Lecture Note 5:
Signaling

- Signaling Games
  - The Intuitive Criterion
  - Forward Induction Equilibrium
  - D1, D2, Divinity, Universal Div.

The Intuitive Criterion (Cho & Kreps, 1987)

- Two players: a Sender (S) and a Receiver (R).
- The timing of the game is:
  - (1) nature draws a type for S, denoted $t \in T$, according to the probability distribution $p(t)$;
  - (2) S privately observes the type t and then sends the message $m \in M$ to R; and
  - (3) R observes m and then takes the action $a \in A$.
- T, M, and A are all finite.
- Payoffs are US(t,m,a) and UR(t,m,a).
- Everything but t, is common knowledge.

Identifying Unreasonable Sequential Equilibria

- Test:
- Consider the game: T={t,t'} and M={m,m'}.
- Suppose a pooling equilibrium: t, t’ send m with probability one.
- Then the message m' is off the equilibrium path, so R's beliefs after observing m' cannot be derived from Bayes' rule.
- Instead, these beliefs need only satisfy Kreps and Wilson's definition of consistency in order to be part of a sequential equilibrium.

Identifying Unreasonable Sequential Equilibria

- Test:
- By sequential rationality, the action R takes after observing m' must be optimal given R's beliefs. That is,

$$a(m') \in \operatorname*{argmax}_{a \in A} \sum_{t \in T} \mu(t|m')U^R(t, m', a).$$

Identifying Unreasonable Sequential Equilibria

- Suppose:
- (1) $\forall$ beliefs, the action a(m') makes type t worse off than t is in the equilibrium, and
- (2) if R infers from m' that S is type t', then R's optimal action will make t' better off than t' is in the equilibrium.

Identifying Unreasonable Sequential Equilibria

- Then, if S is type t', the following speech should be believed by R:

I am t'. To prove this, I am sending m' instead of the equilibrium m. Note that if I were t I would not want to do this, no matter what you might infer from m'. And, as t', I have an incentive to do this provided it convinces you that I am not t.

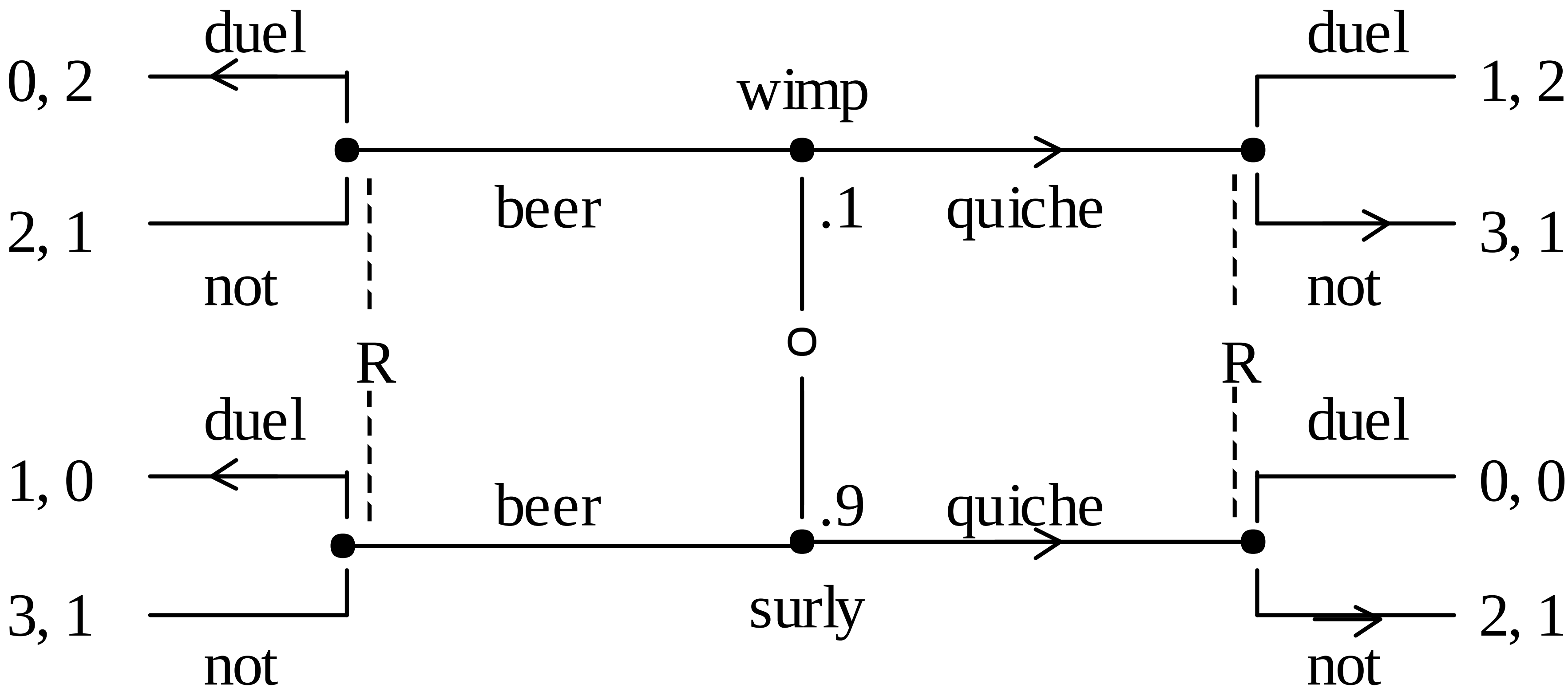
Identifying Unreasonable Sequential Equilibria

- Given (1) and (2), t' should deviate from the sequential equilibrium in which m is sent with probability one.
- On this ground, Cho & Kreps reject the equilibrium.

Beer and Quiche: The Entry-Deterrence Problem

- The incumbent firm can be either "surly" or a "wimp.”
- A surly firm prefers to have "beer" for breakfast, whereas a wimp prefers "quiche.”
- Having the preferred breakfast is worth 1 to the incumbent, but avoiding a duel with the entrant is worth 2.
- The entrant's payoffs are independent of the incumbent's breakfast: the entrant prefers to duel with the wimp but not to duel with the surly incumbent.
- The prior probability that the incumbent is surly is 0.9.

Beer and Quiche: The Entry-Deterrence Problem



Sequential Equilibria

- Two kinds of SE:
- Pooling: Both types drink beer, and the entrant duels if quiche is observed but declines to duel if beer is observed.

In such an equilibrium, the decision to duel following quiche is rationalized by any off-the-equilibrium-path belief that puts sufficiently high probability (at least 1/2) on the incumbent being wimpy. [Note: If Pr(wimp) > 1/2, then equilibrium involves mixed strategies for Wimp and R if Beer.]

Sequential Equilibria

- Pooling: both types have quiche, and the entrant duels if beer is observed but declines to duel if quiche is observed.

The beliefs that support the decision to duel are those that attach high probability to the wimp. But here such beliefs seem unnatural: the prior belief is .9 that the incumbent is surly, but when conditioned on the observation of beer—which is preferred if surly but not if wimpy—the posterior belief is at least .5 that the incumbent is wimpy.

Sequential Equilibria

How can we reject the second equilibrium?

- Using forward induction one can show that surly will find it optimal to deviate from the proposed equilibrium (both eat quiche):

  If the entrant concludes that the beer-drinker is surly, then declining to duel is the optimal decision. This yields a payoff of 3 for surly, which is better than the 2 earned in equilibrium. Then, the second supposition is met.

Formalizing the "Intuitive Criterion"

Notation:

- After hearing $m \in M$, R's beliefs are $\mu(t|m)$.
- Sequential rationality requires that R's subsequent action a(m) maximize the expectation of $U^R(t,m,a)$ with respect to these beliefs. Define the set of such best responses as

$$BR(\mu, m) \equiv \underset{a \in A}{\operatorname{argmax}} \sum_{t \in T} \mu(t|m) U^R(t, m, a).$$

- Then R's (behavior) strategy $\pi^R(a|m)$ is greater than zero only if $a \in BR(\mu,m)$.

Formalizing the "Intuitive Criterion"

Notation:

- For subsets I of T, let BR(I,m) denote the set of best responses for R to beliefs concentrated on I:

$$\mathrm{BR}(\mathrm{I},\mathrm{m}) \equiv \bigcup_{\{\mu:\mu(\mathrm{I})=1\}} \mathrm{BR}(\mu,\mathrm{m}).$$

- Given the equilibrium strategies $\pi = \{\pi^S(m|t), \pi^R(a|m)\}$, the equilibrium payoff to an S of type t is

$$\mathrm{U}^*(\mathrm{t}) \equiv \sum_{\mathrm{a}\in\mathrm{A}} \sum_{\mathrm{m}\in\mathrm{M}} \pi^{\mathrm{R}}(\mathrm{a}|\mathrm{m})\pi^{\mathrm{S}}(\mathrm{m}|\mathrm{t})\mathrm{U}^{\mathrm{S}}(\mathrm{t},\mathrm{m},\mathrm{a}).$$

Intuitive Criterion

- An equilibrium *fails* to satisfy the intuitive criterion if there exist

(a) an unsent message m' $\in$ M (i.e., π^S(m'|t)=0 for all t $\in$ T),

(b) a subset J of T, and

(c) a type t' $\in$ T~J

such that

(1) for all t $\in$ J, for all a $\in$ BR(T,m'), $U^*(t) > U^S(t,m',a)$, and

(2) for all a $\in$ BR(T~J,m'), $U^*(t') < U^S(t',m',a)$.

Spence's Signaling Model

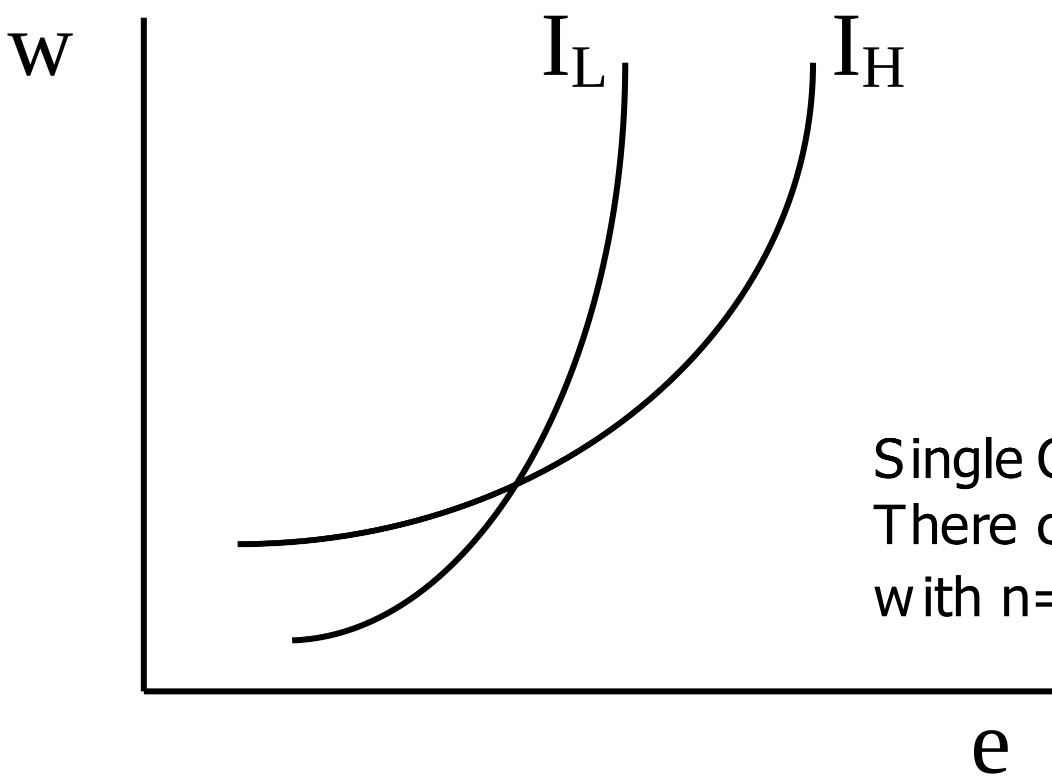
- A worker privately observes her productive ability and then chooses an amount of education e; $e \in [0,\infty)$.
- The market observes e and then offers a wage w; $w \in [0,\infty)$.
- Worker's preferences:

$$U^S(t,e,w) = w - c(t,e), \quad t \in \{H,L\}$$

- c(t,e) is the (psychic) cost for worker t of acquiring education e.

Spence's Signaling Model

- The low-ability worker has higher marginal cost of education than does the high-ability worker.



Single Crossing Property:
There cannot be pooling with n=2.

Spence's Signaling Model

- Let a worker of ability t and education e produce output y(t,e), where

$$y(H,e) > y(L,e) \text{ for all } e \text{ and}$$
$$\partial y(t,e)/\partial e > 0 \text{ for all } t,e.$$

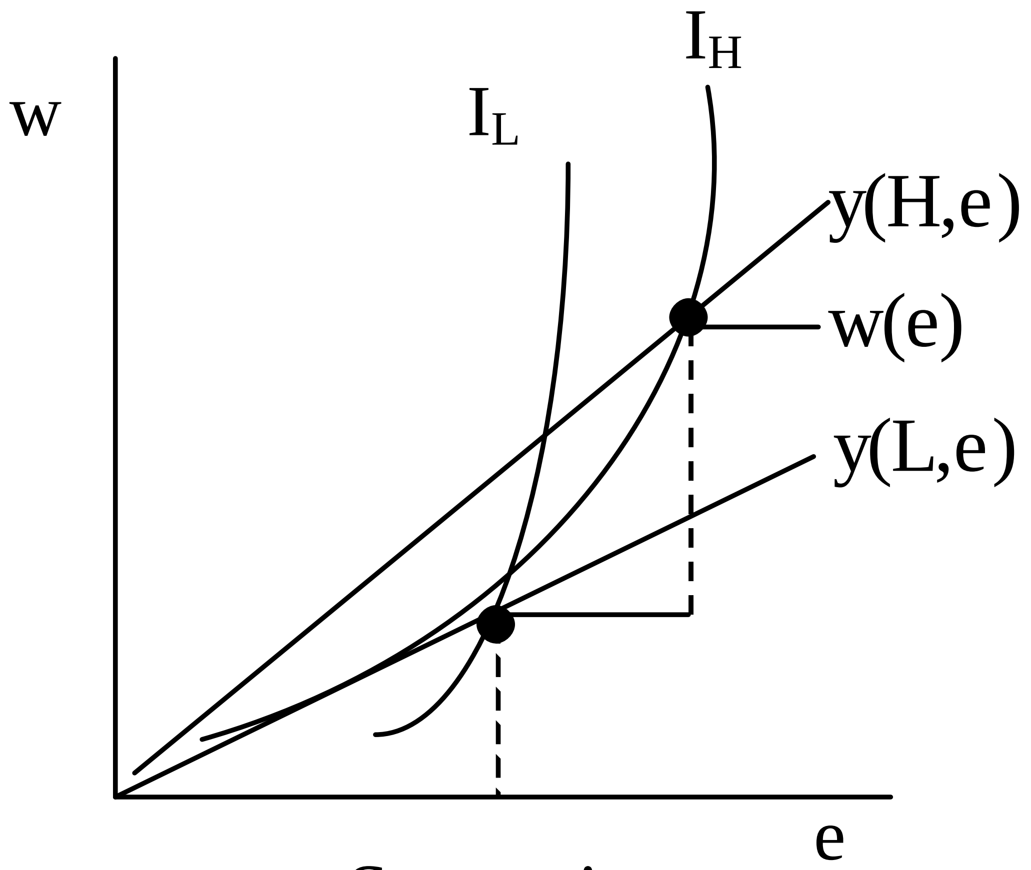
- Spence (1973) argues that competition among firms will drive profits to zero. In terms of a Bayesian Nash equilibrium, this means that given a conjecture e(t) about the worker's education choice, the market wage will satisfy w(e) = y(t,e).

Spence's Signaling Model

Three categories of equilibria:

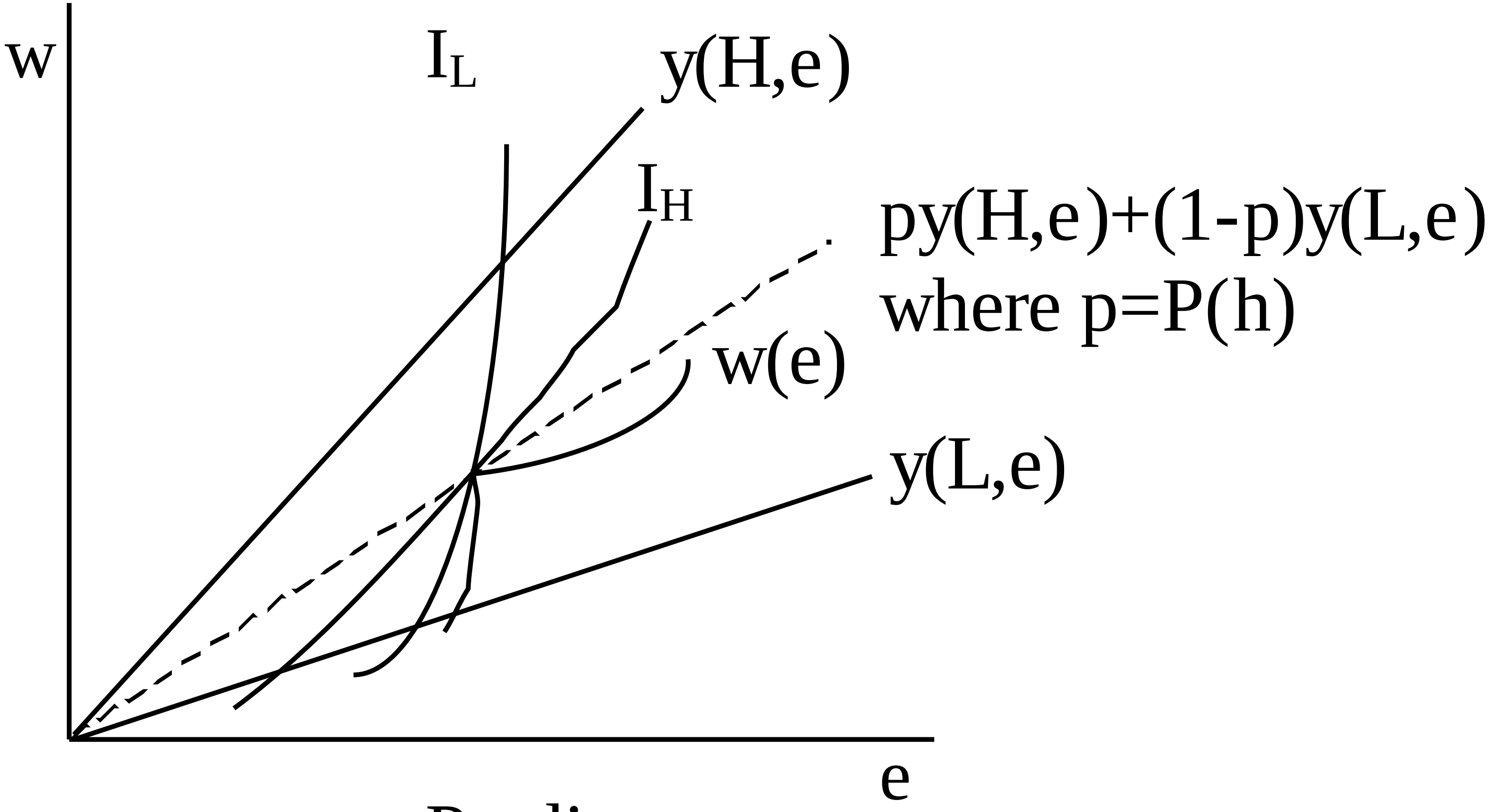
- Separating (in which the two types choose different levels of education).
- Pooling (in which they choose the same education).
- Hybrid (in which at least one type randomizes between pooling with the other type and distinguishing itself).

Separating Equilibrium



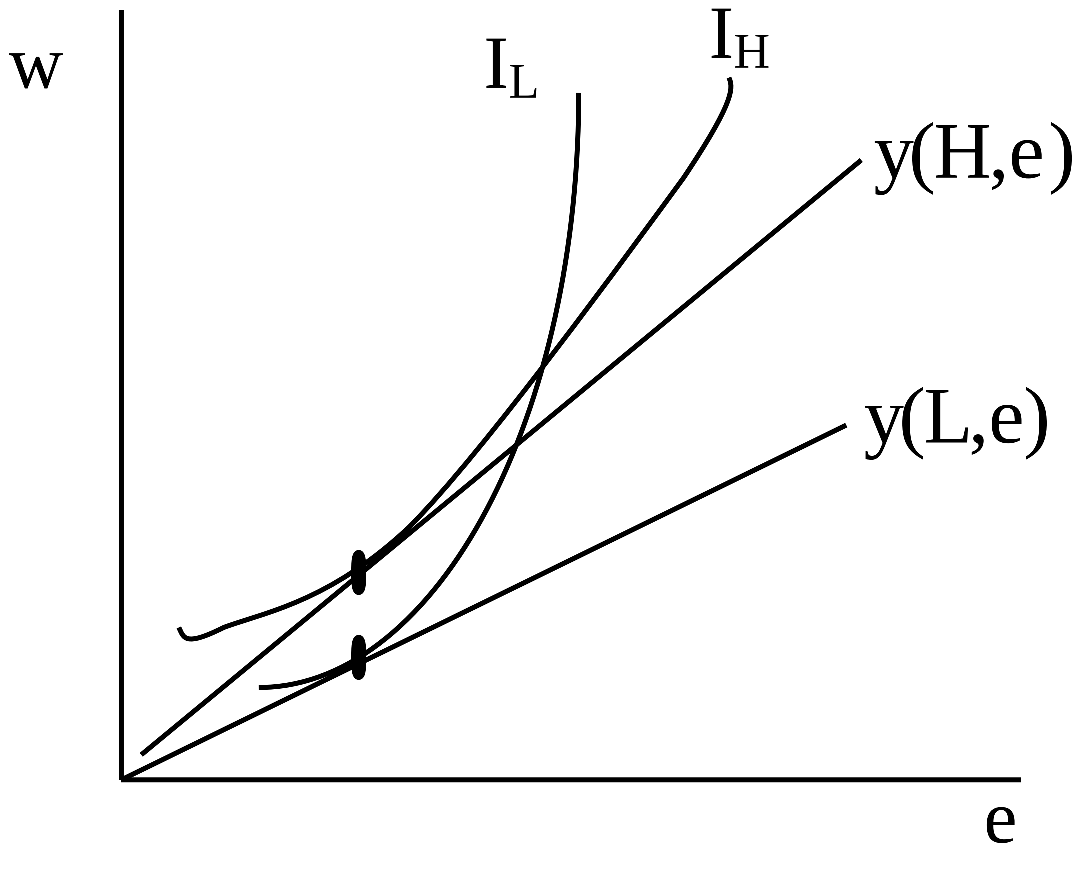
Separating

Pooling Equilibrium



Pooling

Envy Case



Steps for Selecting Equilibria

- First, some of the Nash equilibria can be rejected because they are not sequential equilibria.
- Second, some sequential equilibria can be rejected because they do not remain equilibria after weakly dominated strategies have been eliminated.
- Third, other sequential equilibria can be rejected because they do not conform to the intuitive criterion.

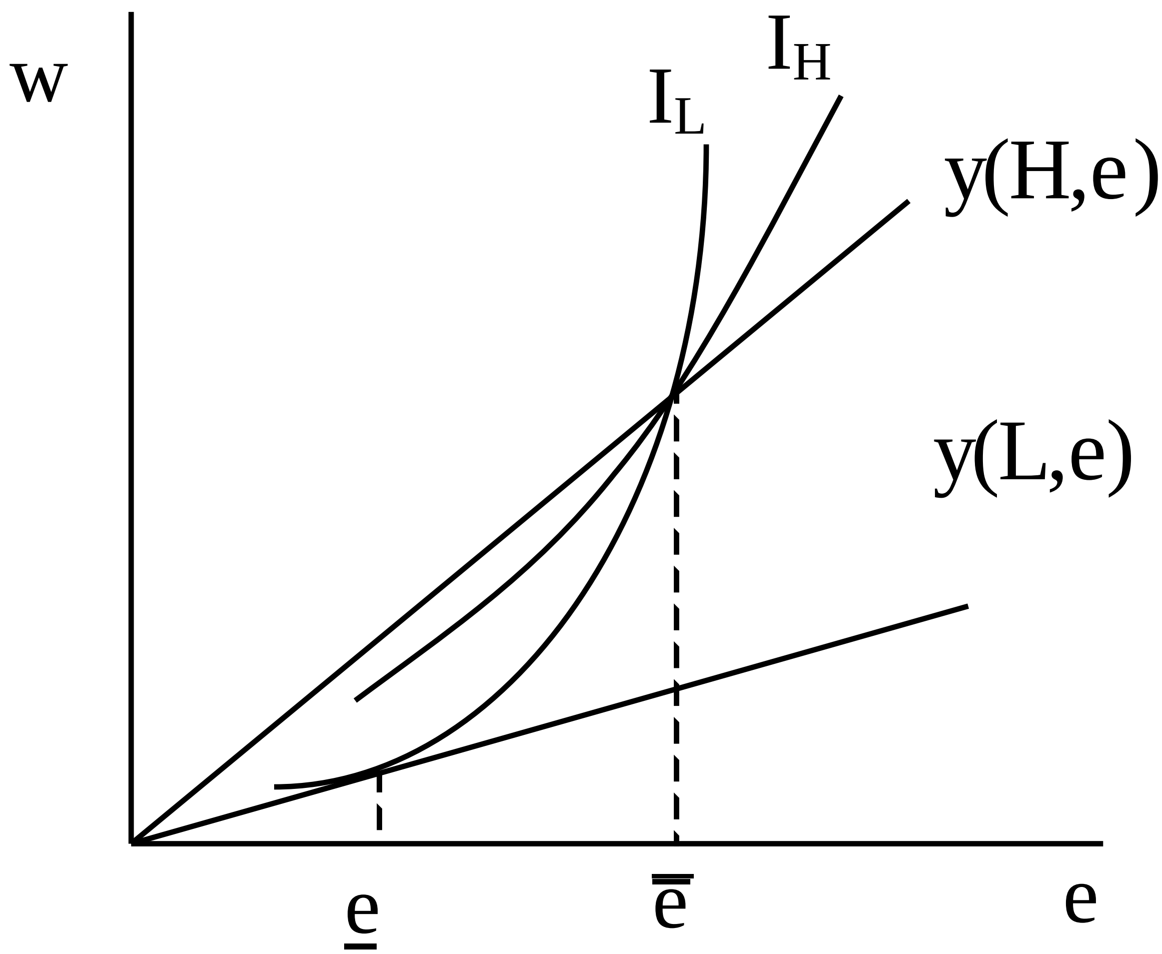
Step 1: Sequential Rationality

- In a sequential equilibrium the market must have beliefs μ(t|e) following any signal e.
- Wage is the expected productivity given these beliefs,

$$y(L,e) \leq w(e) \leq y(H,e), \text{ for each } e.$$

- Therefore, the wage schedules in the separating and pooling equilibria drawn above are Nash but not sequential.

Figure 1



Step 1: Sequential Rationality

- Suppose worker L separates with positive probability. Then
  - (a) it happens at the tangency of y(L,e) and I_L, hereafter $\underline{e}$.

    $w(\underline{e}) \geq y(L, \underline{e}) \Rightarrow w(\underline{e}) - c(L, \underline{e}) > y(L,e) - c(L,e)$

    for all $e \neq \underline{e}$.
  - (b) L accepts no utility less than $y(L, \underline{e}) - c(L, \underline{e})$ by individual rationality (an implication of the equilibrium); and
  - (c) any hybrid equilibrium must do its pooling on the indifference curve I_L through $(\underline{e}, y(L, \underline{e}))$.

Step 2: Elimination of Weakly Dominated Strategies

- L earns at least $y(L, \underline{e}) - c(L, \underline{e})$ for any equilibrium, whether or not L separates. Therefore, in any equilibrium

  (d) education levels above $\overline{e}$ (determined by the intersection of the productivity curve y(H,e) and the indifference curve I_L through $(\underline{e}, y(L, \underline{e}))$ are weakly dominated for L;

  (e) market beliefs $\mu(t|e)$ for $e > \overline{e}$ in the pruned game tree must be degenerate on H;

  (f) wages must be $w(e) = y(H,e)$ for $e > \overline{e}$; and

  (g) H accepts no utility less than $y(H,\overline{e}) - c(H,\overline{e})$.

Step 2: Elimination of Weakly Dominated Strategies

- Returning to the assumption that L separates with positive probability yields

  (h) the only possible hybrid is at ($\bar{e}$,y(H,$\bar{e}$)), but this wage earns negative profits unless the probability that L accepts is zero; and

  (i) there are no hybrid equilibria in which L separates with probability less than one.
- There is a unique equilibrium in which L separates with positive probability. In it, both types separate with probability one, L at ($\underline{e}$,y(L,$\underline{e}$)) and H at ($\bar{e}$,y(H,$\bar{e}$)).

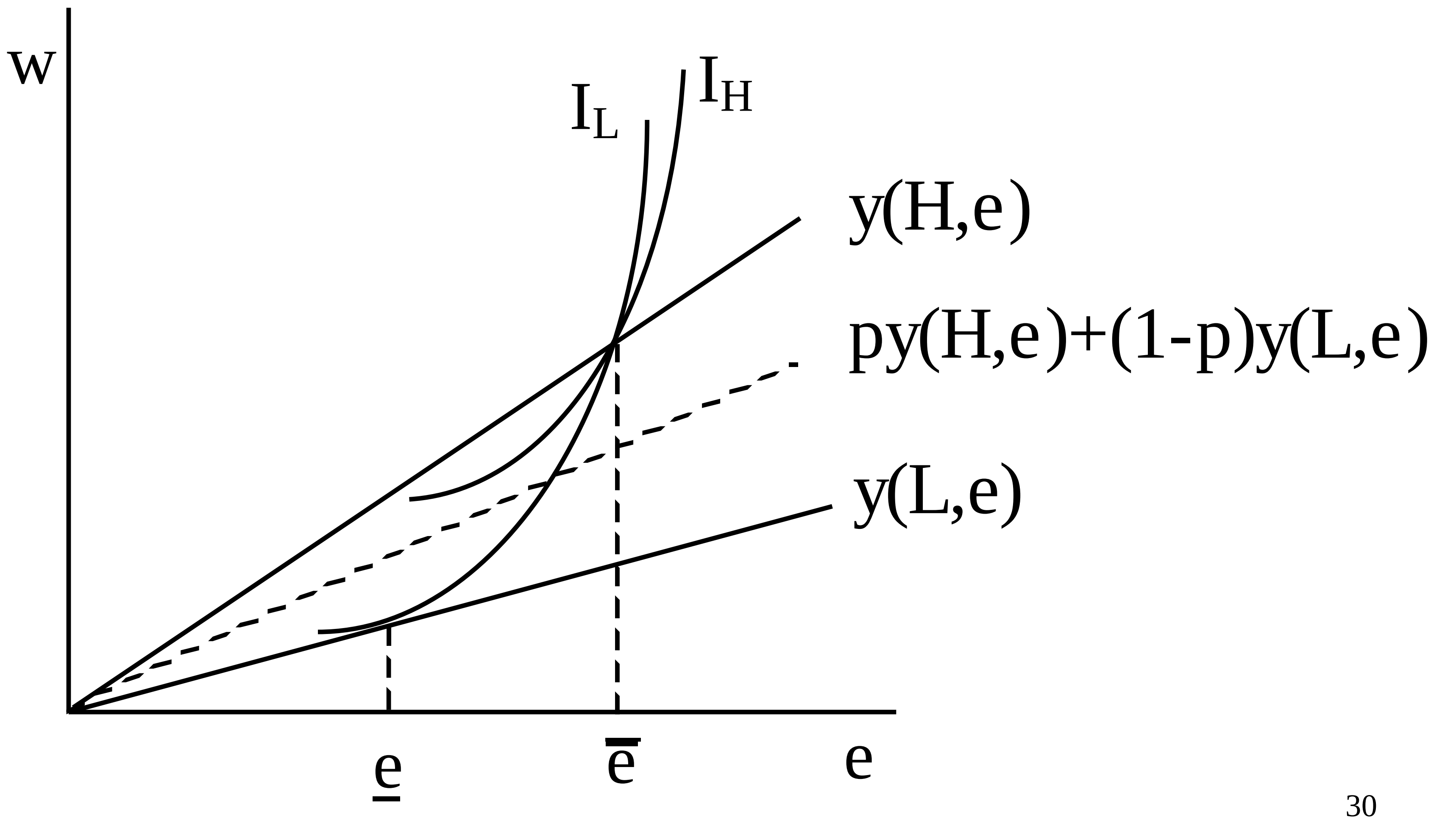
Step 3: Intuitive Criterion

- Any alternative equilibrium must have L separating with probability zero. There are two cases: pooling and hybrid equilibria. In such equilibria,

  (*) $w(e) \leq p(H)y(H,e) + p(L)y(L,e)$,

  with equality for pooling equilibria and inequality for hybrids.
- If this prevents H from achieving the utility $y(H,\bar{e}) - c(H,\bar{e})$, then by (g) these kinds of equilibria do not exist.

Figure 2

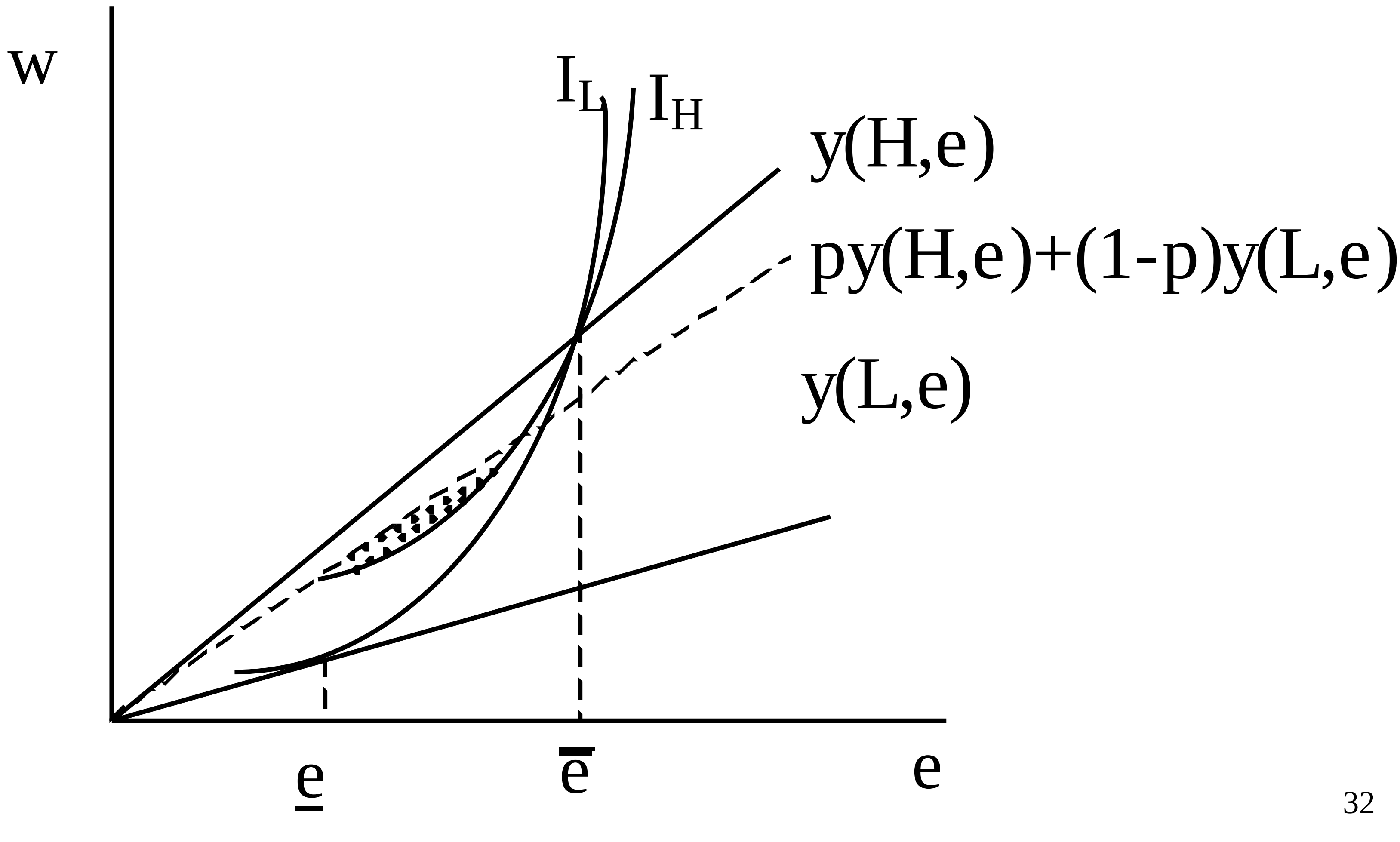


Step 3: Intuitive Criterion

If H can achieve the requisite utility then such equilibria exist but are rejected by the intuitive criterion.

- Pick a point satisfying (*) and H's utility constraint.
- Consider the indifference curves I_L and I_H through this point. By assumption, I_L is steeper, so the intersection of I_L and y(H,e) is to the left of the intersection of I_H and y(H,e).

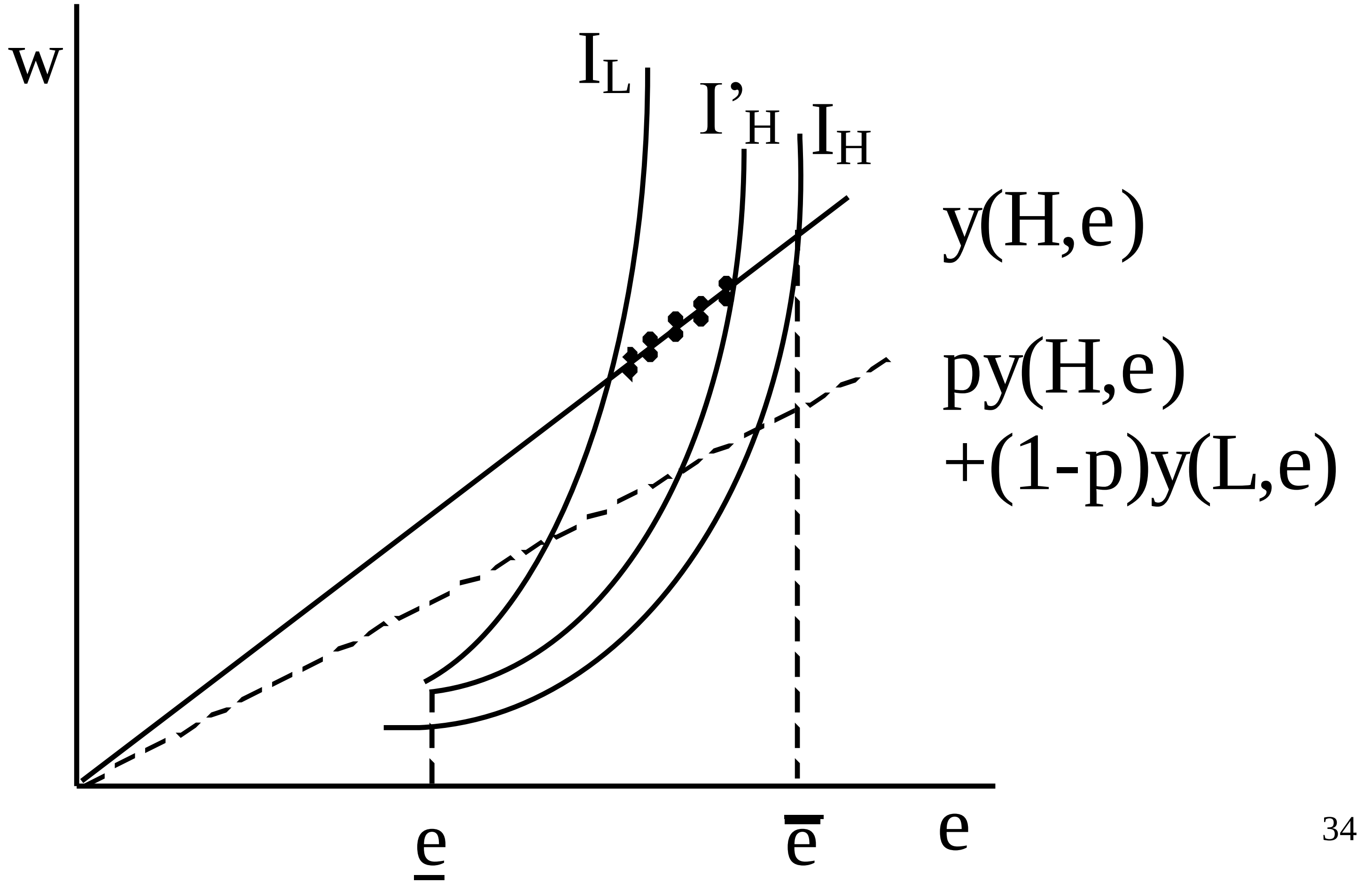
Figure 3



Step 3: Intuitive Criterion

- Any e between these two points of intersection is an unsent message that fulfills the requirements of the intuitive criterion: the market should infer that the worker is H because such signals are worse than the equilibrium payoff for L, but if it is sure to be H then the wage must be w(e)=y(H,e), which makes H better off than in the equilibrium.

Figure 4



Signaling Games and the Intuitive Criterion

- Recall the signaling game of Cho & Kreps (1987). The timing is:

  1. nature draws a type $t \in T$ for the Sender, S;

  2. S learns t and sends a message $m \in M$ to the Receiver, R; and

  3. R observes m and takes an action $a \in A$.

- The payoffs are $U^S(t,m,a)$ and $U^R(t,m,a)$. Everything is common knowledge except nature's choice of t for S.

Signaling Games and the Intuitive Criterion

- Intuitive criterion: reject any sequential equilibrium satisfying the following conditions:

there exists an unsent message m' and a subset of types J such that

(1) for all $t \in J$, for all $a \in BR(T,m')$, $U^*(t) > U^S(t,m',a)$, and

(2) there exists $t' \in T \sim J$ such that for all $a \in BR(T \sim J,m')$,
$U^*(t') < U^S(t',m',a)$,

where $U^*(t)$ is t's expected payoff in the equilibrium under consideration.

Steps for Rejecting Equilibria

- *Step 1*: Condition (1) suggests that R's belief μ(t|m') should put no probability on types t $\in$ J: reasonable μ's should be concentrated on T~J.

- *Step 2*: If there is a type t' satisfying (2), then surely this type would deviate from the proposed equilibrium, since t' is better off deviating no matter what reasonable belief R will hold.

Forward Induction Equilibrium (Cho, 1987)

- Let

  $J(m'|\pi) \equiv \{t \in T | U^*(t) > U^S(t,m',a)$ for all $a \in BR(T,m')\}$,

  where π is the sequential-equilibrium strategy in question.This is the largest set J that satisfies (1) above.
- Reasonable beliefs following the deviation m' are then those that assign zero probability to $t \in J(m'|\pi)$:

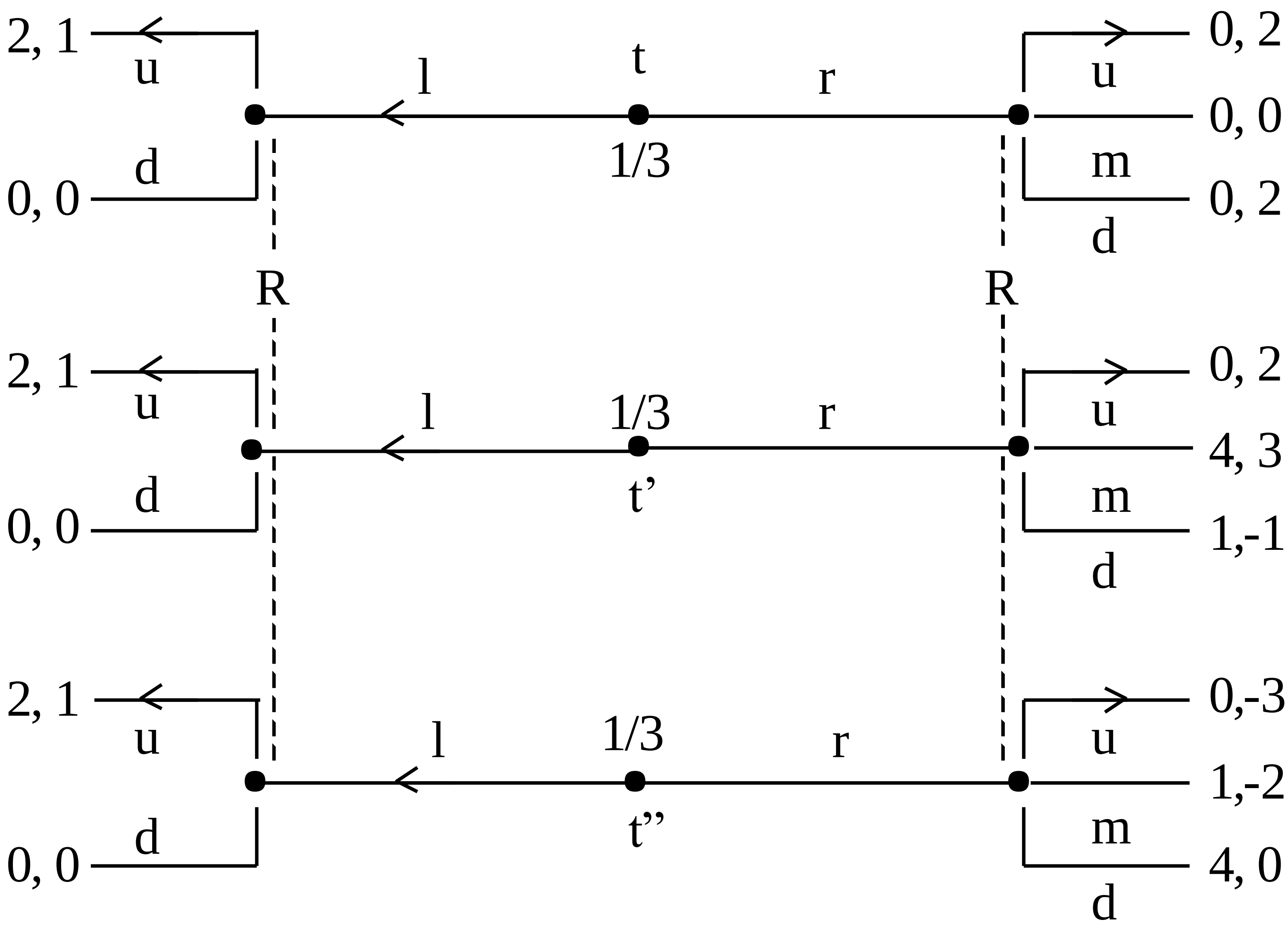
  $$\mu(J(m'|\pi) \mid m') = 0,$$

  provided $J(m'|\pi)$ is a proper subset of T.

Forward Induction Equilibrium (Cho, 1987)

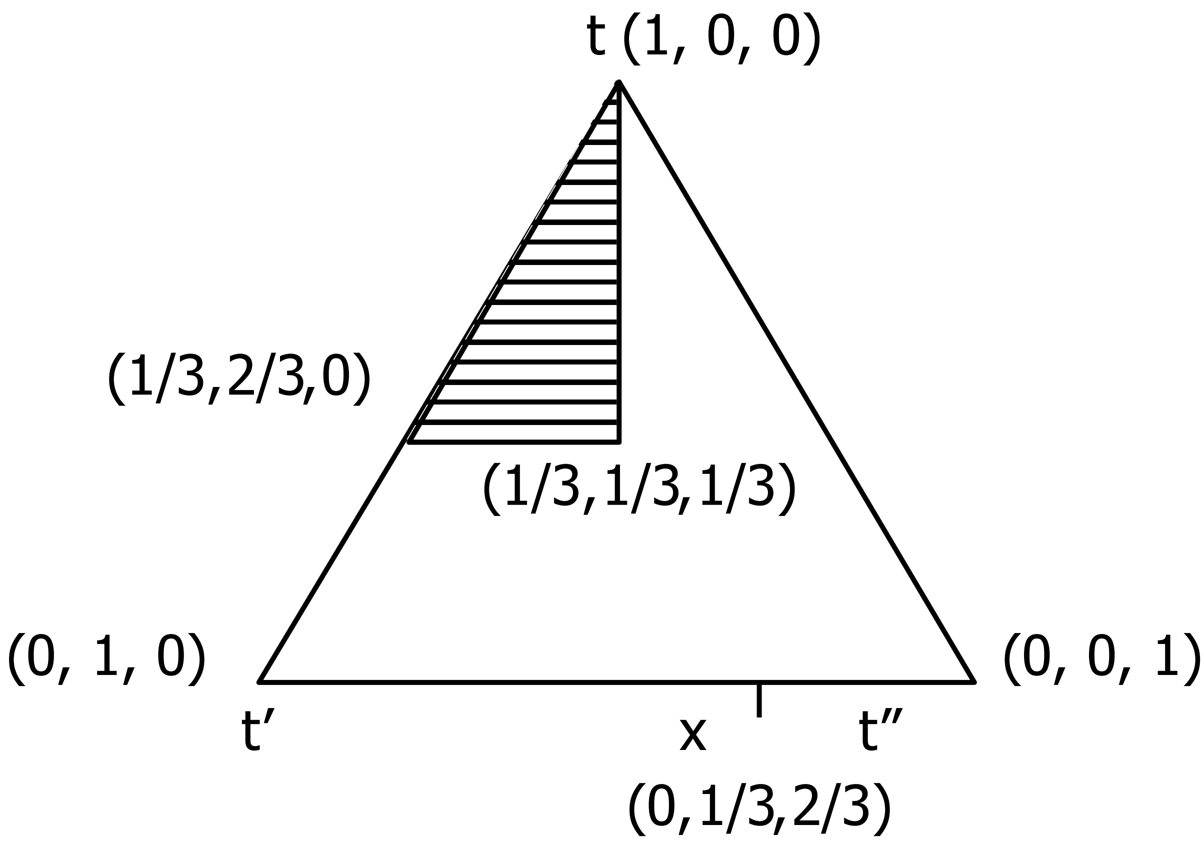
- Cho says that such beliefs satisfy *introspective consistency*. Further, a sequential equilibrium is a *forward induction equilibrium* if it is supported by beliefs satisfying introspective consistency.

Example


2, 1
u
l
t
r
u
0, 2
0, 0
d
1/3
m
0, 0
0, 2
d
R
R
2, 1
u
l
1/3
r
u
0, 2
4, 3
d
t'
m
0, 0
1,-1
d
2, 1
u
l
1/3
r
u
0,-3
1,-2
d
t''
m
0, 0
4, 0
d

Example

- R's choice off the equilibrium path is rationalized by beliefs in the shaded portion of the simplex below



Intuitive Criterion

- Let J = {t} and m' = r.
- Then (1) holds, but neither t' nor t'' satisfies (2).
- The beliefs that support the equilibrium do not satisfy introspective consistency: beliefs on the t'-t'' axis to the left of x cause R to play m, while beliefs to the right cause d; there is no belief over {t',t''} that rationalizes u for R.

Extended Intuitive Criterion

- The extended intuitive criterion consists of the following conditions:

  (1) for all $t \in J$, for all $a \in BR(T,m')$, $U^*(t) > U^S(t,m',a)$, and

  (2') for all $a \in BR(T\sim J,m')$ there exists $t' \in T\sim J$ such that $U^*(t') < U^S(t',m',a)$.

- A sequential equilibrium is to be rejected if it satisfies condition (1) and (2').

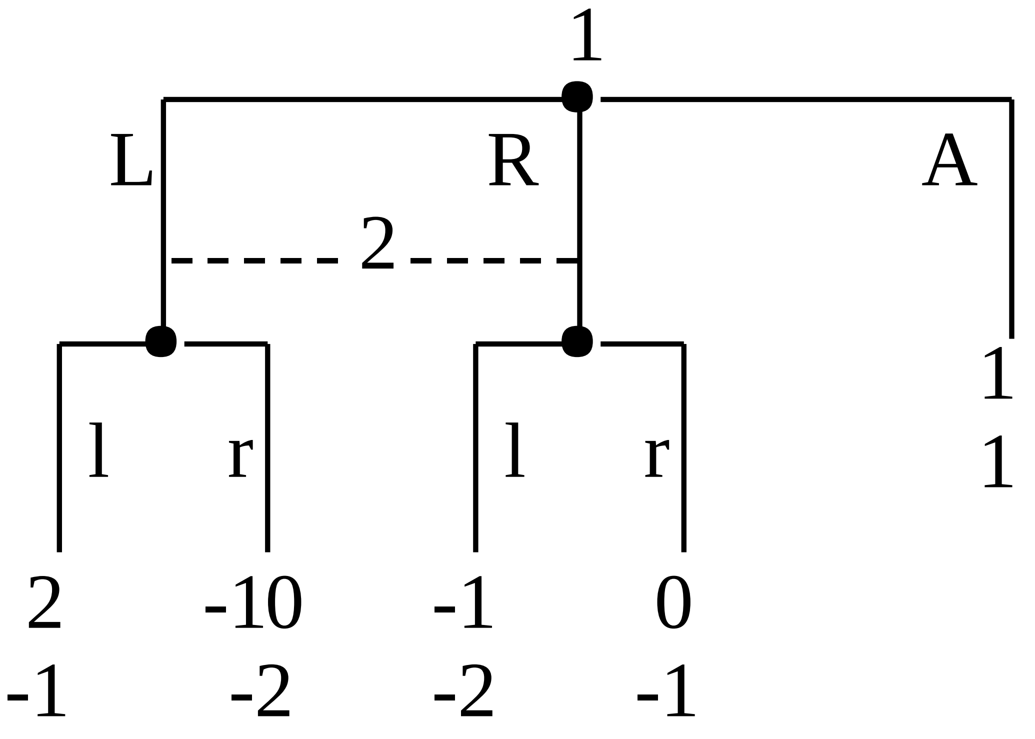
Proposition

- *A strategy profile π is a forward induction equilibrium if and only if π is a sequential equilibrium satisfying the extended intuitive criterion.*

Forward Induction and Sequential Equilibria

- For a general game, a deviation from a specified equilibrium is said to be "bad" if it *always* yields the deviator less than her equilibrium payoff in *every* circumstance.

Example



Forward Induction and Sequential Equilibria

- The SE (A,r) seems unreasonable because it requires player 2 to believe with high probability that player 1 has made a ridiculous deviation from the equilibrium.

- R would be a bad deviation for 1. Thus, (A,r) is not a forward induction equilibrium because it can be supported only by beliefs that assess positive probability that a bad deviation has occurred.

Forward Induction and Sequential Equilibria

Limitations:

In more complex games the set of bad deviations often is empty, in which case every sequential equilibrium is a forward induction equilibrium, and we must resort once again to ad hoc arguments to capture forward induction.

Divinity, D1, D2, Universal Divinity

Sender-receiver game

- S learns her type $t \in T$ and sends a message $m(t) \in M$ to R, who takes an action $a(m) \in A$.

"What should R infer from the message m?

Equilibrium Dominance

- The intuitive criterion and the forward induction equilibrium are based on the following dominance argument.

  Dominance: Eliminate t if m is sent and m is dominated by m' for t:

- This is too weak a requirement. At the very least, we should require that R's action a is a best response for some beliefs: $a \in BR(T,m)$. This is known as equilibrium dominance.

Dominance

- Fix an equilibrium with payoff to S of $U^*(t)$.
- For each (t,m) find the set of best responses by R that cause S to defect.
- Define D_t to be the set of best responses by R that make S strictly prefer defection:

$$D_t = \{a \in BR(T(m),m): U^*(t) < U(t,m,a)\}.$$

- Define D°_t to be the set of best responses by R that make S indifferent between defection and the equilibrium:

$$D^\circ_t = \{a \in BR(T(m),m): U^*(t) = U(t,m,a)\}.$$

D1 Refinement

D1. If $\exists$ t' with $D_t \cup D^\circ_t \subseteq D_{t'}$ then prune (t,m).

- D1 requires that zero weight be put on the type t if m is sent if there is another type t' such that t' always strictly benefits from the deviation whenever t benefits from the deviation.
- Cho and Sobel (1988) demonstrate that, for monotonic signaling games, the set of D1 equilibria is the same as the set of stable equilibria.
- If the single crossing property is satisfied, then D1 yields a unique equilibrium.

D2 Refinement

D2. If $D_t \cup D^{\circ}_t \subseteq \cup_{t' \neq t} D_{t'}$ then prune (t,m).

- D2 requires that zero weight be put on t when m is sent if for every best response that causes t to deviate there is a t' that strictly benefits from the deviation.

Divinity and Universal Divinity Banks and Sobel (*Econometrica*, 1988)

- *Divinity:* It is a weakening of D1. Rather than put zero weight on types t satisfying D1, divinity simply requires that the posterior belief after m is sent cannot increase the likelihood ratio of t to t'.
- *Universal divinity*: It is a strengthening of D2. It requires that t be eliminated, using an iterative application of D2.
- With universal divinity the updated beliefs do not depend on the prior; whereas, divine beliefs do depend on the prior.

Never Weak Response

Prune (t,m) if $D^{\circ}_{t} \subseteq \cup_{t' \neq t} D_{t'}$.

- This refinement is a slight strengthening of D2.
- t is given zero weight if whenever t is indifferent between deviating and following the equilibrium, there is a t' that strictly benefits from the deviation.

Spence Signaling Game

Which refinement do we need in the Spence signaling game to get a unique equilibrium?

- Two types: the intuitive criterion is enough to guarantee that the separating equilibrium is unique.
- More than two types: pooling can occur and it cannot be eliminated by the intuitive criterion. Applying D1 gives us a unique equilibrium.

Ordering of Refinements

It is possible to order the equilibrium refinements in terms of the set of equilibria they produce.

- For general games, we have

  NE $\supset$ SE $\supset$ PE $\supset$ ProperE $\supset$ IC $\supset$ EIC.

- For signalling games, we have

EIC$\supset$Div$\supset$D1$\supset$D2$\supset$UniDiv$\supset$NWBR$\supset$Stable$\neq\emptyset$.

Lecture Note 5:
Signaling

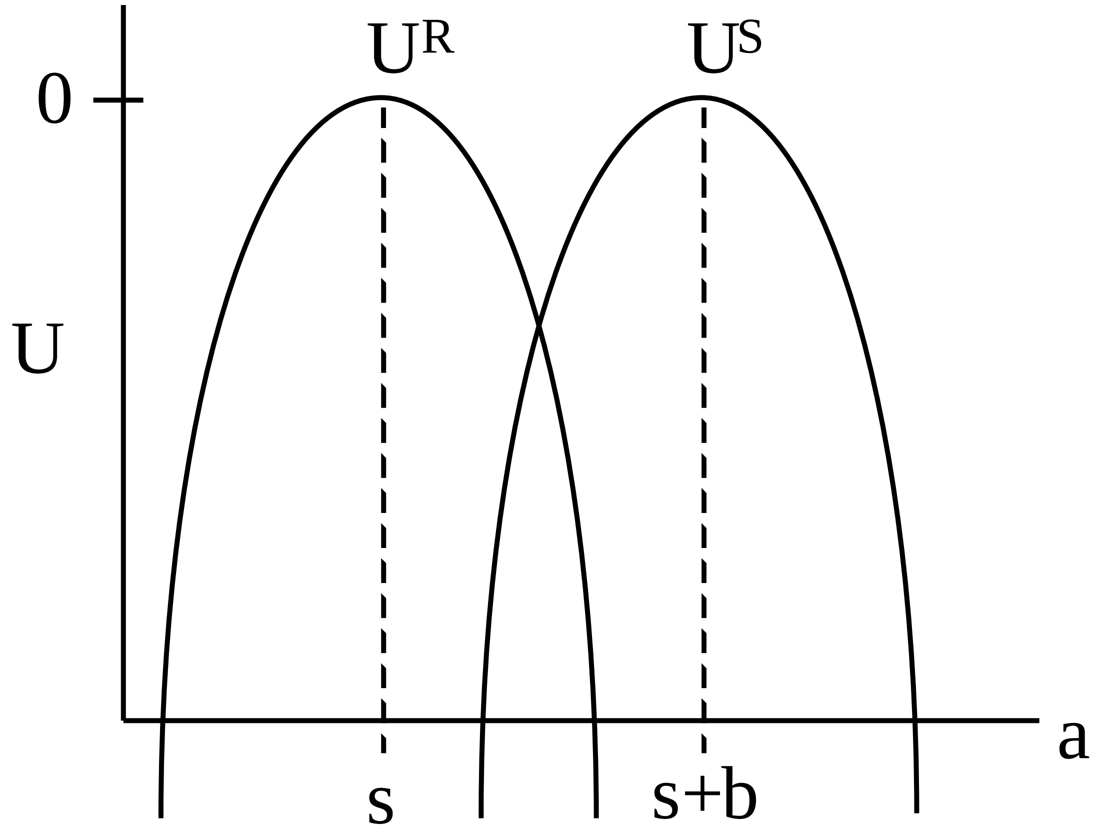
- Cheap Talk
  - Strategic Information Transmission
  - Neologisms
  - Perfect Sequential Equilibria

Strategic Information Transmission

- There are two parties, a Sender (S) and a Receiver (R) of information.
- Timing:
  1) S privately observes the state of the world, $s \in [0,1]$;
  2) S sends a message $m \in M$ to R; and
  3) R takes an action $a \in (-\infty,\infty)$.
- R does not observe the state s, but holds the prior belief that s has distribution F(s) on [0,1].
- Payoffs are $U^S(a,s,b)$ and $U^R(a,s)$; b measures how nearly the agents' interests coincide.

Example

$U^S(a,s,b) = -[a - (s+b)]^2$ and $U^R(a,s) = -[a - s]^2$



Payoffs

- The signal $m \in M$ is irrelevant to the payoff functions ("talk is cheap").
- Assumptions on U^i, for i= R, S
  - (i) $U^i_1 = 0$ for some a,
  - (ii) $U^i_{11} < 0$, and
  - (iii) $U^i_{12} > 0$.
- (i), (ii) and (iii) imply that given s there is a unique action, a, that maximizes U^i. Moreover, these optimal actions, $a^R(s)$ and $a^S(s,b)$, are continuous and strictly increasing functions of s.

Bayesian Game

- $\Gamma = \{A_S, A_R; T_S, T_R; p_S, p_R; U_S, U_R\}$

where

- $A_S = M$,
- $A_R = (-\infty, \infty)$
- $T_S = [0,1]$
- $T_R = \{0\}$
- $p_S = 1$ (no uncertainty)
- $p_R = f(s)$ (the density of F)
- U^S and U^R are as defined above.

Strategies

A strategy is a plan of action covering every contingency that might arise.

- For S, about whom R has incomplete information, a strategy is a function from types to actions: let $q(m|s)$ be the density of S's choice of m when the state is s.
- For R, whose type is known, a strategy must specify an action $a(m)$ for each signal $m \in M$ that R might receive.

Beliefs

- Beliefs are updated using Bayes' rule.
- If R conjectures that S chooses m according to the density $q(m|s)$ when the state is s, then Bayes' rule yields

$$p(s|m) = \frac{q(m|s)f(s)}{\int_0^1 q(m|t)f(t)dt}.$$

Bayesian Equilibrium

- The strategies $\{q(m|s), a(m)\}$ form a Bayesian equilibrium if the usual Nash conditions hold:

(1) for each $s \in [0,1]$, $\int_M q(m|s)dm = 1$ and if $m^* \in M$ is in the support $q(\cdot|s)$ then m^* solves

$$\max_{m \in M} U^S(a(m), s, b)$$; and

(2) for each $m \in M$, $a(m)$ solves ,

$$\max_a \int_0^1 U^R(a, s)p(s|m)ds$$

where $p(s|m)$ is given by Bayes' rule.

Existence

- Given an action rule a(m) and an arbitrary action $\bar{a} \in (-\infty, \infty)$ define the set $\bar{M} \equiv \{m \in M: a(m) = \bar{a}\}$.

- Say that an S-type $\bar{s}$ induces the action $\bar{a}$ in the equilibrium {q(m|s),a(m)} if

$$\int_{\bar{M}} q(m|\bar{s})dm > 0.$$

- $a^S(s,b)$ and $a^R(s)$ maximize U^S and U^R, respectively, and that these are continuous and increasing functions of s.

Existence

Lemma:

Suppose that b is such that no $s \in [0,1]$ satisfies $a^S(s,b) = a^R(s)$. Then there exists $\varepsilon > 0$ such that if u and v are actions induced in equilibrium then $|u - v| \geq \varepsilon$. Further, the set of actions induced in equilibrium is finite.

Existence

Proof:

- Take $u < v$. Let s_u and s_v be S-types that induce u and v, respectively. Then revealed preference yields

  $U^S(u,s_u,b) \geq U^S(v,s_u,b)$ and $U^S(v,s_v,b) \geq U^S(u,s_v,b)$.

- By the continuity of U^S in s there exists $\bar{s}$ such that

$$U^S(u,\bar{s},b) = U^S(v,\bar{s},b).$$

Existence

By the concavity of U^S in a,

(i) $u < a^S(\bar{s}, b) < v$.

(ii) u is not induced by any $s > \bar{s}$, and

(iii) v is not induced by any $s < \bar{s}$.

- The last two statements, together with the assumption that $U^R_{12} > 0$, imply

(iv) $u \leq a^R(\bar{s}) \leq v$.

- This is because: $a^R(s)$ is increasing; $a^R(\bar{s})$ would be R's action if S's type were certain to be $\bar{s}$; and u(v) signals to R that $s <(>) \bar{s}$.

Existence

- For instance, $\underset{a}{\text{argmax}}\, E_s[U^R(a,s) \mid s < \bar{s}] \le a^R(\bar{s})$.
- Since $a^R(s)$ and $a^S(s,b)$ are continuous functions of s, $|a^R(s) - a^S(s,b)|$ attains its minimum over $s \in [0,1]$. By hypothesis, this minimum is positive. Therefore, there exists ε such that $v - u \ge |a^R(\bar{s}) - a^S(\bar{s},b)| \ge \varepsilon > 0$.
- Since $a^R(s)$ is continuous and increasing, the set A of actions induced in equilibrium is bounded by $a^R(0)$ and $a^R(1)$, both of which must be finite.
- Therefore A is a finite set.

Remarks

- The Lemma applies to any message space M.
- It shows that in cheap-talk games with imperfectly aligned preferences there cannot be a separating equilibrium: unlike in Spence's signaling model, equilibrium communication is necessarily imperfect. This happens because there are no exogenous signaling costs. The only costs (or benefits) from signaling arise endogenously because different signals induce different actions.

Notation

- Let $\sigma(M) \equiv (\sigma_0(M),\ldots,\sigma_M(M))$ denote a partition of [0, 1] into M steps, where $0=\sigma_0(M)<\sigma_1(M)<\ldots<\sigma_M(M) = 1$.
- Where there is no possibility of confusion, denote $\sigma_i(M)$ simply by σ_i.
- For $0 \le \underline{\sigma}, \overline{\sigma} \le 1$, define

$$\bar{a}(\underline{\sigma},\overline{\sigma}) \equiv \begin{cases} \underset{a}{\operatorname{argmax}} \int_{\underline{s}}^{\bar{s}} U^R(a,s)f(s)ds & \text{if } \underline{\sigma} < \overline{\sigma}, \\ a^R(\underline{\sigma}) & \text{if } \underline{\sigma} = \overline{\sigma}. \end{cases}$$

Theorem

Suppose b is such that no $s \in [0,1]$ satisfies $a^S(s,b) = a^R(s)$. Then there exists a positive integer M(b) such that for every integer $M \in [1,M(b)]$ there exists an equilibrium {q(m|s),a(m)}, where for all $i \in \{1,...,M-1\}$

(i) $q(m|s) \sim U\,[\sigma_i,\sigma_{i+1}]$ when $s \in (\sigma_i,\sigma_{i+1}]$,

(ii) $U^S(\bar{a}(\sigma_i,\sigma_{i+1}),\sigma_i,b) = U^S(\bar{a}(\sigma_{i-1},\sigma_i), \sigma_i, b)$,

(iii) $a(m) = \bar{a}(\sigma_i,\sigma_{i+1})$ when $m_i \in (\sigma_i,\sigma_{i+1})$,

and $\sigma_0 = 0$ and $\sigma_M = 1$.

Proof (Sketch)

- Given (i), if R hears the message $m \in (\sigma_i, \sigma_{i+1})$, the posterior belief is simply

$$p(m|s) = \frac{f(s)}{\int_{s_i}^{s_{i+1}} f(t)dt}.$$

Proof (Sketch)

- R's strategy specified in (iii) is a best response to S's strategy given in (i). As for S, (ii) guarantees that the S-type $s=\sigma_i$ is indifferent between the actions $\bar{a}(\sigma_i,\sigma_{i+1})$ and $\bar{a}(\sigma_{i-1},\sigma_i)$.

- Types s >(<) σ_i strictly prefer the latter (former) to the former (latter). Moreover, S-types $s \in (\sigma_i,\sigma_{i+1})$ strictly prefer $\bar{a}(\sigma_i,\sigma_{i+1})$ to any of the other actions $\bar{a}(\sigma_j,\sigma_{j+1})$ induced by (iii). That is, S's strategy is a best response to R's.

Proof (Sketch)

- Let the set of these actions be $A = \{a_j\}^J_{j=1}$, where $a_j < a_{j+1}$ for all $j < J$. As shown in the proof of the Lemma, for each pair (a_j,a_{j+1}) there exists an S-type s_j satisfying

  (*) $$U^S(a_j,s_j,b) = U^S(a_{j+1},s_j,b).$$

- Concavity of U^S implies that s_j strictly prefers either of a_j or a_{j+1} to any other $a_k \in A$, and that S-types $s \in (s_{j-1},s_j)$ strictly prefer a_j to any other $a_k \in A$, including a_{j+1}.

Proof (Sketch)

- Given the conjecture $a(m)$ about R's behavior, each S-type $s \in (s_{j-1},s_j)$ will send the signal $m_j \in M$ that induces the action a_j via $a_j = a(m_j)$.
- In equilibrium, R holds a correct conjecture about which S-types send m_j, so $a(m_j)$ must be R's best response to this belief, namely $\bar{a}(s_{j-1},s_j)$. Thus (*) is exactly (ii), and (i) and (iii) are rephrased in terms of the general signal space M as

  (i) S sends m_j when $s \in (s_{j-1},s_j)$, and

  (ii) $a(m_j) = a(s_{j-1},s_j)$.

Example

Determination of M(b) and construction of solutions to the difference equation (ii) for integers $M \in [1, M(b)]$:

– $U^S(a,s,b) = -[a - (s+b)]^2$
– $U^R(a,s) = -[a - s]^2$
– $F(s) \sim U\ [0,1]$

- Then $\bar{a}(\sigma_i, \sigma_{i+1}) = (\sigma_i + \sigma_{i+1})/2$, so (ii) becomes

$$\sigma_{i+1} = 2\sigma_i - \sigma_{i-1} + 4b.$$

- Solution: $\sigma_i = i\sigma_1 + 2i(i - 1)b$, for any σ_i.

Example

- Substituting i = M and $\sigma_M = 1$ yields

$$1 = M\sigma_1 + 2M(M - 1)b.$$

- M(b) is the largest integer less than

$$[1+(1+2/b)^{1/2}]/2.$$

- $M(b) \to \infty$ as $b \to 0$, but M(b) = 1 for $b \geq 1/4$.

In this example, more communication is possible when preferences are more similar

Example

- Suppose b = 1/20. Then M(b) = 3.
- Two-step equilibrium: {0, 2/5, 1}
- Three-step equilibrium: {0, 2/15, 7/15, 1}.

- None of the equilibria yields expected payoffs that Pareto-dominate the payoffs associated with the other equilibria: players have a problem knowing which equilibrium to play.

Forward Induction Equilibria in Cheap Talk Games

In a signaling game where talk is cheap, every sequential equilibrium is a forward induction equilibrium.

- Cheap talk means $U^S(t,m,a)$ and $U^R(t,m,a)$ are independent of m.
- The set of types for whom m' is a bad deviation from the equilibrium strategy π is empty, so communicational consistency has no cutting power.

Forward Induction Equilibria in Cheap-Talk Games

Mathematically:

- $J(m'|\pi) \equiv \{t \in T | U^*(t) > U^S(t,m',a)$ for all $a \in BR(T,m')\}$.

- Since $U^R(t,m,a)$ is independent of m, BR(T,m') becomes BR(T). Among the actions in BR(T) is the action a(t) that type t induces in equilibrium by sending the message m(t). Since $U^S(t,m,a)$ is independent of m,

$$U^S(t,m',a(t)) = U^S(t,m(t),a(t)) \equiv U^*(t).$$

- So $J(m'|\pi) = \varnothing$ for all m' and π.

Multiple Equilibria and Refinements

- Crawford and Sobel show that there typically are multiple sequential equilibria in cheap-talk games.

- None of the refinements considered before work for cheap-talk games.

Neologisms (Farrel, 1985)

- *Neologism*: it is an unsent message in a signaling game.
- Roughly speaking, a neologism is *credible* if those S-types that might send this unexpected message can make a persuasive speech to R, along the lines envisioned by Cho & Kreps.

Game

- Timing:

  1. Nature draws a type t from a finite set T for the Sender, S;
  2. S learns t and sends a message $m \in M^*$ to the Receiver, R; and
  3. R observes m and takes an action $a \in A$, where A is finite.
- Because talk is cheap, the payoffs are $U^S(t,a)$ and $U^R(t,a)$, independent of m.
- M^* is infinite but discrete.

Test for Selecting Equilibria

- A sequential equilibrium is reasonable if and only if it is neologism-proof.

- Equivalently, a sequential equilibrium should be rejected if and only if there is a credible neologism S could sent to R. It remains to define a credible neologism.

Test for Selecting Equilibria

- Let X be a non-empty subset of T.
- Let μ(t|X) be the distribution over types t $\in$ X that results from restricting the prior distribution of types μ(t) to X:

$$\mu(\mathrm{t}|\mathrm{X}) = \begin{cases} \dfrac{\mu(\mathrm{t})}{\sum_{\tau \in \mathrm{X}} \mu(\tau)} & \text{if t} \in \mathrm{X} \\ 0 & \text{if t} \notin \mathrm{X}. \end{cases}$$

Test for Selecting Equilibria

- Let $a^*(X)$ solve

$$\max_{a \in A} \sum_{t \in X} \mu(t|X)U^R(t,a),$$

  and assume $a^*(X)$ is unique.

- If R's beliefs are $\mu(t|X)$ then $a^*(X)$ will follow.

- S's payoff: $V(X,t) \equiv U^S(a^*(X),t)$.

Equilibrium Payoff

- Let R's behavioral strategy be $\pi^R(a|m)$, which specifies a distribution over $a \in A$ for each $m \in M^*$ that might be observed.

- Then t's best response m(t) yields the payoff

$$\max_{m \in M^*} \sum_{a \in A} \pi^R(a|m) \cdot U^S(a,t) \equiv U^*(t)$$

Notation

Define:

- $K(X|\pi) \equiv \{t \in T \mid U^*(t) < V(X,t)\}$:

  Set of types who would deviate from the equilibrium π if in so doing they led R to hold the belief $\mu(t|X)$.

- $J(m'|\pi) \equiv \{t \in T \mid U^*(t) > U^S(t,m',a) \text{ for all } a \in BR(T,m')\}$:

  Set of types who would *not* deviate from π by sending m', no matter what belief this induced R to hold.

Notation

- Farrell says that a subset X of T is *self-signaling* given the equilibrium π if K(X|π) = X, and that the neologism (unsent message) "t is in X" is *credible* if X is self-signaling.

Notation

If t $\in$ X then S says (or R reasons):

"My type is in X. Moreover, every other type in X and no type outside X has an incentive to make this speech. For if you believe it then your belief should be μ(t|X), so your action should by a^*(X), which only we types in X would prefer to our equilibrium payoff."

Selecting Equilibria

- Given a sequential equilibrium π, if there exists a credible neologism then Farrell rejects the equilibrium.
- If there does not exist a credible neologism then Farrell says that π is *neologism-proof* and accepts it.

- Problems: there may not exist a neologism-proof equilibrium

Example

- Let $T=\{t_1,t_2\}$, $\mu(t_1)=\mu(t_2)=1/2$, $A=\{a_1,a_2,a_3\}$, and let the payoffs be as given below.

U^S

	t_1	t_2
a_1	2	-1
a_2	-1	-2
a_3	0	0

U^R

	t_1	t_2
a_1	3	0
a_2	0	3
a_3	2	2

Example

Pooling Equilibrium:

- In any pooling equilibrium, R will play a_3.

- The set $X = \{t_1\}$ is self-signaling: if R believes the neologism "$t \in X$" then a_1 will replace a_3 as a best response; this yields a payoff of 2 for t_1 (-1 for t_2), which is better (worse) than the equilibrium payoff of 0. Thus, Farrell rejects all the sequential equilibria in this game.

Perfect Sequential Equilibria
(Grossman and Perry, *JET*, 1986)

- In refinements based on equilibrium dominance, the beliefs following a deviation do not rationalize the deviation in an equilibrium sense.
- G&P take the view that once a deviation has occurred, the other should try to rationalize the deviation by trying to find a set of types K $\subseteq$ T that benefit from the deviation if it is thought K deviated, but t $\notin$ K lose from the deviation.
- If such a K exists, then the beliefs following the deviation should require that the receiver infers that K deviated.

Perfect Sequential Equilibria (PSE)

PSE can be motivated from NE and SE concepts:

- NE requires best responses along the equilibrium path (players can threaten with actions).
- SE requires best responses at every information set given beliefs (players can threaten with beliefs).
- PSE requires best responses at every information set *for all* beliefs (it limits a player's ability to threaten with beliefs).

Definitions

- A *metastrategy* is an action to take at each information set and all beliefs ($\sigma^R(m,\mu) \in A$).

- An *updating rule* maps the message m and prior p into a posterior belief $\mu = g(m,p)$.

The heart of the PSE refinement is to place a restriction on the updating rule.

Definitions

- A strategy profile and updating rule (σ,g) is a *PSE* if $\forall$ information sets and $\forall$ beliefs, σ is a best response and g is *credible*.
- The updating rule g is *credible* if:

(a) the support of the posterior is contained in the support of the prior,

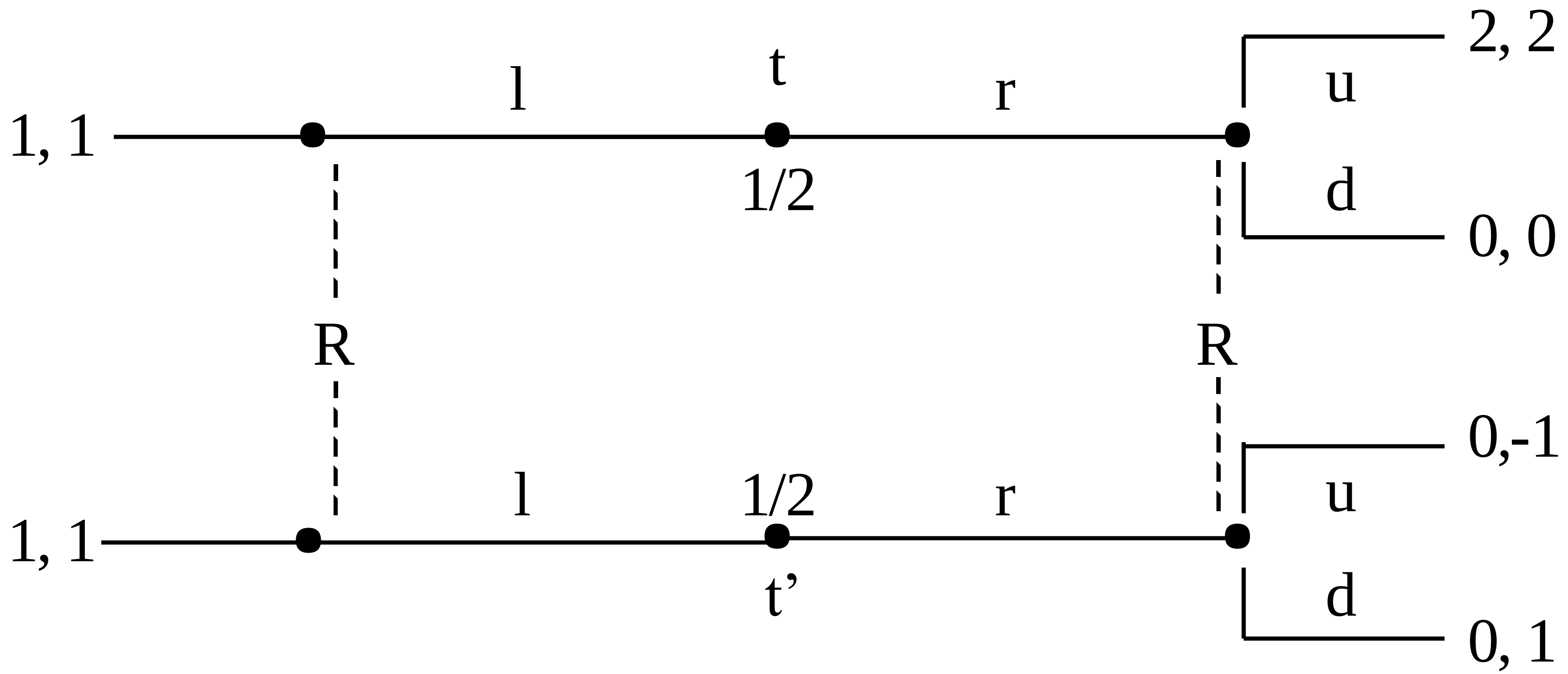
(b) if $\exists$ K such that

$$U^S(t,m,\sigma^R(m,p_K)) \geq U^*(t) \ \forall \ t \in K$$
$$U^S(t,m,\sigma^R(m,p_K)) \leq U^*(t) \ \forall \ t \notin K,$$

then $g(m,p) = p_K = p(t)/(\Sigma_{\tau \in K} p(\tau))$.

(c) use Bayes rule when possible.

Example 1

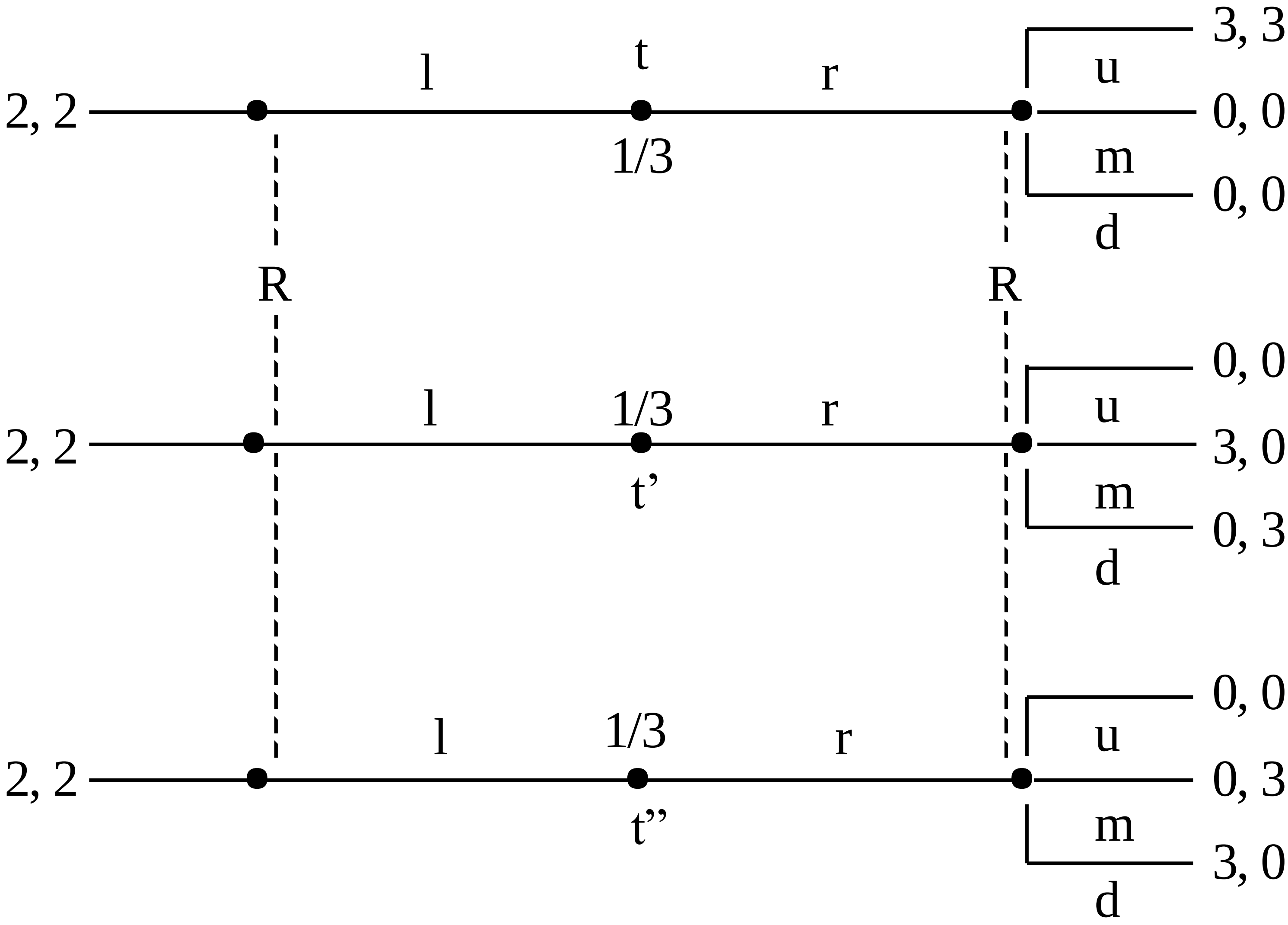


Example 1

SE: (1) ll,d, and (2) rl,u. Only (2) is a PSE.

- (1) requires that R put sufficiently high weight on t' if r is played, so that d is a best response. But K={t} rationalizes the deviation r, since t benefits if thought to be t by playing r and t' would not want to deviate if thought to be t'.
- (2) is the only equilibrium satisfying the intuitive criterion as well, since r is a bad deviation for t': r yields 0 for t' regardless of R's response vs. 1 in the equilibrium.

Example 2

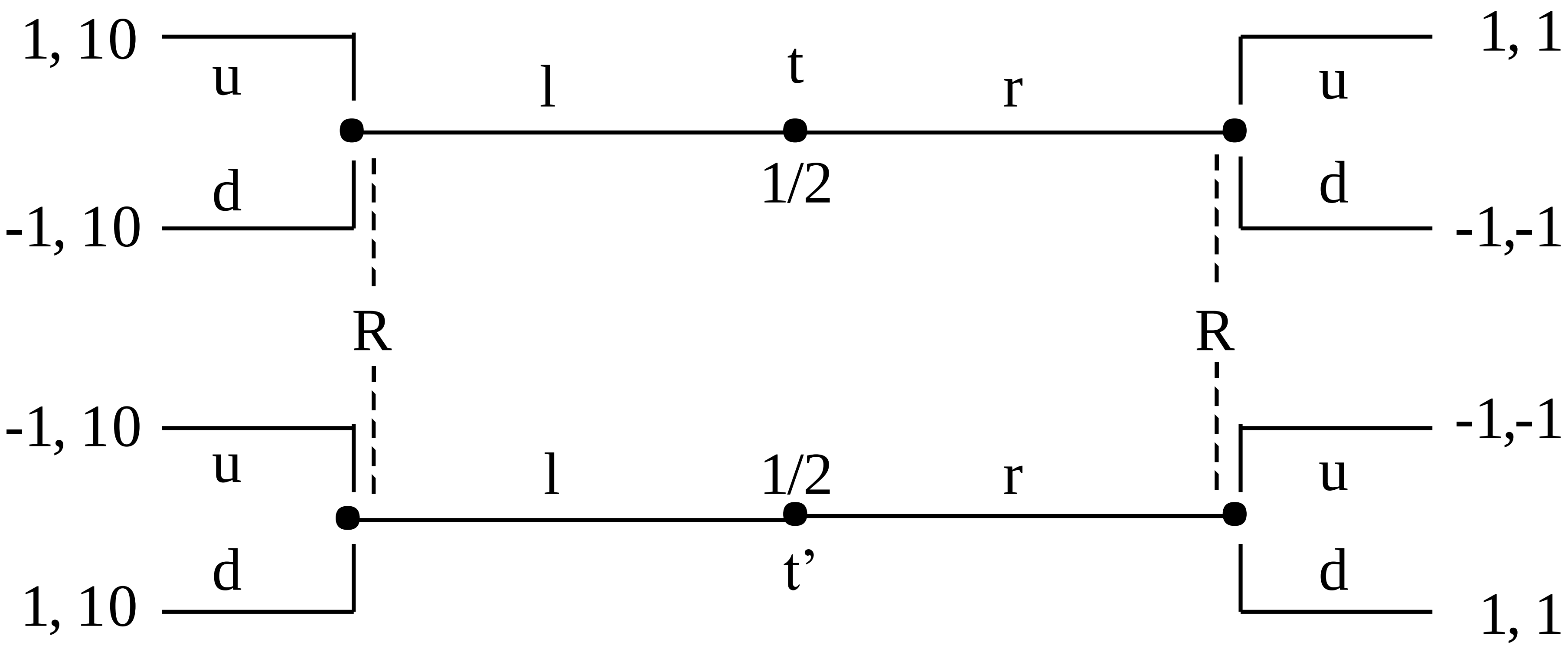

3, 3
t
l
r
u
2, 2
0, 0
1/3
m
0, 0
d
R
R
0, 0
l
1/3
r
u
2, 2
3, 0
t'
m
0, 3
d
0, 0
l
1/3
r
u
2, 2
0, 3
t''
m
3, 0
d

Example 2

SE: (1) lll, and (2) rll,u. Only (2) is a PSE.

- (1) is not a PSE, because there is a unique rationalization of the deviation r. K={t} rationalizes r, since if R infers t from r then R's best response is u, yielding 3 rather than the equilibrium payoff of 2 to S. Neither t' nor t'' benefit from the deviation, since they would get 0 rather than 2.
- (1) is not rejected by the intuitive criterion, since r is not a bad deviation for either t' or t”.

Example 3



Example 3

Three SE: (1) (rl; ud), (2) (lr,ud), (3) (ll,1/2(ud)). All three are PSE.

- The two separating equilibria, (1) and (2), are PSE, since they are SE and all messages are sent in equilibrium, so beliefs are uniquely defined from Bayes' rule.
- The pooling equilibrium is PSE, since the deviation can be rationalized by the inference K = {t,t'}.

The three different credible updating rules lead to three different PSE.

Example 3

- Farrell would reject the pooling equilibrium, since the updating rules that lead to the separating equilibria *strictly* rationalize the deviation in the sense that the deviator strictly gains from the deviation with an inference of either K = {t} or K = {t'}, whereas no type strictly gains from deviating with K = {t,t'}.

- Farrell's concept in some sense lets the sender pick the credible updating rule by sending the neologism "my type is in K." This may be reasonable if the setting allows the sender to communicate in this way.