-strategy best responses and the set of mixed-strategy best responses.

To formally define the Cho-Kreps intuitive criterion, given any proposed PBE, let $u_1^*(\theta)$ be the equilibrium payoff of the type-θ player 1, and define for all $a_1 \in A_1$,

$$J(a_1) = \{\theta : \ u_1^*(\theta) > \max_{a_2\in BR(\Theta,a_1)} u_1(a_1, a_2, \theta)\}.$$

The Cho-Kreps criterion states that player 2, upon seeing a_1, should hold posterior beliefs $\mu \in \mathcal{M}$ such that $\mu(\Theta \setminus J(a_1)|a_1) = 1$, and select a_2 accordingly from $BR(\Theta \setminus J(a_1), a_1)$. Now, if there exist $a_1 \in A_1$ and θ' such that

$$u_1^*(\theta') < \min_{a_2\in BR(\Theta\setminus J(a_1),a_1)} u_1(a_1, a_2, \theta'),$$

then we say the PBE fails the Cho-Kreps criterion. (We have assumed that the above maximum and minimum both exist.)

There is an enhanced version of the Cho-Kreps criterion, referred to as the *iterated intuitive criterion*, which involves defining for all $a_1 \in A_1$, $\Theta_1(a_1) = \Theta$ and for all $n \in \mathbf{Z}_+$,

$$J_n(a_1) = \{\theta : \ u_1^*(\theta) > \max_{a_2\in BR(\Theta_n(a_1),a_1)} u_1(a_1, a_2, \theta)\},$$

and

$$\Theta_{n+1}(a_1) = \Theta_n(a_1) \setminus J_n(a_1).$$

The PBE under consideration fails the *iterated intuitive criterion* if for some $a_1 \in A_1$, for some n, there exists $\theta' \in \Theta_{n+1}(a_1)$ such that

$$(\Delta) \quad u_1^*(\theta') < \min_{a_2 \in BR(\Theta_{n+1}(a_1), a_1)} u_1(a_1, a_2, \theta').$$

To see that the *iterated intuitive criterion* is a stronger criterion than the *intuitive criterion*, note that given $a_1 \in A_1$, $\{\Theta_n(a_1); n \in \mathbf{Z}_+\}$ is a decreasing sequence of subsets of Θ, so that by definition $\{$ BR$(\Theta_{n+1}(a_1), a_1)$; $n \in \mathbf{Z}_+\}$ is also a decreasing sequence of subsets of A_2. This implies that given a_1, for all $\theta' \in \Theta$ and for all $n \in \mathbf{Z}_+$,

$$\min_{a_2 \in BR(\Theta_n(a_1), a_1)} u_1(a_1, a_2, \theta') \leq \min_{a_2 \in BR(\Theta_{n+1}(a_1), a_1)} u_1(a_1, a_2, \theta').$$

Hence it is possible that a PBE is intuitive but not iteratedly intuitive.[20] [21]

[20]Consider the following signaling game. Player 1 has three equally probable types t_1, t_2, t_3, and 2 feasible signals m_1 and m_2. Player 2, upon seeing player 1's signal, can respond in three ways (r_1, r_2, r_3). If player 1 sends m_1, then the game ends with players 1 and 2 getting respectively 0 and 10, irrespective of player 1's type. The following table summarizes the two players' payoffs if player 1 sends m_2.

type/response	r_1	r_2	r_3
t_1	(-1,1)	(-2,1)	(-3,0)
t_2	(1,0)	(-1,2)	(-2,0)
t_3	(-1,0)	(-2,0)	(1,1)

Show that this game has a pooling equilibrium where all three types of player 1 send m_1 and upon seeing m_2, player 2 plays r_j with probability x_j, where $x_1 \leq x_2 + 2x_3$ and $x_3 \leq x_1 + 2x_2$. This equilibrium is intuitive, but fails to be iteratedly intuitive. Indeed, BR$(\{t_1, t_2, t_3\}, m_2)$ contains each and every mixed strategy that player 2 can use, but upon seeing m_2, intuitive criterion suggests that m_2 was not sent by t_1. Then, one can show that BR$(\{t_2, t_3\}, m_2)$ contains each and every mixed strategy that player 2 can use satisfying $x_1 = 0$, but if we restrict player 2's best response this way, then t_2 cannot have sent m_2. It follows that BR$(\{t_3\}, m_2)$ contains nothing but r_3, which then induce t_3 to deviate for sure. Thus the pooling equilibrium is intuitive, but fails to be iteratedly intuitive.

[21]A related concept is the *equilibrium domination criterion*, which involves replacing

39. Now we introduce the Grossman-Perry equilibrium. Consider the following signaling game. Player 1 has three equally probable types t_1, t_2, t_3, and 2 feasible signals m_1 and m_2. Player 2, upon seeing player 1's signal, can respond in three ways (r_1, r_2, r_3). If player 1 sends m_1, then the game ends with both players getting 2, irrespective of player 1's type. The following table summarizes the two players' payoffs if player 1 sends m_2.

type/response	r_1	r_2	r_3
t_1	(3,3)	(0,0)	(0,0)
t_2	(0,0)	(0,3)	(3,0)
t_3	(0,0)	(3,0)	(0,3)

(i) Show that this game has a pooling PBE where player 1 always sends m_1, and upon seeing m_2, player 2 believes that all three types of player 1 are equally likely to have made this deviation, and hence he randomizes over the three responses with equal probabilities. Verify that this PBE passes the Cho-Kreps criterion.
(ii) Show that this game has another PBE where player 1 sends m_2 if and only if his type is t_1, and upon seeing m_2, player 2 responds by choosing r_1. Show that this PBE also passes the Cho-Kreps criterion.
(iii) To define Grossman-Perry equilibrium, given any PBE, define T_i as the set of player 1's types that player 2 considers likely to have sent m_i. (Note that this definition does not distinguish equilibrium signals from off-the-equilibrium signals.)

Definition 3. A PBE is a *Grossman-Perry equilibrium* if for all i, $t \in T_i$ implies that sending m_i is *one* best response for the type-t player 1.

Show that the PBE in part (ii) is a Grossman-Perry equilibrium, but the PBE in part (i) is not.

(Δ) above by: for some $a_1 \in A_1$, for some n, for all $a_2 \in A_2$ there exists $\theta'(a_1)$ such that $u_1^*(\theta'(a_1)) < u_1(a_1, a_2, \theta'(a_1))$. Apparently, the iterated intuitive criterion is stronger than the equilibrium domination criterion, in that if a PBE fails the latter then it fails the former.

Solution. Consider part (i). It is easy to see that the asserted pooling PBE does exist. To see that it is a Cho-Kreps equilibrium, note that $J(m_2) = \emptyset$: type t_1 would prefer m_2 to m_1 if r_1 would be taken after player 2 sees m_2, which is optimal if player 2 holds the belief that $\mu(t_1|m_2) = 1$; type t_2 would prefer m_2 to m_1 if r_3 would be taken after player 2 sees m_2, which is optimal if player 2 holds the belief that $\mu(t_3|m_2) = 1$; and type t_3 would prefer m_2 to m_1 if r_2 would be taken after player 2 sees m_2, which is optimal if player 2 holds the belief that $\mu(t_2|m_2) = 1$.

Consider part (ii). It is easy to see that the asserted PBE does exist. To see that the PBE is a Cho-Kreps equilibrium, note that under this PBE both m_1 and m_2 are *equilibrium signals*.

Consider part (iii). It is easy to see that the PBE stated in part (ii) satisfies the Grossman-Perry criterion. To see that the PBE stated in part (i) fails the Grossman-Perry criterion, note that for player 2 to choose r_2 with a strictly positive probability after seeing m_2 (in this PBE r_2 is played with probability $\frac{1}{3}$ following m_2), it is necessary that

$$\mu(t_2|m_2) \geq \max(\mu(t_1|m_2), \mu(t_3|m_2)) \Rightarrow \mu(t_2|m_2) > 0,$$

so that $t_2 \in T_2$. However, by sending m_1 type t_2 would get 2, implying that sending m_2 type t_2 should get an expected payoff of 2 also, which in turn requires that player 2 chooses r_3 with probability $\frac{2}{3} > 0$, but for the latter to happen, it must be that

$$\mu(t_3|m_2) \geq \max(\mu(t_1|m_2), \mu(t_2|m_2)),$$

so that

$$0 < \mu(t_2|m_2) = \mu(t_3|m_2) \Rightarrow t_3 \in T_2,$$

which in turn implies that by sending m_2 type t_3 can get an expected payoff of 2. Note that the latter requires that player 2 should choose r_2 with probability $\frac{2}{3}$ after seeing m_2. This cannot be possible.

40. The following examples will show that the PBE solution concept may be too weak and render unreasonable solutions. These examples motivate Kreps and Wilson's sequential equilibrium concept.

- **Example 4**. Consider the following game with imperfect information.[22] Nature first chooses a type $t \in \{a, b\}$ for player 1 with equal probability, and then without seeing her own type, player 1 chooses between x and y, which is observed by player 2 before the latter chooses between l and r. The payoff vectors (u_1, u_2) for player 1 and player 2 are respectively $(2, 10), (0, 5), (5, 2), (0, 5), (5, 10)$, and $(2, 10)$ in respectively the events $(t, s_1, s_2) = (a, x, \cdot), (a, y, l), (a, y, r), (b, y, l), (b, y, r)$, and $(b, x, \cdot)$. One PBE for this game consists of player 1 playing x and player 2 playing l if y is observed, and player 2's posterior belief in the latter zero-probability event is that player 1 is of type a with probability 0.9. Note that the latter supporting belief is unreasonable, because player 1 cannot make her trembling to y contingent on her type; recall that she does not know her own type when choosing between x and y!
- **Example 5**. Consider the following game where E (the entrant) decides to stay out or to enter an industry, which is observed by I (the incumbent). Then E and I simultaneously choose to prey or to accomodate. The payoff vectors (u_E, u_I) are respectively $(0, 2), (-3, -1), (1, -2), (-2, -1)$, and $(3, 1)$ when the strategy profiles are respectively (Out,$\cdot$), (prey, prey), (prey, accomodate), (accomodate, prey), and (accomodate, accomodate). This game has a PBE where E chooses Out and E would accomodate with probability 1 after the zero-probability event that E chooses In takes place, and yet firm I thinks that E preys with probability 1. This PBE is not even subgame perfect!

41. The above problems pertaining to the PBE solution concept motivate Kreps and Wilson's *sequential equilibrium*, where Kreps and Wilson extend subgame perfection (P) to sequential rationality (S) (that extends the idea of perfection to non-singleton information sets) and refine the Bayesian updating rule (B) by requiring beliefs be consistent (C). Formally, an *assessment* is a pair (σ, μ), where σ is an (equilibrium) strategy profile and μ a set of beliefs that assigns beliefs to each and every information set in the game tree. The assessment is *sequentially*

[22] Harsanyi points out that an extensive Bayesian game can be modeled as an extensive game with imperfect information, and so the solution concept of PBE can be applied to games with imperfect information as well.

rational, if at any information set h where it is player $i(h)$'s turn to move, given the player's beliefs $\mu(h)$ at that information set, no other strategy $\sigma'_{i(h)}$ can make the player better off:

$$u_{i(h)}(\sigma|h, \mu(h)) \geq u_{i(h)}((\sigma'_{i(h)}, \sigma_{-i(h)})|h, \mu(h)).$$

This definition extends subgame perfection because (i) it does not require h to be a singleton set; and (ii) just like subgame perfection, it assumes that all other players will follow σ even if h ought to be reached with zero probability. Let Σ^0 be the set of totally mixed strategy profiles. Note that if $\sigma \in \Sigma^0$, then $\sigma_i(a_i|h) > 0$ for all h and for all $a_i \in A_{i(h)}(h)$ (where $A_{i(h)}(h)$ is the set of feasible actions for player $i(h)$ at information set h). In this case, for each node x on the game tree, $\mu(x) = \frac{P^\sigma(x)}{P^\sigma(h(x))}$, where $P^\sigma(\cdot)$ is the probability distribution over nodes induced by the strategy profile σ. That is, given that $\sigma \in \Sigma^0$, μ is completely determined by Bayes law. Let Ψ^0 be the set of assessments where $\sigma \in \Sigma^0$ and μ is derived from σ via Bayes law. We say that an assessment is *consistent* if it is the limit of a sequence of assessments $\{(\sigma^n, \mu^n)\}$ in Ψ^0. Now we can define a sequential equilibrium.

Definition 4. An assessment is a *sequential equilibrium* (SE) if it is *sequentially rational* (S) and *consistent* (C).

Thus like a PBE, a sequential equilibrium consists of a set of strategies plus a set of beliefs. Indeed, an SE is a PBE, but the converse is not necessarily true. These two equilibrium concepts coincide with each other if the game is a signaling game or if the informed player has only two possible types.

42. Note that we did not require in the definition of an SE that for all n $\{(\sigma^n, \mu^n)\}$ be an equilibrium in any sense, which distinguishes SE from the (trembling-hand) perfect equilibrium in extensive games. In fact, every perfect equilibrium in extensive games is an SE, and hence the existence of SE is implied by the existence of perfect equilibria. Although there generally may exist infinitely many SE's for a finite extensive game with incomplete information, as Kreps and Wilson show, the number of distinct equilibrium payoff profiles is generically finite.

43. Now, let us show that the PBE's obtained in Examples 4 and 5 are not SE's. In the former example, if for all $n \in \mathbf{Z}_+$, σ_1^n is a totally mixed strategy of player 1, then the posterior belief of player 2 upon seeing y should be that player 1 is of type a with probability $\mu^n(a) = \frac{1}{2}$, which converges to $\mu(a) = \frac{1}{2}$ for sure, and with this belief, player 2 is better off choosing r. However, expecting player 2 to choose r upon seeing y, player 1 is better off choosing y in the first place, upsetting the PBE. Thus the PBE is not a sequential equilibrium.

    Similary, in the latter example, suppose that for all $n \in \mathbf{Z}_+$, E chooses In with probability $\epsilon^n > 0$, where $\epsilon^n \downarrow 0$, and that given each n, $e^n > 0$ is the probability that E will prey given that he has entered the industry, with $e^n \downarrow 0$ also. Then firm I's belief upon seeing E's entry should be that

$$\mu^n(\text{prey}) = \frac{e^n \epsilon^n}{\epsilon^n} = e^n \downarrow 0 = \mu(\text{prey}),$$

    which contradicts the supporting belief specified by the PBE; namely, $\mu(\text{prey}) = 1$. That is, the assessment specified by the original PBE fails to be *consistent*, and hence that PBE is not a sequential equilibrium.[23] [24]

44. Now we introduce the divine equilibrium. Let

$$D(\theta, T, a_1) = \bigcup_{\{\mu:\mu(T|a_1)=1\}} \{\alpha_2 \in MBR(\mu, a_1) : \ u_1^*(\theta) < u_1(a_1, \alpha_2, \theta)\}$$

    and

$$D^0(\theta, T, a_1) = \bigcup_{\{\mu:\mu(T|a_1)=1\}} \{\alpha_2 \in MBR(\mu, a_1) : \ u_1^*(\theta) = u_1(a_1, \alpha_2, \theta)\}.$$

[23] The equilibrium concept SE also has some drawbacks itself. First, it is not robust against additions or deletions of irrelevant moves. Second, it may allow for equilibria involving some players playing weakly dominated strategies.

[24] Recall from Lecture 1, Part II that the perfect equilibrium defined for strategic games does not remove all subgame-imperfect NE's in extensive form. Selten's remedy (1975) is to define agent's normal form and require the original definition of perfect equilibrium to adapt to this new concept. This way, all subgame-imperfect NE's are removed. With this new definition, all perfect equilibria are sequential, and the two coincide generically.

Definition 5. Bank and Sobel's *divine equilibrium* is a sequential equilibrium that can be supported by the following beliefs: if for θ, θ',

$$D(\theta, \Theta, a_1) \bigcup D^0(\theta, \Theta, a_1) \subset (\neq) D(\theta', \Theta, a_1),$$

then

$$\frac{p(\theta')}{p(\theta)} \leq \frac{\mu(\theta'|a_1)}{\mu(\theta|a_1)},$$

where recall that $p(\cdot)$ defines the prior beliefs on Θ.[25]

45. To demonstrate the concept of divine equilibrium, consider the following game of lawsuit settlement.

Example 6. Player 1 (the defendant) can be either guilty or not guilty with equal probability, which is her private information, and she can offer to settle at either 3000 or 5000 dollars with player 2 (the plaintiff). Player 2 can either accept or reject the offer. If player 2 rejects the offer, then the two go to the court and the truth (the type of the defendant) will reveal with player 1 paying player 2 nothing if the verdict is "not guilty" and 5000 if the verdict is "guilty." In any case, going to the court will cost player 1 6000 regardless of the outcome of the verdict. The game has two pooling PBE's. In one, both types of player 1 offer to settle at 3000 and player 2 is willing to accept either 3000 or 5000. In the other, both types of player 1 offer to settle at 5000 and player 2

[25]The strengthened definition of divine equilibrium can be found in Fubenberg and Tirole (1991), where a divine equilibrium must have supporting beliefs that satisfy the following D1 criterion: Let $\Theta_1 = \Theta$ and G_1 the subset of Θ such that $\theta \in G_1$ if and only if there exists some other $\theta' \in \Theta$ such that

$$D(\theta, \Theta, a_1) \bigcup D^0(\theta, \Theta, a_1) \subset (\neq) D(\theta', \Theta, a_1).$$

Then, for $n \geq 2$, define $\Theta_n = \Theta \setminus \bigcup_{j=1}^{n-1} G_j$, where G_j is the subset of Θ such that $\theta \in G_j$ if and only if there exists some other $\theta' \in \Theta_j$ such that

$$D(\theta, \Theta_j, a_1) \bigcup D^0(\theta, \Theta_j, a_1) \subset (\neq) D(\theta', \Theta_j, a_1).$$

Then the supporting beliefs must be such that, following any off-the-equilibrium signal a_1, the posterior probability for any element contained in $\Theta \setminus \bigcup_{n \geq 2} G_{n-1}$ is zero. Note that this definition is indeed stronger than the intuitive criterion, because by the intuitive criterion, G_1 contains only those θ with empty $D(\theta, \Theta, a_1)$ and $D^0(\theta, \Theta, a_1)$.

stands ready to accept 5000 but to reject 3000. Both PBE's pass the Cho-Kreps criterion.[26]

The second PBE is however not divine. Note that off the equilibrium player 2 can get 3000 if she accepts the deviation offer 3000, and $5000\mu(\text{guilty}|3000)$ if she rejects that deviation offer. Thus her MBR contains accepting the deviation offer (if $\mu(\text{guilty}|3000) \leq \frac{3}{5}$), rejecting the deviation offer (if $\mu(\text{guilty}|3000) \geq \frac{3}{5}$), and a mixed strategy (if $\mu(\text{guilty}|3000) = \frac{3}{5}$). It can be shown that[27]

$$D(\text{not guilty}) = \{\text{prob.(accepting)} > \frac{1}{3}\},$$

$$D(\text{guilty}) = \{\text{prob.(accepting)} > \frac{3}{4}\},$$

$$D^0(\text{guilty}) = \{\text{prob.(accepting)} = \frac{3}{4}\}.$$

Thus in a divine equilibrium, following Bank and Sobel's argument,

[26] Apparently, in the first equilibrium both types of player 1 strictly prefer the equilibrium signal 3000 to the off-the-equilibrium signal 5000, implying that $J(5000) = \Theta$. The intuitive criterion does not add any restrictions on the set of posterior beliefs in this case. In the second equilibrium, $J(3000) = \emptyset$: if 3000 would be accepted with probability one then both the guilty and the non-guilty types would prefer offering 3000 to the equilibrium signal of offering 5000. Note that accepting 3000 can be optimal if player 2's posterior beliefs are such that $\mu(\text{not guilty}|3000) = 1$.

[27] Recall that in equilibrium both types of the defendant offer 5000 without going to the court. If offering 3000 implies an expected payoff less than (or equal to) 5000 for a given type of defendant, then he would strictly (weakly) prefer the deviation to his equilibrium signal. That is, to make a guilty type weakly prefer offering 3000, it is necessary and sufficient that

$$q \cdot 3000 + (1-q) \cdot (5000 + 6000) \leq 5000 \Rightarrow q \geq \frac{3}{4},$$

where q is the plaintiff's prob. of accepting 3000. Similarly, to make a non-guilty type weakly prefer offering 3000, it is necessary and sufficient that

$$q \cdot 3000 + (1-q) \cdot (0 + 6000) \leq 5000 \Rightarrow q \geq \frac{1}{3}.$$

In other words, knowing that he does not need to pay the plaintiff after going to the court, the non-guilty type of defendant is more willing to try the low offer 3000 and run the risk of being turned down and having to go to the court.

when the off-the-equilibrium signal 3000 appears, we must require

$$\mu(\text{not guilty}|3000) \geq \frac{1}{2},$$

where $\frac{1}{2}$ is the prior probability that the defendant is not guilty. However, with the Bank-Sobel posterior beliefs, player 2 had better accept the 3000 when it appears (if going to the court, player 2's expected payoff is $\mu(\text{guilty}|3000) \cdot 5000 < 3000$), and knowing this player 1 will not offer 5000 in equilibrium, showing that the PBE is not a divine equilibrium.