Solutions Key to
Problem Sets in Game Theory

Marina Sanchez del Villar *

Winter 2020

Problem Set 4: Bayesian Games

Exercise 1: The Chain Store Paradox Paradox

Consider the Kreps-Wilson version of the chain store paradox: An entrant may stay out and get nothing (0), or he may enter. If he enters, the incumbent may fight or acquiesce. The entrant gets b if the incumbent acquiesces, and $b-1$ *if he fights, where* $0<b<1$*. There are two types of incumbent, both receiving* $a>1$ *if there is no entry. If there is a fight, the strong incumbent gets* 0 *and the weak incumbent gets* -1*; if a strong incumbent acquiesces he gets* -1*, a weak incumbent* 0*. Only the incumbent knows whether he is weak or strong; it is common knowledge that the entrant a priori believes that he has a chance* p_0 *of facing a strong incumbent. Define:* [1]

$$\gamma = \frac{p_0}{1-p_0}\frac{1-b}{b}$$

a. *Sketch the extensive form of this game.*

b. *Define a sequential equilibrium of this game.*

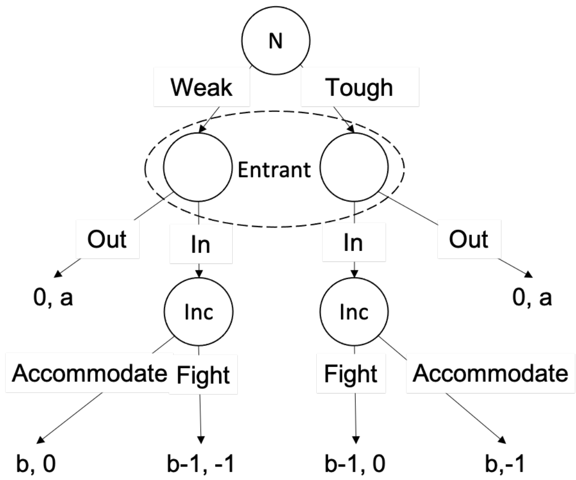
A sequential equilibrium of this game is a pair of strategies (σ) and beliefs (μ_E):

$$\Big(\sigma = (\sigma_E \in \Delta\{O, I\}, \sigma_I \in \Delta\{F, A\}); \mu_E(S) \in \Delta\{S, W\}\Big) \qquad \text{such that:}$$

- σ is sequentially rational given $\mu_E(S)$;
- Consistency of beliefs: there exists a fully mixed strategy converging to the actual strategy $[\{\sigma^k\}_{k=1}^{\infty}$ with $\lim_{k\to\infty}\sigma^k = \sigma]$ such that the corresponding sequence of beliefs converging to that belief is derived through Bayes rule: $\mu = \lim_{k\to\infty}\mu^k$

*This version builds on the solutions provided by Damiano Argan and Konuray Mutluer.

[1] p_0 is the probability that the incumbent is strong, assigned by nature, while $\mu_E(S)$ is a belief by the entrant. In a pooling equilibrium we will see that $\mu_E(S) = p_0$.



c. *Show that if $\gamma \neq 1$, there is a unique sequential equilibrium, and that if $\gamma > 1$ entry never occurs, while if $\gamma < 1$ entry always occurs.*

For which beliefs does the entrant enter? Note that, when it is the incumbent's time to move:

- Incumbent plays F if strong: $BR_I(In|S) = F$
- Incumbent plays A if weak: $BR_I(In|W) = A$

Hence, the entrant faces the following choice:

$$\mathbb{E}\Big[u_E(In, BR_I(I))\Big] \gtrless \mathbb{E}\Big[u_E(Out)\Big]$$
$$\mu_E(S)(b-1) + (1-\mu_E(S))b \gtrless 0$$
$$\frac{\mu_E(S)}{1-\mu_E(S)}\frac{1-b}{b} \gtrless 1$$

The entrant decides to enter if:
$$\frac{\mu_E(S)}{1-\mu_E(S)} < \frac{b}{1-b}$$

The beliefs supporting this strategy must be derived from Bayes Rules:

$$\mu_E(S) = \Pr(S|\text{nature's move}) = p_0$$

If the game is played once, the Incumbent's strategy is of no consequence for consistent beliefs. Hence, the belief must be equal to the prior. [2]

Substituting these beliefs in the previous condition for the entrant to choose In:

$$\frac{p_0}{1-p_0}\frac{1-b}{b} < 1$$

Therefore, $\forall \gamma \neq 1$, $\exists$ a unique sequential equilibrium where:

[2]Should we check for a converging sequence? In this case, as there are no beliefs involved in the BR of the players, it is not needed. Furthermore, notice that nature is already mixing: with p_0 incumbent is strong, with $(1-p_0)$ it is weak. In the next question, where beliefs enter the BRs, we will not need to check for a fully mixed sequence because it is already provided by nature.

- $\gamma < 1$: Entrant always enters:

$$\Big(\sigma = (In; F \text{ if strong}, A \text{ if weak}); \frac{\mu_E(S)}{1-\mu_E(S)}\frac{1-b}{b} < 1, \mu_E(S) = p_0\Big)$$

- $\gamma > 1$: Entrant never enters:

$$\Big(\sigma = (O; F \text{ if strong}, A \text{ if weak}); \frac{\mu_E(S)}{1-\mu_E(S)}\frac{1-b}{b} > 1, \mu_E(S) = p_0\Big)$$

d. *What are the sequential equilibria if $\gamma = 1$?*

The entrant is now indifferent between entering or not given the prior beliefs on the incumbent's type.

$$\mathbb{E}\Big[u_E(In|BR_I(I))\Big] = \mathbb{E}\Big[u_E(Out)\Big]$$
$$\mu_E(S)(b-1) + (1-\mu_E(S))b = 0$$
$$\frac{\mu_E(S)}{1-\mu_E(S)}\frac{1-b}{b} = 1$$

Any strategy profile where the incumbent fights if strong and gives in if weak is a sequential equilibrium supported by beliefs identical to the prior:

$$\Big(\sigma = (\sigma_E \in \Delta\{Out, In\}; F \text{ if strong}, A \text{ if weak}); \frac{\mu_E(S)}{1-\mu_E(S)}\frac{1-b}{b} = 1, \mu_E(S) = p_0\Big)$$

e. *Now suppose that the incumbent plays a second round against a different entrant who knows the result of the first round. The incumbent's goal is to maximize the sum of his payoffs in the two rounds. Show that if $\gamma > 1$ there is a sequential equilibrium in which the entrant enters on the first round and both types of incumbents acquiesce. Be careful to specify both the equilibrium strategies and beliefs.*

We are looking at an equilibrium that, with $\gamma > 1$, the entrant will enter In_1 in the first round (as opposed to the solution in part *c.*), stay out in the second round and both types of incumbents in the first round will accommodate A_1 (pooling equilibrium!).

<u>Second round</u>:

- Incumbent: will play A_2 if weak and F_2 if strong.
- Entrant $E2$: his information set will be the action he observes by the incumbent in period 1: accommodate or fight.

  We need to construct some beliefs that support the entrant staying out in the second round, while $\gamma > 1$ and $E1$ entering. The beliefs have to be the opposite of those in question b), where the entrant E_1 stayed out if $\gamma > 1$.

  - Information set when A_1: on path, part of the pooling equilibrium. the posterior must be equal to the prior. [3]

$$\mu_E(S|A_1) = p_0 \qquad \mu_E(W|A_1) = 1 - p_0$$

[3] Notice that the beliefs in each information set sum up to one.

Given beliefs, the BR by the entrant in the second round upon observing accommodate is to stay out. Note that this is the case because the question tells us that $\gamma > 1$. $BR_{E2}(A_1, \mu_E(S|A_1)) = O_2$

  - Information set when F_1: off path. These beliefs will help us construct the story we want to tell. We need beliefs on types supporting that if the entrant sees fight, he will choose E_2. This is why the incumbent will end up playing accommodate in the first round.

$$\mu_E(S|F_1) = 0 \qquad \mu_E(W|F_1) = 1$$

  Given these beliefs, $BR_{E2}(F_1, \mu_E(W|F_1)) = In_2$

First round:

- Incumbent: we need to consider the sum of the payoffs in both cases.
  - Strong: it is optimal for the strong incumbent to give in in the first round.

$$u_{I1}(A_1; In, BR_{E1}(A_1)|S) = -1 + a > 0 = u_{I1}(F_1; In, BR_{E1}(F_1)|S)$$

  - Weak: it is also optimal for the weak incumbent to give in.

$$u_{I1}(A_1; In, BR_{E1}(A_1)|W) = +a > -1 + 0 = u_{I1}(F_1; In, BR_{E1}(F_1)|W)$$

- Entrant: since the incumbent will always play give in, the entrant finds it profitable to play In in the first round. This strategy is sequentially rational.

Given these beliefs, we found a strategy that with $\gamma > 1$ the first entrant enters and the incumbent always accommodates.

We now need to check two things for these beliefs: that they are derived using Bayes rule / Bayesian updating and that they are consistent (via a converging sequence).

Bayesian updating: on-equilibrium path $\mu_E(S|A_1)$
We check that the beliefs are induced by Bayes rules according to the first round of incumbent strategies.

$$\mu_E(S|A_1) = \Pr(S|A_1, \sigma) = \frac{\Pr(A_1|S, \sigma)\Pr(S|\sigma)}{\Pr(A_1|\sigma)} = \frac{1 \cdot p_0}{1} = p_0$$

Since both types play A_1, no new information is revealed by playing A_1 in the first round. In WPBE we do not need to check for convergence of the nodes on the information sets that are on the equilibrium path. Since $\mu_E(S|A_1)$ is on the equilibrium path, we do not need to check for convergence of a sequence. It exists by definition.

Converging sequence: off-equilibrium path $\mu_E(S|F_1)$
$\mu_E(S|F_1)$ is off the equilibrium path. We cannot compute its probability with Bayes, so we check for consistency using the converging sequence. The sequence must rely on a fully mixed strategy.

Consider:

$$\sigma^n = \{\sigma^n_{I1}(F_1|W), \sigma^n_{I1}(F_1|S)\} = \{\varepsilon^n, \varepsilon^{2n}\} \qquad \varepsilon \text{ small}$$

So that $\lim_{n\to\infty} \varepsilon^n = 0$. This means that, at the limit, both types of incumbents will play accommodate. This sequence therefore converges to the equilibrium strategy of the incumbent. The induced beliefs μ_E^n must be calculated via Bayesian updating according to σ^n. [4]

$$\mu_E^n(S|F_1) = \Pr(S|F_1, \sigma^n) = \frac{\Pr(S|\sigma^n)\Pr(F|S_1, \sigma^n)}{\Pr(F|\sigma^n)} \underset{\text{LOTP}}{=} \frac{p_0 \cdot \varepsilon^{2n}}{p_0\varepsilon^{2n} + (1-p_0)\varepsilon^n}$$
$$= \frac{\varepsilon^n}{\varepsilon^n} \frac{p_0\varepsilon^n}{p_0\varepsilon^n + (1-p_0)}$$
$$\lim_{n\to\infty} \mu_E^n(S|F_1) = \lim_{n\to\infty} \frac{p_0\varepsilon^n}{p_0\varepsilon^n + (1-p_0)} = 0$$

This concludes the proof of the sequential equilibrium.

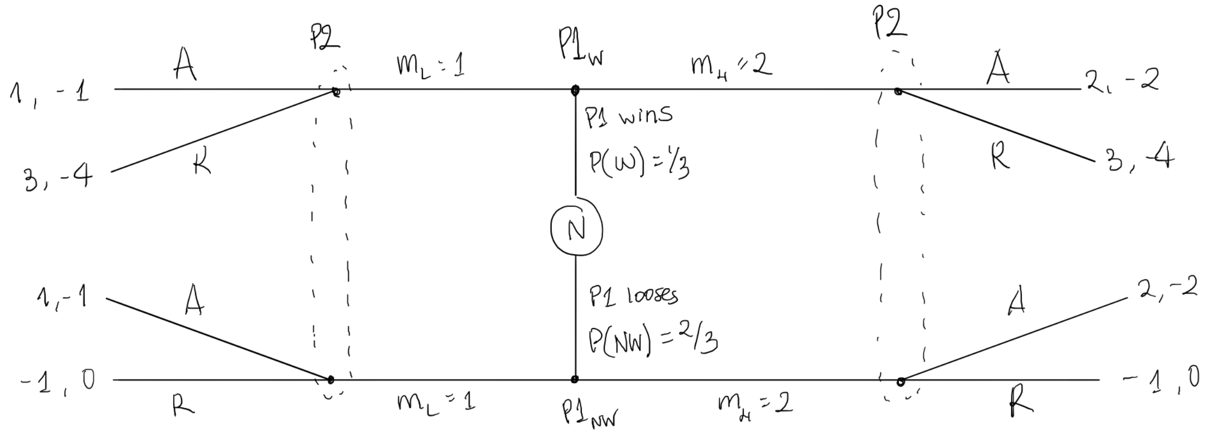
To wrap-up, let us state the sequential equilibrium we have derived:

$$\{\sigma = (A_1|S, A_1|W, F_2|S, A_2|W; In_1, In_2|F_1, Out_2|A_1)\,; \mu_{E1}(S) = p_0, \mu_{E2}(S|A_1) = p_0, mu_{E2}(S|F_1) = 0\}$$

Exercise 2: Courtroom Drama

Two players: plaintiff and defendant, in a civil suit. The plaintiff knows whether or not he will win the case if it goes to trial, but the defendant does not. The defendant's beliefs are Pr(*plaintiff wins*) = 1/3*. This is common knowledge. The cost of the trial is* 1*. The loser of the trial bears this cost. If the plaintiff wins the trial, then the defendant will have to pay the plaintiff* 3 *and also pay for the cost of the trial. If the plaintiff loses, he'll have to pay for the cost of the trial, and the defendant will neither win anything nor lose anything. The plaintiff has two actions: ask for a low settlement,* $m = 1$*, or ask for a high settlement,* $m = 2$*. If the defendant accepts* m*, then the defendant is agreeing to pay* m *to the plaintiff out of court. If the defendant rejects* m*, the case goes to court.*

(a) *Draw the game tree.*



---

[4]Bayesian updating is updating following Bayes rule:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

(b) *Find all sequential equilibria.*

Any sequential equilibria is also a WPBE. [5] We will first look at the WPBE where both types of plaintiffs P1 play pure strategies, to then move to mixed strategies.

Pure strategies:

Let's compute the BR function of the defendant P2 at each information set given his beliefs. An information set will refer to the settlement offer he receives m_H or m_L. [6]

- High settlement: the defendant accepts if:

$$u_D(A, \mu(W|m_H)) = -2 > -4\mu(W|m_H) + 0(1 - \mu(W|m_H)) = u_D(R, \mu(W|m_H))$$

Accept when $\mu(W|m_H) > 1/2$.

$$BR_D(m_H, \mu(W|m_H)) = \begin{cases} A & \text{if } \mu(w|m_H) > 1/2 \\ \Delta\{A, R\} & \text{if } \mu(w|m_H) = 1/2 \\ R & \text{if } \mu(w|m_H) < 1/2 \end{cases}$$

- Low settlement: the defendant accepts when:

$$u_D(A, \mu(W|m_L)) = -1 > -4\mu(W|m_L) + 0(1 - \mu(W|m_L)) = u_D(R, \mu(W|m_L))$$

Accept when $\mu(W|m_L) > 1/4$.

$$BR_D(m_L, \mu(W|m_L)) = \begin{cases} A & \text{if } \mu(w|m_L) > 1/4 \\ \Delta\{A, R\} & \text{if } \mu(w|m_L) = 1/4 \\ R & \text{if } \mu(w|m_L) < 1/4 \end{cases}$$

Depending on the plaintiff's strategy, we will have two types of equilibria.

1) Separating equilibria:
   A. $\sigma_P^A = (m_H \text{ when W}, m_L \text{ when NW})$
   B. $\sigma_P^B = (m_L \text{ when W}, m_H \text{ when NW})$
2) Pooling equilibria
   A. $\sigma_P^A = (m_L \text{ when W}, m_L \text{ when NW})$
   B. $\sigma_P^B = (m_H \text{ when W}, m_H \text{ when NW})$

Let's go through each in turns.

1) Separating equilibria:

[5] Remember that sequential equilibria are a subset of WPBE. That is, if an equlibrium is not WBPE it cannot be a sequential equilibrium.

[6] We had a discussion in terms of notation during the TA class on whether we could write $BR_D(m|\mu(W|m))$. After reading some papers over the break, I have decided to write the solutions key in the same way as those papers: $BR_D(m,\mu(W|m))$

A. $\sigma_P^A = (m_H \text{ when W}, m_L \text{ when NW})$

The beliefs of the defendant have to be derived by Bayes rules. Given the plaintiff's equilibrium strategy, the beliefs should have $\mu(W|m_L) = 1$, else there would be a profitable deviation.

$$\begin{aligned} \mu(W|m_H) &= \Pr\left(W|m_H, \sigma_P^A\right) = 1 \quad \rightarrow \quad BR_D = (m_H|\mu(W|m_H)) = A \\ \mu(W|m_L) &= \Pr\left(W|m_L, \sigma_P^A\right) = 0 \quad \rightarrow \quad BR_D = (m_L|\mu(W|m_L)) = R \end{aligned}$$

The defendant will play A when he receives m_H, believing that the plaintiff will win, and R when seeing m_L.

The plaintiff, however, wants to be rejected when he has $\Pr(W) = 1$, and accepted if he is a looser type.

With this strategy, both types have a profitable deviation. [7] The strategy σ_P^A for the plaintiff is not sequentially rational.

B. $\sigma_P^B = (m_L \text{ when W}, m_H \text{ when NW})$

We will derive the defendant's beliefs by Bayes rule:

$$\begin{aligned} \mu(W|m_H) &= \Pr\left(W|m_H, \sigma_P^B\right) = 0 \quad \rightarrow \quad BR_D = (m_H|\mu(W|m_H)) = R \\ \mu(W|m_L) &= \Pr\left(W|m_L, \sigma_P^B\right) = 1 \quad \rightarrow \quad BR_D = (m_L|\mu(W|m_L)) = A \end{aligned}$$

The defendant will A when he sees m_L, as he beliefs the plaintiff will win, and R when he sees m_H.

Both types of plaintiffs will have a profitable deviation. For example, the looser has the option to offer m_L and the defendant will accept, yielding him 1 instead of -1.

From these two cases we conclude that there is no separating sequential equilibrium in pure strategies.

2) Pooling equilibria:

A. $\sigma_P^A = (m_L \text{ when W}, m_L \text{ when NW})$

Derive the beliefs of the defendant by Bayes rule:

$$\begin{aligned} \mu(W|m_L) = \Pr\left(W|m_L, \sigma_1^A\right) = \frac{\Pr(m_L|W,\sigma)\Pr(W)}{\Pr(m_L|\sigma)} = \frac{1 \cdot 1/3}{1} = 1/3 \\ BR_D(m_L, \mu(W|m_L) = 1/3 > 1/4) = A \\ u_P\left(\left(\sigma_P^A; A\right); \mu(\cdot)\right) = 1 \end{aligned}$$

The beliefs with respect to a high offer are off-path. Instead of using a sequence, as we are in pure strategies, consider the following two deviations:

- In the first deviation, when offered m_H, the defendant always accepts. The plaintiff would get a utility of 2 for both types (compared to 1 if he follows the

[7] Recall that you only need one of the types having a profitable deviation to rule out a candidate equilibrium.

strategy σ_P^A). With this profitable deviation, σ_P^A cannot be part of a sequential equilibrium.

- In the second deviation, when offered m_H, the defendant always rejects. The winning plaintiff gets a utility of 3, so he has an incentive to deviate. This strategy is not sequentially rational for winners, and thus cannot be part of a sequential equilibrium.

B. $\sigma_P^B = (m_H \text{ when W}, m_H \text{ when NW})$

Deriving the beliefs of the defendant.

$$\mu(W|m_H) = \Pr\left(W|m_H, \sigma_1^B\right) = \frac{\Pr(m_G|W,\sigma)\Pr(W)}{\Pr(m_H|\sigma)} = \frac{1\cdot 1/3}{1} = 1/3$$
$$BR_D(m_H, \mu(W|m_H) = 1/3 < 1/2) = R$$
$$u_P\left(\left(\sigma_P^B; R\right); \mu(\cdot)|W\right) = 3$$
$$u_P\left(\left(\sigma_P^B; A\right); \mu(\cdot)|NW\right) = -1$$

The beliefs with respect to a low offer are off-path. Consider the following deviations:

- When offered m_L, the defendant always accepts: then the looser plaintiff will deviate to offering m_L and get $u_P((m_L; A), \mu(\cdot)|NW) = 2 > -1$. This is not sequentially rational, and hence does not constitute a sequential equilibrium.
- When offered m_L, the defendant will always reject: neither type of plaintiff has a profitable deviation. We need to find off-equilibrium path beliefs that sustain rejection as a BR (recall that $\mu(W|m_L) < 1/4$). Such beliefs would sustain the unique pure sequential equilibrium.

  Let's look for a sequence, based on the plaintiffs beliefs, that maintains $\mu(W|m_L) < 1/4$ at the limit:

$$\sigma^n = \left\{\sigma_P^n(m_L|W) = \varepsilon^{2n}, \sigma_P^n(m_L|NW) = \varepsilon^n\right\} \qquad \lim_{n\to\infty}\varepsilon^n = 0$$

  The induced beliefs μ^n must be calculated by Bayesian updating according to σ^n. At the limit, we want that $\mu^n \to \mu$.

$$\begin{aligned}
\mu^n(W|m_L) &= \frac{\Pr(m_L|W,\sigma)\Pr(W|\sigma)}{\Pr(m_L,\sigma)} = \frac{\Pr(m_L|W,\sigma)\Pr(W|\sigma)}{\Pr(m_L|W,\sigma)\Pr(W) + \Pr(m_L|NW,\sigma)\Pr(NW)} \\
&= \frac{\varepsilon^{2n}\cdot 1/3}{\varepsilon^{2n}\cdot 1/3 + \varepsilon^n \cdot 2/3} = \frac{\varepsilon^n}{\varepsilon + 2} \underset{\lim}{=} 0 < 1/4 \\
\mu^n(W|m_H) &= \frac{\Pr(m_H|W,\sigma)\Pr(W|\sigma)}{\Pr(m_H,\sigma)} = \frac{\Pr(m_H|W,\sigma)\Pr(W|\sigma)}{\Pr(m_H|W,\sigma)\Pr(W) + \Pr(m_H|NW,\sigma)\Pr(NW)} \\
&= \frac{1-\varepsilon^{2n}}{1-\varepsilon^{2n} + 2(1-\varepsilon^n)} = \frac{1}{3}
\end{aligned}$$

  This sequence converges to some beliefs that conform with the ones needed to sustain our sequential equilibrium.

There is a unique pure strategy pooling sequential equilibrim:

$$\{\sigma = (m_H|W, m_H|NW\colon R|m_H, R|m_L); \mu(W|m_L) < 1/4, \mu(W|m_H) = 1/3\}$$

Mixed strategies:

First, we will show that unless the defendant plays $\{R, R\}$, there is no way to make both types indifferent at the same time:

- Indifference condition for plaintiff of type W:

$$\sigma_D(A|m_L) + 3(1 - \sigma_D(A|m_L)) = 2\sigma_D(A|m_H) + 3(1 - \sigma_D(A|m_H))$$
$$2\sigma_D(A|m_L) = \sigma_D(A|m_H)$$

- Indifference condition for plaintiff of type NW:

$$\sigma_D(A|m_L) - (1 - \sigma_D(A|m_L)) = 2\sigma_D(A|m_H) - (1 - \sigma_D(A|m_H))$$
$$\frac{2}{3}\sigma_D(A|m_L) = \sigma_D(A|m_H)$$

The two indifference conditions cannot be satisfied at the same time. We thus proceed to analyse each case separately.

- Defendant makes type W plaintiff indifferent. This means that $2\sigma_D(A|m_L) = \sigma_D(A|m_H)$. For the NW type, it means that he prefers to play m_H.

  For the defendant, it implies that $\mu(W|m_L) = 1$. But at the beginning of the exercise we found that $BR_D(m_L, \mu(W|m_L) = 1 > 1/4) = A$. So the defendant is not fully mixing, a contradiction. This is not a WPBE.
- Defendant makes type NW plaintiff indifferent by mixing $\frac{2}{3}\sigma_D(A|m_L) = \sigma_D(A|m_H)$. For the W type, it means that he finds it optimal to offer m_H.

  The defendant then expects that $\mu(W|m_L) = 0$. With these beliefs, $BR_D(m_L, \mu(W|m_L) = 0 < 1/4) = R$. So player 2 is not fully mixing, a contradiction. This is not WPBE.

Hence, we restrict our attention to the case where $\{R|m_H, R|m_L\}$ is played by the defendant: $\sigma_D(A|m_L) = \sigma_D(A|m_H) = 0$. By the defendant's BRs, this means that the beliefs have to be such that $\mu(W|m_H) < 1/2$ and $\mu(W|m_L) < 1/4$. Our mixed strategy must induce by Bayes these beliefs:

$$\mu(W|m_H) = \frac{\Pr(m_H|W, \sigma)\Pr(W)}{\Pr(m_H|\sigma)} = \frac{\sigma_P(m_H|W) \cdot 1/3}{1/3\sigma_P(m_H|W) + 2/3\sigma_P(m_H|NW)} < \frac{1}{2}$$
$$\sigma_P(m_H|W) < 2\sigma_P(m_H|NW)$$
$$\mu(W|m_L) = \frac{\Pr(m_L|W, \sigma)\Pr(W)}{\Pr(m_L|\sigma)} = \frac{\sigma_1(m_L|W)1/3}{1/3\sigma_P(m_L|W) + 2/3\sigma_P(m_L|NW)} < \frac{1}{4}$$
$$\sigma_P(m_H|W) > \frac{1}{3} + \frac{2}{3}\sigma_P(m_H|NW)$$

The mixed strategy sequential equilibrium is:

$$\left\{\sigma = \left(\frac{2}{3}\sigma_P(m_H|NW) + \frac{1}{3} < \sigma_P(m_H|W) < 2\sigma_P(m_H|NW); (R|m_H, R|m_L)\right);\right.$$
$$\left.\mu(W|m_L) < \frac{1}{4}, \mu(W|m_H) < \frac{1}{2}\right\}$$

Exercise 3: Education and Employment

There is a single firm and a continuum of workers. The firm moves first and sets a wage schedule. The workers move second and choose whether to apply for the job and, if so, how much education to get. There are two levels of education: none and some. Hence the firm in the first stage offers non-negative wages $\bar{w}$ for the educated and $\underline{w}$ for the uneducated. You may wish to think of this as a mechanism design problem where the firm designs the mechanism by choosing the wages. The worker has two types, good workers and bad workers: $\underline{\theta}, \bar{\theta}$ where $0 < \underline{\theta} < \bar{\theta}$ and the proportion of good workers is p. Education costs $\bar{c}$ for good workers and $\underline{c}$ for bad workers, where $0 < \bar{c} < \underline{c}$, so that it is cheaper for good workers to get an education than it is for bad workers.

Worker utility is the difference between the wage and the cost of education, or zero if the worker does not apply to the job. The firm can only hire one individual and it will choose a candidate from the pool of potential applicants, and will choose depending on the education level. Firm utility is the difference between the productivity and wage paid. For what values of p will the firm choose a pooling equilibrium? A separating equilibrium?

Notice how this exercise is different from the first two: in this case, the uninformed player (firm) moves first and restricts the set of choices available to the informed player (worker). Hence, the informed player is setting up a screening mechanism, with which it can force the informed player to reveal its type. We will hence study the maximization problem of the firm when it implements the screening equilibrium.

The firm has two choices in pure strategies: either it sets the menu of wages such that the two worker types separate, or the menu of wages allows the candidates to choose the same level of education per type. Let us first state the candidate equilibria:

1. Separating equilibria
    A. Low type gets educated, high type does not.
    B. High type gets educated, low type does not.
2. Pooling equilibria
    C. Both types get educated.
    D. No type gets educated.

Let's study each candidate in turns. We will write down the maximization problem of the firm subject to the constraints that the two type of workers want to participate in the mechanism (PC) and comply with it (IC).

1. Separating

A. Low type gets educated, high type does not.

$$\max_{\underline{w},\bar{w}} \Pi = \max_{\underline{w},\bar{w}} \bar{\theta} - \underline{w}$$

$$PC_{\bar{\theta}}\colon \underline{w} \geq 0 \qquad IC_{\bar{\theta}}\colon \underline{w} \geq \bar{w} - \bar{c}$$
$$PC_{\underline{\theta}}\colon \bar{w} - \underline{c} \geq 0 \qquad IC_{\underline{\theta}}\colon \bar{w} - \underline{c} \geq \underline{w}$$

However, note that the two ICs imply a contradiction.

$$\overbrace{\bar{w} - \underline{c} \geq \underline{w}}^{IC_{\underline{\theta}}} \underbrace{\geq \bar{w} - \bar{c}}_{IC_{\bar{\theta}}}$$

$$\bar{w} - \underline{c} \geq \bar{w} - \bar{c} \qquad \bar{c} \geq \underline{c}$$

The question tells us that $\bar{c} < \underline{c}$. This equilibrium is not implementable as it is not incentive compatible.

B. High type gets an education, while the low type does not.

$$\max_{\underline{w},\bar{w}} \Pi = \max_{\underline{w},\bar{w}} \bar{\theta} - \bar{w}$$

$$PC_{\bar{\theta}}\colon \bar{w} - \bar{c} \geq 0 \qquad IC_{\bar{\theta}}\colon \bar{w} - \bar{c} \geq \underline{w}$$
$$PC_{\underline{\theta}}\colon \underline{w} \geq 0 \qquad IC_{\underline{\theta}}\colon \underline{w} \geq \bar{w} - \underline{c}$$

Notice that in this case, we have four equations (2PCs, 2ICs) and two unknowns ($\underline{w}, \bar{w}$). We can solve for the equilibrium level of wages by inspecting the constraints, without the need to explicitly maximize. [8] If we put together the two incentive constraints, we get the following relation:

$$\overbrace{\bar{w} - \bar{c} \geq \underline{w}}^{IC_{\bar{\theta}}} \underbrace{\geq \bar{w} - \underline{c}}_{IC_{\underline{\theta}}}$$

Notice how $\underline{w}$ only appears in the $PC_{\underline{\theta}}$ and firms enter the profit of the firm negatively. Hence, the firm can push $\underline{w}$ to its lower bound until $\underline{w} = 0$ and the two ICs would still be satisfied. Furthermore, once we plug in the value of $\underline{w} = 0$, out of the two ICs, the one that is binding the fastest is $IC_{\bar{\theta}}$. Hence, from $IC_{\bar{\theta}}$, we find that $\bar{w} = \bar{c}$.

With these equilibrium level of wages, the profits of the firm from separating are:

$$\Pi^S = \bar{\theta} - \bar{w} = \bar{\theta} - \bar{c}$$

2. Pooling

[8] If you prefer to maximize with the constraints, you will find that some are binding (Lagrange multiplier non-negative) and some others are not. In this solutions key, I skip this formal derivation of binding constraints and directly argue why some are binding.

C. Both types of workers get an education.

Since the firm can no longer distinguish the high ability worker, its profits will be expressed in terms of expectations.

$$\max_{\underline{w},\bar{w}} \mathbb{E}[\Pi] = \max_{\underline{w},\bar{w}} p\left(\bar{\theta} - \bar{w}\right) + (1-p)\left(\underline{\theta} - \bar{w}\right)$$

$$PC_{\bar{\theta}}\colon \bar{w} - \bar{c} \geq 0 \qquad\qquad IC_{\bar{\theta}}\colon \bar{w} - \bar{c} \geq \underline{w}$$

$$PC_{\underline{\theta}}\colon \bar{w} - \underline{c} \geq 0 \qquad\qquad IC_{\underline{\theta}}\colon \bar{w} - \underline{c} \geq \underline{w}$$

Once again, let's look at each constraint in turns. We can see that, since the cost of education for the low ability worker is higher, $PC_{\underline{\theta}}$ will bind faster than $PC_{\bar{\theta}}$. Hence, $\bar{w} = \underline{c}$. If we plug these values into the ICs, we will find that:

$$IC_{\bar{\theta}}\colon \bar{w} - \bar{c} = \underline{c} - \bar{c} \geq \underline{w}$$

$$IC_{\underline{\theta}}\colon \bar{w} - \underline{c} = \underline{c} - \underline{c} \geq \underline{w}$$

Hence, $IC_{\underline{\theta}}$ binds faster, which means that the firm will set the lowest possible $\underline{w}$ at $\underline{w} = 0$. These wages yield a profit for the firm of:

$$\Pi^C = p\left(\bar{\theta} - \underline{c}\right) + (1-p)\left(\underline{\theta} - \underline{c}\right) = p\bar{\theta} + (1-p)\underline{\theta} - \underline{c}$$

D. No worker gets an education

$$\max_{\underline{w},\bar{w}} \mathbb{E}[\Pi] = \max_{\underline{w},\bar{w}} p\left(\bar{\theta} - \underline{w}\right) + (1-p)\left(\underline{\theta} - \underline{w}\right)$$

$$PC_{\bar{\theta}}\colon \underline{w} \geq 0 \qquad\qquad IC_{\bar{\theta}}\colon \underline{w} \geq \bar{w} - \bar{c}$$

$$PC_{\underline{\theta}}\colon \underline{w} \geq 0 \qquad\qquad IC_{\underline{\theta}}\colon \underline{w} \geq \bar{w} - \underline{c}$$

We can set the PCs binding so that $\underline{w} = 0$. Plugging this value in the ICs we find $IC_{\bar{\theta}}$ is binding faster:

$$IC_{\bar{\theta}}\colon \underline{w} = 0 \geq \bar{w} - \bar{c}$$

$$IC_{\underline{\theta}}\colon \underline{w} = 0 \geq \bar{w} - \underline{c}$$

This implies that $\bar{w} = \bar{c}$. The expected profits of the firm are:

$$\Pi^D = p\left(\bar{\theta} - 0\right) + (1-p)\left(\underline{\theta} - 0\right) = p\bar{\theta} + (1-p)\underline{\theta}$$

If we compare Π^C and Π^D, the firm makes higher profits when neither type gets an education. Hence, the relevant pooling equilibrium that we will compare to the separating equilibrium is candidate D, so that:

$$\Pi^P = p\bar{\theta} + (1-p)\underline{\theta}$$

The final step is to compare the profits of pooling and separating and asses for which values of p the firm prefers to implement one or the other. We will look for the threshold value $\hat{p}$.

$$\begin{aligned}
\Pi^P &\gtrless \Pi^S \\
p\bar{\theta} + (1-p)\underline{\theta} &\gtrless \bar{\theta} - \bar{c} \\
\hat{p}\bar{\theta} + (1-\hat{p})\underline{\theta} &= \bar{\theta} - \bar{c} \\
\hat{p} &= 1 - \frac{\bar{c}}{\bar{\theta} - \underline{\theta}}
\end{aligned}$$

Therefore, when the fraction of high type individuals p is higher than $\hat{p}$, the probability that a high type individual is selected at random is high and the firm prefers to implement a pooling equilibrium. On the other hand, when p is lower than $\hat{p}$, the firm would rather separate the types and get a high ability candidate for sure.

Side comment: do we need restrictions on the values of $\frac{\bar{c}}{\bar{\theta} - \underline{\theta}}$*?*
From the parameter values, we know that $\frac{\bar{c}}{\bar{\theta} - \underline{\theta}} > 0$, but we know nothing about whether $\frac{\bar{c}}{\bar{\theta} - \underline{\theta}} \gtrless 1$. However, given that we are looking for a threshold value, this is not essential. It could be the case that $\frac{\bar{c}}{\bar{\theta} - \underline{\theta}} > 1$, so that $\hat{p} < 0$, which would mean that $p > \hat{p}$ and hence the firm will implement a pooling equilibrium.

Exercise 4: Decreasing Absolute Risk Aversion

A continuum of consumers has utility function $u(x) = 78x - x^2$. Each consumer has a 50% chance of getting $x = 30$ and a 50% chance of $x = 10$. Consider the following "mechanism": a consumer that announces he has $x = 30$ pays τ. A consumer that announces $x = 10$ receives a lottery with a 50% chance of winning g and a 50% chance of winning b where $0.5g + 0.5b = \tau$. Suppose that "rich" consumers ($x = 30$) can lie and say that they are poor ($x = 10$). Find the mechanism that maximizes the expected utility of a consumer before he knows his type, subject to the constraint that the rich consumer does not wish to lie.

Let's approach this question from a mechanism design approach. We want to maximize the expected utility of the agent, subject to the incentive compatibility constraint and $0.5g + 0.5b = \tau$.

$$\begin{aligned}
\max \mathbb{E}[u(\cdot)] &= \max_{\tau,b,g} \frac{1}{2}u(30-\tau) + \frac{1}{2}\left[\frac{1}{2}u(10+b) + \frac{1}{2}u(10+g)\right] \text{ st} \\
u(30-\tau) &\geq \frac{1}{2}u(30+b) + \frac{1}{2}u(30+g) \\
\tau &= \frac{g}{2} + \frac{b}{2}
\end{aligned}$$

Notice that we have two constraints for three unknowns: τ, b, g. From inspecting the IC, we can conclude that it must be binding: otherwise we could increase τ, which would decrease the utility of the rich consumer while satisfying the constraint, and would also increase the transfer sent to the poor consumer. [9] I will use the two constraints to derive expressions of b, g depending on τ, so that I can later perform an unconstrained maximization.

$$\left.\begin{aligned} b &= 2\tau - g \\ 2u(30-\tau) &= u(30+b) + u(30+g) \end{aligned}\right\} 2u(30-\tau) = u(30+2\tau-g) + u(30+g)$$

$$\begin{aligned} 2u(30-\tau) &= u(30+2\tau-g) + u(30+g) \\ 2\Big[78(30-\tau) - (30-\tau)^2\Big] &= 78(30+g) - (30+g)^2 + 78(30+2\tau-g) - (30+2\tau-g)^2 \\ -2g^2 + 4\tau g - 2\tau^2 + 72\tau &= 0 \\ g &= \tau \pm 6\sqrt{\tau} \end{aligned}$$

Let me focus on the case where $g = \tau + 6\sqrt{\tau}$, which implies that $b = 2\tau - g = \tau - 6\sqrt{\tau}$. This means that, without loss of generality, we can ignore the case where $g = \tau - 6\sqrt{\tau}$, as it implies that $b = \tau + 6\sqrt{\tau}$ and it is analogue to the analysis that will follow. We now plug in these two expressions of b, g into the objective function and maximize:

$$\max_{\tau} \frac{1}{2}u(30-\tau) + \frac{1}{2}\left[\frac{1}{2}u\left(10+\tau-6\sqrt{\tau}\right) + \frac{1}{2}u\left(10+\tau+6\sqrt{\tau}\right)\right]$$

$$\frac{\partial}{\partial\tau}: \frac{1}{2}\left[(-1)u'(30-\tau) + \frac{1}{2}\left(1-\frac{6}{2}\tau^{-1/2}\right)u'\left(10+\tau-6\sqrt{\tau}\right) + \frac{1}{2}\left(1+\frac{6}{2}\tau^{-1/2}\right)u'\left(10+\tau+6\sqrt{\tau}\right)\right] = 0$$

$$0 = -78 + 2(30-\tau) + \left(1-3\tau^{-1/2}\right)\left[39-(10+\tau-6\sqrt{\tau})\right] + \left(1+3\tau^{-1/2}\right)\left[39-(10+\tau+6\sqrt{\tau})\right]$$

$$4\tau = 4$$

Therefore, the values that characterize the lotteries should be:

$$\begin{aligned} \tau &= 1 \\ g &= \tau + 6\sqrt{\tau} = 1 + 6\sqrt{1} = 7 \\ b &= \tau - 6\sqrt{\tau} = 1 - 6\sqrt{1} = -5 \end{aligned}$$

We can check that these values make the lotteries incentive compatible for the rich consumer, before he knows his type:

$$\begin{aligned} u(30-\tau) &= u(29) = 78 * 29 - 29^2 = 1421 \\ u(30+b) &= u(25) = 78 * 25 - 25^2 = 1325 \\ u(30+g) &= u(37) = 78 * 37 - 37^2 = 1517 \end{aligned}$$

Which means that the IC is satisfied and binding, as $(1325 + 1517)/2 = 1421$.

Exercise 5: Moral hazard

There are 2 states of the world $s = 1, 2$ and 2 possible actions $a = 1, 2$. A risk neutral principal observes only the state and not the action of the agent he hires. The net gain of an agent if he is

[9] Instead of arguing with words, we can also do a constrained maximization using the Lagrangian. If you try this out, you will find that the Lagrange multiplier associated to the constraint is equal to $\lambda = 1/2$.

paid w and takes action a is $v(w) - c(a)$, where $c(1) < c(2)$. Under action a the probability of state s is $p_s(a)$, where $p_2(a)$ is increasing in a. The agent's reservation utility is 0. The output (received by the principal) is y_s where $y_2 > y_1$. [10]

I will use the following (simplified) notation. Call the state $s = 1, 2$ so that $c(s) \equiv c_s$ and $p_s(a) \equiv p(s|a)$. Denote reservation utility by $\bar{u}$. The wage schedule is denoted by $w = (w_1, w_2)$.

- *The principal wishes to induce action $a = 2$ and only "downward" constraints of pretending lower cost are potentially binding. What condition is sufficient for the optimal incentive scheme w_1, w_2 to be monotonic? Prove your claim.*

  Monotonicity of w means that $w_2 > w_1$. We will look for a condition that will guarantee this relation.

  Let us write the maximization problem of the principal. She will maximize her expected profits (expected in terms of state of the world), inducing the agent to exert high effort. As we want the agent to exert high effort, we will include the downward incentive constraint: a contract is downward incentive compatible if it makes the agent better off exerting high effort.

$$\begin{aligned}
&\max_{w_1, w_2} p(1|2)(y_1 - w_1) + p(2|2)(y_2 - w_2) \qquad \text{st} \\
PC_{a=2}&: p(1|2)[v(w_1) - c_2] + p(2|2)[v(w_2) - c_2] \geq \bar{u} = 0 \\
IC_{a=2 \succcurlyeq a=1}&: p(1|2)[v(w_1) - c_2] + p(2|2)[v(w_2) - c_2] \geq p(1|1)[v(w_1) - c_1] + p(2|1)[v(w_2) - c_1]
\end{aligned}$$

  Rearranging the IC, we can see that:

$$[p(2|2) - p(1|2)][v(w_2) - v(w_1)] \geq c_2 - c_1 > 0$$

  If $v(w_s)$ is increasing, this implies $w_2 > w_1$. Therefore, I claim that $v'(w_s) > 0$ is sufficient for the optimal incentive scheme to be monotonic. Let me now formally prove this claim using the Lagrangian.

$$\begin{aligned}
\mathcal{L} &= p(1|2)(y_1 - w_1) + p(2|2)(y_2 - w_2) + \mu[p(1|2)[v(w_1) - c_2] + p(2|2)[v(w_2) - c_2]] \\
&\quad + \lambda[p(1|2)[v(w_1) - c_2] + p(2|2)[v(w_2) - c_2] - p(1|1)[v(w_1) - c_1] - p(2|1)[v(w_2) - c_1]] \\
\frac{\partial \mathcal{L}}{\partial w_s}&: \; -p(s|2) + \mu[p(s|2)v'(w_s)] + \lambda v'(w_s)[p(s|2) - p(s|1)] = 0 \qquad \forall s = 1, 2 \\
\min &\left\{ \sum_{s=1}^{2} p(s|2)[v(w_s) - c_2], \mu \right\} \geq 0 \\
\min &\left\{ \sum_{s=1}^{2} p(s|2)[v(w_s) - c_2] - \sum_{s=1}^{2} p(s|1)[v(w_s) - c_1], \lambda \right\} \geq 0
\end{aligned}$$

[10] This exercise is an introduction to moral hazard, a topic that you will cover in more depth in the last block of the microeconomics sequence. For this exercise, it is enough to know that moral hazard arises when, after signing a contract, the agent can modify his behaviour in a way that hurts the principal's interests. When solving for the optimal mechanism, we assume that the principal designs the contract in a way to induce (or force upon) her preferred level of effort on the agent and eliminate this moral hazard.

The FOCs will give us the standard moral hazard formula, which I will also use to argue why both constraints need to be binding. [11]

$$\frac{1}{v'(w_s)} = \mu + \lambda \left(1 - \frac{p(s|1)}{p(s|2)}\right) \qquad \forall s = 1, 2$$

- Suppose $\mu = 0$, so that PC is slack. This means that our formula becomes:

  $$\frac{1}{v'(w_s)} = \lambda \left(1 - \frac{p(s|1)}{p(s|2)}\right)$$

  Let me focus on the case when $s = 1$.

  $$\frac{1}{v'(w_1)} = \lambda \left(1 - \frac{p(1|1)}{p(1|2)}\right) = \lambda \left(1 - \frac{1 - p(2|1)}{1 - p(2|2)}\right)$$

  As we know, $p(2|a)$ is increasing in a: $p(2|2) > p(2|1)$, so that $1 - p(s|1) < 1 - p(2|2)$. This would mean that $1/v'(w_1) < 0$ and $v'(w_1) < 0$. But this is not consistent with a monotonic wage schedule, as it would imply that the agent would derive less utility from a wage higher than w_1. Therefore, we have a contradiction, as it must be that $v'(w_s) > 0$, $\mu > 0$ and so PC binds.

- Suppose $\lambda = 0$, so that IC is slack. This means our formula becomes:

  $$\frac{1}{v'(w_1)} = \mu \qquad \frac{1}{v'(w_2)} = \mu$$

  But we know that $v'(w_s) > 0 \ \forall s = 1, 2$. Therefore, we have that $w_1 = w_2 = \bar{w}$ and so $v(\bar{w}) = \bar{v}$. If we plug this in the IC, we will find that:

  $$\begin{aligned} p(1|2)[\bar{v} - c_2] + p(2|2)[\bar{v} - c_2] > p(1|1)[\bar{v} - c_2] + p(2|2)[\bar{v} - c_1] \\ \bar{v} - c_2 > \bar{v} - c_1 \\ c_1 > c_2 \end{aligned}$$

  This is a contradiction. Hence, $\lambda > 0$ and IC binding.

Therefore, a sufficient condition for the wage schedule to be monotonic is that $v'(w) > 0$.

Furthermore, note that:

$$\frac{1}{v'(v_2)} = \mu + \lambda \left(1 - \frac{p(2|1)}{p(2|2)}\right) > \mu + \lambda \left(1 - \frac{p(1|1)}{p(1|2)}\right) = \frac{1}{v'(v_1)}$$
$$\frac{p(2|1)}{p(2|2)} < \frac{p(1|1)}{p(1|2)}$$

This last condition is called the monotone likelihood ratio, and it is only achieved if we also assume that $v'(w)$ is decreasing on w: $v''(w) < 0$.

Note that assuming $v'(w) > 0$ and $v''(w) < 0$ is quite standard.

[11] If you want to read more about this formula and how we argue through each constraint, I recommend either MWG 14B or the book you will use in the next block for microeconomics: Reny P. and Jehle, *Advanced Microeconomic Theory*, 2nd edition, Addisen Wesley.

- *Suppose $v(w) = 1 - \exp^{-\gamma w}$. What more can be said about $w_2 - w_1$?*

  I will use the utility function to explicitly calculate the difference between the wages. The conditions we have imposed imply that $\gamma > 0$ Notice that we can express this difference as:

$$w_2 - w_1 = -\frac{1}{\gamma} \ln \left(\frac{\exp^{-\gamma w_2}}{\exp^{-\gamma w_1}} \right) > 0$$

  We have two binding constraints (we proved this in part *a)*) and two unknowns, so I will focus on the constraints.

$$\begin{aligned}
PC\colon & [1 - p(2|2)][v(w_1) - c_2] + p(2|2)[v(w_2) - c_2] = 0 \\
& p(2|2)\left[1 - \exp^{-\gamma w_2}\right] + [1 - p(2|2)]\left[1 - \exp^{-\gamma w_1}\right] - c_2 = 0 \\
& 1 + p(2|2)\left[\exp^{-\gamma w_1} - \exp^{-\gamma w_2}\right] - \exp^{-\gamma w_1} - c_2 = 0 \\
IC\colon & [1 - p(2|2)][v(w_1) - c_2] + p(2|2)[v(w_2) - c_2] = [1 - p(2|1)][v(w_1) - c_1] + p(2|1)[v(w_2) - c_1] \\
& p(2|2)\left[1 - \exp^{-\gamma w_2}\right] + [1 - p(2|2)]\left[1 - \exp^{-\gamma w_1}\right] - c_2 = \\
& p(2|1)\left[1 - \exp^{-\gamma w_2}\right] + [1 - p(2|1)]\left[1 - \exp^{-\gamma w_1}\right] - c_1 \\
& [p(2|2) - p(2|1)]\left[\exp^{-\gamma w_1} - \exp^{-\gamma w_2}\right] = c_2 - c_1
\end{aligned}$$

  I solve for $[\exp^{-\gamma w_1} - \exp^{-\gamma w_2}]$ using the IC and I plug it in the PC.

$$\begin{aligned}
IC\colon\ \exp^{-\gamma w_1} - \exp^{-\gamma w_2} &= \frac{c_2 - c_1}{p(2|2) - p(2|1)} \\
PC\colon \qquad \exp^{-\gamma w_1} &= 1 + p(2|2)\left[\exp^{-\gamma w_1} - \exp^{-\gamma w_2}\right] - c_2 \\
&= 1 + p(2|2)\left[\frac{c_2 - c_1}{p(2|2) - p(2|1)}\right] - c_2 \qquad (1) \\
&= \frac{p(2|2) - p(2|1) - p(2|2)c_1 + p(2|1)c_2}{p(2|2) - p(2|1)}
\end{aligned}$$

  I then use this $\exp^{-\gamma w_1}$ to plug it back into the IC:

$$\begin{aligned}
IC\colon\ \exp^{-\gamma w_2} &= \exp^{-\gamma w_1} - \frac{c_2 - c_1}{p(2|2) - p(2|1)} \\
&= \frac{p(2|2) - p(2|1) - p(2|2)c_1 + p(2|1)c_2}{p(2|2) - p(2|1)} - \frac{c_2 - c_1}{p(2|2) - p(2|1)} \\
&= \frac{c_1 - c_2 + p(2|2) - p(2|1) - p(2|2)c_1 + p(2|1)c_2}{p(2|2) - p(2|1)}
\end{aligned}$$

  We can now calculate the difference in wages as:

$$\begin{aligned}
w_2 - w_1 &= -\frac{1}{\gamma} \ln \left(\frac{\exp^{-\gamma w_2}}{\exp^{-\gamma w_1}} \right) \\
&= -\frac{1}{\gamma} \ln \left(\frac{c_1 - c_2 + p(2|2) - p(2|1) - p(2|2)c_1 + p(2|1)c_2}{p(2|2) - p(2|1) - p(2|2)c_1 + p(2|1)c_2} \right) \\
&= -\frac{1}{\gamma} \ln \left(1 - \frac{c_2 - c_1}{p(2|2) - p(2|1) - p(2|2)c_1 + p(2|1)c_2} \right)
\end{aligned}$$

  Note that this expression is always positive, as $\dfrac{\exp^{-\gamma w_2}}{\exp^{-\gamma w_1}} < 1$.

- *For this special case, discuss the effect on the optimal incentive scheme if there is a change in the agent's reservation utility, assuming the principal still wants to induce the action $a = 2$.*

  We need to understand how a change in $\bar{u}$ affects both w_1 and w_2. To do this, we can take the PC expression in equation (1), considering that now $\bar{u} \neq 0$, and apply the Implicit Function Theorem (IFT):

$$PC(c_1, c_2, \bar{u}, p(2|\cdot))\colon 1 - \exp^{-\gamma w_1} - c_2 + p(2|2)\frac{c_2 - c_1}{p(2|2) - p(2|1)} - \bar{u} = 0$$
$$IFT\colon \frac{dw_1}{d\bar{u}} = -\frac{\partial PC/\partial \bar{u}}{\partial PC/\partial w_1} = -\frac{-1}{(-1)(-\gamma)\exp^{-\gamma w_1}} > 0$$

  So that w_1 increases with $\bar{u}$. I will then check how w_2 changes with w_1.

$$IC(c_1, c_2, p(2|\cdot))\colon [(p(2|2) - p(2|1)] \left[\exp^{-\gamma w_1} - \exp^{-\gamma w_2}\right] - c_2 + c_1 = 0$$
$$IFT\colon \frac{dw_2}{dw_1} = -\frac{\partial IC/\partial w_1}{\partial IC/\partial w_2} = -\frac{(-\gamma)\exp^{-\gamma w_1}}{(-1)(-\gamma)\exp^{-\gamma w_2}} > 1$$

  The IFT shows that not only does w_2 increase with w_1, but that it does so more than proportionately.

  To conclude: both wages increase with the outside option, and the discrepancy between wages increases with the reservation utility.

- *Is there a change in the agents reservation utility that would lead the principal to prefer to induce the action $a = 1$?*

  It can be the case that, if $\bar{u}$ is very high, the agent is better off inducing low effort (the discrepancy between wages as derived in the last question is too large). Call $\mathbb{E}[w^\star|a = 2]$ the expected value of the optimal wage scheme for high effort under the state distribution from high effort $\Big(p(1|2), p(2|2)\Big)$.

  Notice that, when inducing low effort, there is no need to consider the upward IC. A risk neutral principal will impose the lowest acceptable flat wage: $w_1 = w_2 = \bar{w}$ such that:

$$PC_{a=1}\colon v(\bar{w}) = \bar{u} + c_1$$

  Therefore, we can write the profits of the firm from inducing low and high effort as:

$$\pi_1 = \mathbb{E}[y|a = 1] - \bar{w}$$
$$\pi_2 = \mathbb{E}[y|a = 2] - \mathbb{E}[w^\star|a = 2]$$

  If inducing $a = 1$ is optimal, it must be the case that:

$$\pi_1 > \pi_2$$
$$\mathbb{E}[y|a = 1] - \bar{w} > \mathbb{E}[y|a = 2] - \mathbb{E}[w^\star|a = 2]$$
$$\underbrace{\mathbb{E}[y|a = 1] - \mathbb{E}[y|a = 2]}_{\text{constant in } \bar{u}} > \underbrace{\bar{w} - \mathbb{E}[w^\star|a = 2]}_{\text{must be decreasing in } \bar{u}}$$

  Let's look at each wage schedule in turns:

$$IFT\colon = \frac{d\bar{w}}{d\bar{u}} = -\frac{\partial PC_{a=1}/\partial \bar{u}}{\partial PC_{a=1}/\partial \bar{w}} = -\frac{-1}{v'(\bar{w})} = \frac{1}{v'(\bar{w})}$$

Notice that this is exactly the same condition for the low wage in the optimal wage schedule: $\frac{d\bar{w}}{d\bar{u}} = \frac{dw_1^\star}{d\bar{u}}$. The low wage in the optimal contract for high effort increases the same way as the fixed low effort wage with respect to $\bar{u}$. Furthermore, we have shown in the previous point that $w_2^\star$ increases more than proportionally to $w_1^\star$: $\frac{d\bar{w}}{d\bar{u}} < \frac{dw_2^\star}{d\bar{u}}$. Therefore, we have that:

$$\frac{d\,\mathbb{E}[w^\star|a=2]}{d\bar{u}} = \mathbb{E}\left[\frac{dw^\star}{d\bar{u}}\middle| a=2\right] > \frac{d\bar{w}}{d\bar{u}}$$

This means that $\bar{w} - \mathbb{E}[w^\star|a=2]$ is strictly decreasing in $\bar{u}$. For high enough $\bar{u}$, we have that $\pi_1 > \pi_2$ and thus it is optimal to induce low effort.

- *Suppose that the agent's utility is $u(w,a)$ and is not separable. Is it possible to induce the agent to use $a=2$ for arbitrarily large reservation utilities?*

  The implementability result relies on the utility of wages and disutility of effort being separable.

Exercise 6: Adverse selection

Consider a continuum of ex ante identical individuals with utility function for consumption c of $-\exp^{-c}$. Ex post, two states are possible. In state 1 the endowment is 2. In state 2 the endowment is 0. What is the first best allocation? Suppose that the state is privately known. Show that there is no incentive compatible ex ante exclusive contract that gives the low endowment type more utility than at autarky. [12]

First, let's normalize the continuum of consumers to be between zero and one. This way, using the law of large numbers (LLN), we can interpret the proportion p of individuals who have high endowment ($w=2$) as the probability that each agent will be in the high state; and equally the proportion $(1-p)$ of individuals who have low endowment ($w=0$) as the probability of being in a low state. Furthermore, notice that the LLN allows us to consider p as a constant.

Second, let's look at the agent's preferences.

$$\begin{aligned} u(c) &= -\exp^{-c} \\ u'(c) &= \exp^{-c} \\ u''(c) &= -\exp^{-c} < 0 \end{aligned}$$

The agent is risk averse, which already hints towards the first best allocation being full insurance. To see this, let us focus on the maximization problem of a social planner. From the properties of the utility function, we can assume that there is a representative consumer so that:

$$\max_c -\exp^{-c} \qquad \text{st}$$

$$c \leq 2p + 0(1-p)$$

[12] This exercise is an introduction to adverse selection, covered in the last block of the microeconomics sequence, and incomplete markets, a topic that you will cover in more detail in the last block of the macro sequence.

The budget constraint is binding, so that the solution is $c^{FB} = 2p$.

What does this FB allocation have to do with full insurance? In the FB, the state of the world that each agent is in is verifiable. Hence, the agents that receive the high endowment agree to "share" their consumption with the low endowment agents: they do not consume the full $w = 2$ but only a portion $2p$. The agreement implies that, when they are hit by a bad shock, those in the god state will also share their endowment, allowing them to consume something instead of $w = c = 0$. Agents are smoothing consumption over the shocks and markets are complete: we can write a contract stating this mechanism, and it will be enforced because the state is publicly verifiable.

This mechanism fails if we make the the state privately known: it is not incentive compatible. The ones who receive a high endowment have an incentive to claim that they received a low endowment.

Let us write the mechanism to verify this claim. As the state is private information, we will introduce incentive constraints, an ex-ante participation constraint and a feasibility constraint. I will denote τ_h the transfer to the agents who claim $w = 2$ and τ_l the transfer to the agents who claim $w = 0$.

$$\max_c - \exp^{-c} \qquad \text{st}$$

$$PC_{\text{ex ante}}\colon p(2 + \tau_h) + (1 - p)(0 + \tau_l) \geq 2p \qquad IC_{w=0}\colon 0 + \tau_l \geq 0 + \tau_h$$
$$\text{Feasibility}\colon p\tau_h + (1 - p)\tau_l \leq 0 \qquad IC_{w=2}\colon 2 + \tau_h \geq 2 + \tau_l$$

The IC constraints imply that $\tau_h = \tau_l = \tau$. Plugging this in the PC we see that $\tau \geq 0$ and from the feasibility condition $\tau \leq 0$. The only solution is $\tau = 0$, which means that the two types of agents do not insure each other. In other words, the two types of agents do not "trade risk" and hence we have autarky.