Chapter 3. Games of Incomplete Information

In this chapter, we will study games of *incomplete information*, also called *Bayesian games*. In such games, at least one player is uncertain about another player's payoff function.

3.A. Static Games of Incomplete Information

In this section, we will study the simultaneous-move game of incomplete information, also called the *static Bayesian game*. We will first look at the incomplete information Cournot duopoly model in Section 3.A.1. Then we develop the normal-form representation of the general static Bayesian game and the corresponding solution concept *Bayesian Nash Equilibrium* in Section 3.A.2. Lastly, we will study two applications on auctions in Sections 3.A.3 and 3.A.4.

3.A.1. Cournot Competition under Asymmetric Information

Game Setup. Consider the Cournot duopoly model where the two firms choose their quantities simultaneously. Let q_1 and q_2 denote the quantities (of a homogeneous product) produced by firms 1 and 2, respectively. Let $P(Q) = a - Q$ be the market-clearing price when the aggregate quantity on the market is $Q = q_1 + q_2$. Firm 1's cost function is $C_1(q_1) = cq_1$. Firm 2's cost function is

- $C_2(q_2) = c_H q_2$ with probability θ, and
- $C_2(q_2) = c_L q_2$ with probability $1 - \theta$,

where $c_L < c_H$. Furthermore, information is asymmetric:

- Firm 2 knows both its own cost function (i.e, the realization of the marginal cost c_H, c_L) and Firm 1's cost function, but
- Firm 1 knows its own cost function and only that Firm 2's marginal cost is c_H with probability θ and c_L with probability $1 - \theta$.

Analysis. Naturally, Firm 2 may choose different quantities depending on whether its marginal cost is high or low. Moreover, Firm 1 should anticipate this. Let

- $q_2^*(c_H)$ and $q_2^*(c_L)$ denote Firm 2's equilibrium quantity choice;
- q_1^* denote Firm 1's equilibrium quantity choice.

Then, $q_2^*(c_H)$ solves

$$\max_{q_2} \Big[(a - q_1^* - q_2) - c_H\Big]q_2.$$

Similarly, $q_2^*(c_L)$ solves

$$\max_{q_2} \Big[(a - q_1^* - q_2) - c_L\Big]q_2.$$

Finally, q_1^* solves

$$\begin{aligned}
&\max_{q_1} \theta\Big[(a - q_1 - q_2^*(c_H)) - c\Big]q_1 + (1-\theta)\Big[(a - q_1 - q_2^*(c_L)) - c\Big]q_1 \\
\Longrightarrow\ &\max_{q_1}\Big[a - q_1 - \Big(\theta q_2^*(c_H) + (1-\theta)q_2^*(c_L)\Big) - c\Big]q_1
\end{aligned}$$

FOCs of the three optimization problem give

$$\begin{aligned}
q_2^*(c_H) &= \frac{a - q_1^* - c_H}{2} \\
q_2^*(c_L) &= \frac{a - q_1^* - c_L}{2} \\
q_1^* &= \frac{a - \Big(\theta q_2^*(c_H) + (1-\theta)q_2^*(c_L)\Big) - c}{2}
\end{aligned}$$

The solution is (assuming that the solutions are all positive)

$$\begin{aligned}
q_2^*(c_H) &= \frac{a - 2c_H + c}{3} + \frac{1-\theta}{6}(c_H - c_L); \\
q_2^*(c_L) &= \frac{a - 2c_L + c}{3} - \frac{\theta}{6}(c_H - c_L); \\
q_1^* &= \frac{a - 2c + \theta c_H + (1-\theta)c_L}{3}.
\end{aligned}$$

Note that $q_2^*(c_H) > \frac{a-2c_H+c}{3}$ and $q_2^*(c_L) < \frac{a-2c_L+c}{3}$: Firm 2 not only tailors its quantity to its cost but also responds to the fact that Firm 1 cannot do so. For example, consider the case where Firm 2 has high marginal cost. If Firm 1 knows that and could adjust its quantity accordingly, Firm 1 would respond by choosing the quantity $\frac{a-2c+c_H}{3}$, which is higher than q_1^*. Therefore, when Firm 1 cannot tailor its quantity (and chooses the smaller quantity q_1^*), Firm 2 would produce more, i.e., $q_2^*(c_H) > \frac{a-2c_H+c}{3}$.

3.A.2. Static Bayesian Games and Bayesian Nash Equilibrium

To characterize the static Bayesian games, we want to capture the idea that

1. each player knows his/her own payoff function;
2. each player may be uncertain about the other players' payoff functions.

Harsanyi (1967) introduced *type spaces* to model the players' information on payoff-relevant parameters.

Type and Belief. Player i's payoff functions is represented by

$$u_i(a_1, ..., a_n; t_i),$$

where t_i is called Player i's *type* and belongs to T_i, the set of possible types, or *type space*. For the Cournot competition model in Section 3.A.1,

- Firm 2 has two types and its type space is $T_2 = \{c_L, c_H\}$;
- Firm 1 has only one type and its type space is $T_1 = \{c\}$.

Given this definition of *types*,

1. "Player i knows his/her own payoff function" is equivalent to "Player i knows his/her own type".
2. "Player i may be uncertain about the other players' payoff functions" is equivalent to "Player i maybe uncertain about the types of other players, $t_{-i} = \{t_1, ..., t_{i-1}, t_{i+1}, ..., t_n\}$".

We use T_{-i} to denote the set of possible types of other players. And we use $p_i(t_{-i}|t_i)$ to denote Player i's *belief* about t_{-i} when his/her own type is t_i. The belief is computed by *Bayes' rule* from the prior probability distribution $p(t)$:

$$p_i(t_{-i}|t_i) = \frac{p(t_{-i}, t_i)}{p(t_i)} = \frac{p(t_{-i}, t_i)}{\sum_{t_{-i} \in T_{-i}} p(t_{-i}, t_i)}.$$

For concrete examples of applications of Bayes' rule, see the following examples.

Example 3.A.1. Consider a two player game where both players have two types. Player 1's types are denoted as A and B; Player 2's types C and D. The prior distribution of types are given below.

		Player 2	
		Type C	Type D
Player 1	Type A	30%	40%
	Type B	10%	20%

Question 3.1. What is the posterior probability $p_1(C|A)$?

We use Bayes' rule to calculate the probability:

$$p_1(C|A) = \frac{p(C, A)}{p(A)} = \frac{p(C, A)}{p(C, A) + p(D, A)} = \frac{30\%}{30\% + 40\%} = \frac{3}{7}.$$

Example 3.A.2. A certain disease affects about 1 out of 10,000 people. There is a screening test to check whether a person has the disease. The test is quite accurate. In particular, we know that

- When a person has the disease, it gives a positive result 99% of the time.
- When a person does not have the disease, it gives a negative result 98% of the time.

Question 3.2. A random person gets tested for the disease and the result comes back positive. What is the probability that the person has the disease?

Let D be the event that the person has the disease and T be the event that the test is positive. Accordingly, D^c is the event that the person does not have the disease and T^c be the event that the test is negative.

We know that

$$\begin{aligned} p(D) &= \frac{1}{10,000}; \\ p(T|D) &= \frac{99}{100} \quad \& \quad p(T^c|D) = \frac{1}{100}; \\ p(T|D^c) &= \frac{2}{100} \quad \& \quad p(T^c|D^c) = \frac{98}{100}. \end{aligned}$$

And we want to calculate $p(D|T)$. By Bayes' rule:

$$p(D|T) = \frac{p(D,T)}{p(T)} = \frac{p(D,T)}{p(D,T) + p(D^c,T)} = \frac{p(T|D)p(D)}{p(T|D)p(D) + p(T|D^c)p(D^c)} = 0.049$$

Normal-form Representation. The normal-form representation of n-player static Bayesian game specifies

- players' action spaces $A_1, ..., A_n$;
- their type spaces $T_1, ..., T_n$;
- their beliefs $p_1(t_{-1}|t_1), ..., p_n(t_{-n}|t_n)$;
- their payoff functions $u_i(a_1, ..., a_n; t_i)$ for all i.[1]

Following Harsanyi (1967), the timing of a static Bayesian game is as follows:

1. nature draws a type vector $t = (t_1, ..., t_n)$ where t_i is drawn from the set of possible types T_i;
2. nature reveals t_i to Player i but not to any other player;
3. players simultaneously choose actions, Player i choosing $a_i \in A_i$; and then

[1] More generally, a player's payoff function could also depend on the other players' types. In this case, we write $u_i(a_1, ..., a_n; t_1, ..., t_n)$.

4. payoffs $u_i(a_1, ..., a_n; t_i)$ are received.

Remark 3.1. Note that by introducing the fictional moves by nature, the *incomplete* information game is transformed to the *imperfect* information game.

Here, Player i does not know the complete history of the game when actions are chosen in Step 3. In particular, Player i does not know what nature has revealed to the other players.

Bayesian Nash Equilibrium.

Definition 3.A.1 (Strategy)**.** In the static Bayesian game, a **strategy** for Player i is a function $s_i(t_i)$ that specifies the action $a_i \in A_i$ when the type $t_i \in T_i$ is drawn by nature.

For the Cournot competition model in Section 3.A.1,

- Firm 2's strategy is $(q_2^*(c_H), q_2^*(c_L))$;
- Firm 1's strategy is q_1^*.

Next, we define the solution concept in a static Bayesian game, called *Bayesian Nash Equilibrium*. The central idea is the same: each player's strategy must be a *best response* to the other players' strategies.

Definition 3.A.2 (Bayesian Nash Equilibrium)**.** In the static Bayesian game, the strategies $s^* = (s_1^*, ..., s_n^*)$ are a (pure strategy) **Bayesian Nash Equilibrium (BNE)** if for each player i and for each of i's type $t_i \in T_i$, $s_i^*(t_i)$ solves

$$\max_{a_i \in A_i} \sum_{t_{-i} \in T_{-i}} u_i(s_1^*(t_1), ...s_{i-1}^*(t_{i-1}), a_i, s_{i+1}^*(t_{i+1}), ..., s_n^*(t_n); t) p_i(t_{-i}|t_i).$$

3.A.3. First-Price Sealed-Bid Auction

We have learned second-price auction in Chapter 1. Recall that second-price auction is dominant strategy solvable: bidding one's own valuation is a weakly dominant strategy. Now, we will study first-price auction. Note that first-price and second-price auctions only differ in winner's payments. We will consider a simple version of first-price auction with only two bidders.

Game Setup. There is one indivisible good for sale. The valuations of two potential buyers are independently drawn from a uniform distribution with support $[0, 1]$. Denote Buyer i's valuation by v_i. The auction rule is as follows:

- Buyers bid simultaneously and each submits a bid $b_i \in [0, +\infty)$.
- The bidder with the highest bid wins the auction and pays his/her own bid.
- If the two buyers submit the same highest bid, then each of the buyers has 1/2 chance of winning the good. The payment is the highest bid (since there is a tie).

Normal-form Representation.

- Buyer i's *action* is to submit a bid b_i. The action space is $A_i \in [0, \infty)$.
- Buyer i's *type* is her valuation v_i. The type space is $T_i = [0, 1]$.
- Since the valuations are independent, Buyer i's *belief* about Buyer j's type v_j is that v_j is uniformly distributed on $[0, 1]$, given any v_i.
- Buyer i's *payoff* when submitting the bid b_i is

$$u_i = \begin{cases} 0 & \text{if } b_i < b_j \\ \frac{v_i - b_i}{2} & \text{if } b_i = b_j \\ v_i - b_i & \text{if } b_i > b_j \end{cases}$$

Bayesian Nash Equilibrium. A strategy for Buyer i is a function $b_i(v_i)$. In a Bayesian Nash Equilibrium, Buyer 1's strategy $b_1(v_1)$ is a best response to Buyer 2's strategy $b_2(v_2)$, and vice versa. Thus, $b_i(v_i)$ solves

$$\max_{b_i} (v_i - b_i)\text{Prob}\{b_i > b_j(v_j)\} + \frac{1}{2}(v_i - b_i)\text{Prob}\{b_i = b_j(v_j)\}.$$

We focus on symmetric Bayesian Nash equilibrium where the two players adopt the same strictly increasing, continuous and differentiable bidding strategy $b(\cdot)$.

Suppose Buyer j adopts $b(\cdot)$. Then

- $\text{Prob}\{b_i = b(v_j)\} = 0$ since v_j is uniformly distributed and $b(\cdot)$ is strictly increasing.
- $\text{Prob}\{b_i > b(v_j)\} = \text{Prob}\{b^{-1}(b_i) > v_j\} = b^{-1}(b_i)$ since v_j is uniformly distributed on $[0, 1]$ and $b(\cdot)$ is strictly increasing, continuous and differentiable.

So Buyer i solves

$$\max_{b_i} (v_i - b_i) b^{-1}(b_i).$$

FOC gives

$$-b^{-1}(b_i) + (v_i - b_i)\frac{db^{-1}(b_i)}{db_i} = 0.$$

For $b(\cdot)$ to be a symmetric Bayesian Nash equilibrium, the solution to the above FOC is

$b_i = b(v_i)$, which gives

$$\begin{aligned}
& -v_i + (v_i - b(v_i))\frac{1}{b'(v_i)} = 0 \\
\Longrightarrow\ & b'(v_i)v_i + b(v_i) = v_i \ \Longrightarrow\ \frac{d(b(v_i)v_i)}{dv_i} = v_i \\
\Longrightarrow\ & b(v_i)v_i = \frac{1}{2}v_i^2 + c, \text{ where } c \text{ is a constant.}
\end{aligned}$$

Using the boundary condition $b(0) = 0$,[2] we obtain $c = 0$. Thus, the equilibrium bidding strategy is

$$b(v_i) = \frac{1}{2}v_i.$$

3.A.4. Common Value Auction

In a common value auction, the value of good for sale is the same for all bidders. "Oil well" is an often cited example of common value auctions.

Jar of coins. Let us play the following auction game. There is a jar with some coins. Every bidder bids for the coins in the jar. Rules of the auction are as follows:

- Do not open the jar.
- The winner is the bidder with the highest bid.
- The winner pays his/her own bid and gets the coins in the jar.

This is a common value auction: the amount of money in the jar is certain.

Question 3.3. What is your bidding strategy? Should you bid less or more than your estimate?

Winner's curse. In the common value auctions, the winning bid tends to be higher than the true value of the good. Such a phenomenon is called *winner's curse*.

Question 3.4. Why winner's curse exists?

Let v be the common value, and b_i be Bidder i's bid. Then Bidder i's payoff is

$$\begin{cases} v - b_i & \text{if } b_i \text{ is the highest bid;} \\ 0 & \text{otherwise.} \end{cases}$$

[2] No player should bid more than his/her valuation, i.e., $b(v_i) \leq v_i$ for every v_i. In particular, $b(0) \leq 0$. Besides, bids are non-negative. So $b(0) = 0$.

Bidders only have estimates of the value of the good. Let y_i be Player i's estimate:

$$y_i = v + \tilde{\varepsilon}_i,$$

where $\tilde{\varepsilon}_i$ is Bidder i's estimation error. y_i is also Bidder i's type. Suppose that on average bidders estimate correctly. Then around half of the bidders overestimate. If the bidders bid roughly the same as their estimate, the winner would be the bidder with the largest $\tilde{\varepsilon}_i$. Then, the winning bid would be higher (actually much higher) than the true value.

Question 3.5. After learning winner's curse, how should you bid?

1. If everyone bids roughly their own estimates, then when you (Player i) win, you know that $y_j < y_i$ for all j.
2. You only care how many coins are in the jar if you win.

So, you should bid based not only on your initial estimate y_i but also on the fact that $y_i > y_j$ for all j. Put differently, you should bid as if you know you win.

3.B. Dynamic Games of Incomplete Information

In this section, we will study three specific models of dynamic games of incomplete information: asymmetric information Cournot model with verifiable information in Section 3.B.1, job market signaling model in Section 3.B.2 and a screening model in Section 3.B.3.

The solution concept associated with dynamic games of incomplete information is the *Perfect Bayesian Equilibrium (PBE)*. PBE was invented in order to refine BNE in a similar way that SPE refined NE. We will not study PBE in detail in this course. The definition of PBE is given below for your reference.

Definition 3.B.1 (Perfect Bayesian Equilibrium)**.** A **Perfect Bayesian Equilibrium (PBE)** is a strategy profile σ and a belief system $\mu = (\mu_1, \mu_2, ..., \mu_n)$ where μ_i specifies Player i's belief at each of his/her information sets, satisfying

- Sequential rationality: At each information set, σ_i is a best response to σ_{-i}, given belief μ_i at that information set.
- Belief consistency (on the equilibrium path): At information sets on the equilibrium path, beliefs are determined by Bayes' rule and the players' equilibrium strategies.

- Belief consistency (off the equilibrium path): At information sets off the equilibrium path, beliefs are determined by Bayes' rule and the players' equilibrium strategies where possible.

3.B.1. Asymmetric Cournot with Verifiable Information

Game Setup. Consider the Cournot duopoly model where the two firms choose their quantities simultaneously. Let q_1 and q_2 denote the quantities (of a homogeneous product) produced by firms 1 and 2, respectively. Let $P(Q) = a - Q$ be the market-clearing price when the aggregate quantity on the market is $Q = q_1 + q_2$. Firm 1's cost function is

$$C_1(q_1) = c_M q_1.$$

Firm 2's cost function is

$$C_2(q_2) = \begin{cases} c_H q_2 = (c_M + x)q_2 & \text{with probability } 1/3 \\ c_M q_2 & \text{with probability } 1/3 \\ c_L q_2 = (c_M - x)q_2 & \text{with probability } 1/3 \end{cases}$$

Furthermore, information is asymmetric:

- Firm 2 knows both its own cost function (i.e, the realization of the marginal cost) and Firm 1's cost function, but
- Firm 1 knows its own cost function and only that Firm 2's marginal cost is c_H, c_M or c_L, each with 1/3 probability.

Before the firms choose quantities, Firm 2 can costlessly and verifiably reveal its cost information to Firm 1.

Question 3.6. Should Firm 2 reveal its cost information?

Perhaps it is easier to first consider the following question:

Question 3.7. Would Firm 2 want Firm 1 to know if it has high, middle, or low cost?

First, notice that in the Cournot model, one firm's profit would be higher if the other firm produces less. Then the question becomes "would Firm 1 produce less if it knows that Firm 2 has high, middle or low cost?" To answer this question, we could solve for Firm 1's equilibrium quantities under different scenarios. It is also clear from the reaction curves

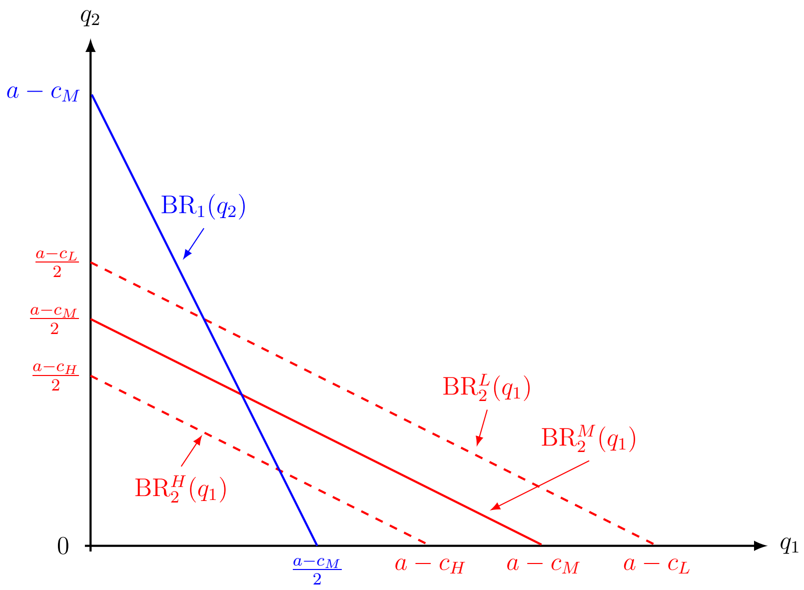

Figure 3.1: Cournot Duopoly

in Figure 3.1. The result is: compared to not knowing Firm 2's cost, Firm 1 produces less (more) if it knows that Firm 2 has low (high) cost. Thus, Firm 2 would want Firm 1 to know if it has low cost. That is, **Firm 2 with low cost would reveal its cost information.**

But the above argument is not over. Let us now consider whether Firm 2 should reveal its cost information when it has middle cost. You might think that Firm 2's decision depends on what Firm 1 thought. However, if Firm 2 doesn't reveal that it has middle cost, then Firm 1 knows that the cost is not low. This is because Firm 2 would reveal its cost information if it has low cost, as is argued in the previous paragraph. Put it differently, Firm 1 knows that the cost is either middle or high. As a result, Firm 2 with middle cost would want Firm 1 to know it so that Firm 1 would produce less. That is, **Firm 2 with middle cost would also reveal its cost information.**

Then, **for Firm 2 with high cost, it really doesn't matter whether it reveals or not.** Because even if it does not reveal, since Firm 2 with middle or low costs would reveal, the fact of no revealing reveals that Firm 2 has high cost.

Remark 3.2. The same argument goes through if Firm 2 has more types.

This idea is called *information unraveling*.

3.B.2. Job-Market Signaling

Suppose that there are two types of workers, high-ability and low-ability. They differ in productivity: high-ability worker has productivity of 100 whereas low-ability worker has productivity of 60. In the population, 20% of workers are high-ability and 80% are low-ability.

	Productivity	Proportion
High-ability Worker	100	20%
Low-ability Worker	60	80%

Suppose that firms are competitive. Thus, firms would offer 100 to a high-ability worker and 60 to a low-ability worker if they could identify the worker's types. If firms cannot identify the worker's types, they would offer $100 * 20\% + 60 * 80\% = 68$.

Question 3.8. Suppose that you are a high-ability worker, how can you make the firms know it? In particular, would it work if you simply tell the firms "I am a high-ability worker"?

Spence (1973) brings up the idea that "education" could be used as a costly signal to differentiate high-ability workers from low-ability ones. The crucial assumption in Spence's model is that low-ability workers find education more costly than high-ability workers. Suppose that it costs high-ability workers 9 for a year of education and low-ability workers 21.

	Cost
High-ability Worker	9
Low-ability Worker	21

We argue that there exists an equilibrium where

- High-ability workers take the three-year graduate education but low-ability workers do not.
- Employers identify those workers with graduate degrees as high-ability workers and those without degrees as low-ability workers. Employers offer 100 to a worker with degree and 60 to a worker without degree.

We have not gone through the concept of PBE in detail. In essence, PBE requires the strategies to be best responses given the belief system and the beliefs to be consistent with the strategy profile. For this particular game, we need to check

1. Both types of workers would not deviate in their respective education choices.
2. Employers' beliefs are consistent with the equilibrium behavior.

The second point is obvious. For the first point,

- A high-ability worker obtains $100 - 9 * 3 = 73$ if he/she takes the education and 60 if not.
- A low-ability worker obtains $100 - 21 * 3 = 37$ if he/she takes the education and 60 if not.

Thus, a high-ability worker would not deviate to not taking the education and a low-ability worker would not deviate to taking the education.

Remark 3.3. This is called a *separating equilibrium* because in equilibrium the types separate and get identified.

Question 3.9. What is the education program only takes two years? How about one year?

Remark 3.4. For the separation to work, there must be enough differences in costs for the two types of workers.

Remark 3.5. If the standard of obtaining education becomes lower, then probably we will see qualification inflation.

Remark 3.6. Education increases inequality: Compared to the no education outcome, a three-year education program makes high-ability workers better-off ($73 > 68$) and low-ability workers worse-off ($60 < 68$).

Remark 3.7. It is possible that high-ability workers are also worse-off. To see this, consider a four-year education program. High-ability workers still have the incentive to get educated, since in the separating equilibrium, no education is interpreted as evidence of low ability and the no signaling outcome is no longer available.

3.B.3. Screening

In the last section, we have seen a *signaling* model in which the informed parties (i.e., the workers) move first. Signaling models are closely related to *screening* models, in which the uninformed parties take the lead. Classic references of screening models concern the context of insurance markets. But in this course, we still take the job market as an example.

Job market example. In the previous signaling model, the workers choose education first. And after observing the education choices, the firms offer wages accordingly. Now consider the following timing, which corresponds to a screening setting:

1. Two firms simultaneously announce a menu of contracts specifying the required years of education and the wage offer (e, w).
2. Given these contracts, the workers choose which contract to accept, if any.

The productivity and proportion of workers, as well as the cost of a year of education to each types of workers are all the same as in the signaling case.

Question 3.10. Is it an equilibrium that both firms offer the same two contracts $(e_H = 3, w_H = 100)$ and $(e_L = 0, w_L = 60)$? The outcome is the same as a three-year education program in the signaling model.

For the workers, similar arguments as in the signaling model apply. Both types of workers would self-select the contracts designed for them.

- A high-ability worker obtains $100 - 9 * 3 = 73$ if he/she takes the contract $(e_H = 3, w_H = 100)$ and 60 if takes $(e_L = 0, w_L = 60)$.
- A low-ability worker obtains $100 - 21 * 3 = 37$ if he/she takes the contract $(e_H = 3, w_H = 100)$ and 60 if takes $(e_L = 0, w_L = 60)$.

But for the firms, if the contracts $(e_H = 3, w_H = 100)$ and $(e_L = 0, w_L = 60)$ are offered and accepted by the respective types of workers, each firm obtains a payoff of 0 since the wages and the productivities are the same. A firm could be better-off by offering $(e'_H = 2, w'_H = 95)$ and $(e_L = 0, w_L = 60)$.

- The high-ability workers prefer $(e'_H = 2, w'_H = 95)$ to $(e_H = 3, w_H = 100)$: they obtain $95 - 9 * 2 = 77$ (> 73) if taking $(e'_H = 2, w'_H = 95)$.
- The low-ability workers would not take $(e'_H = 2, w'_H = 95)$: they obtain $95 - 21 * 2 = 53$ (< 60) if taking $(e'_H = 2, w'_H = 95)$.

Thus, by deviating to offering $(e'_H = 2, w'_H = 90)$ and $(e_L = 0, w_L = 60)$, the firm obtains $(100 - 95) * 20\% = 1 > 0$.

Question 3.11. How about both firms offer the same two contracts $(e_H = 2, w_H = 100)$ and $(e_L = 0, w_L = 60)$?

Similar to the previous argument, if such contracts are offered, workers would self-select the contract designed for them.

For the firms, a firm could deviate to offering contracts that each attracts only one types of workers or offering a contract that attracts both types of workers. Moreover, when a contract attracts only one types of workers, offering $w_L > 60$ or $w_H > 100$ results in a loss. On the other hand, for a contract with a lower wage to be accepted by the workers, the education requirement must be lower. Therefore, a firm could not benefit from offering a contract that attracts the low-ability workers alone. And we only need to consider the following two types of deviations:

1. A firm could deviate to an offer $(e'_H = 1, w'_H < 100)$ to attract the high-ability workers. The high-ability workers prefer $(e'_H = 1, w'_H)$ to $(e_H = 2, w_H = 100)$ if $w - 9 \geq 100 - 9 * 2 \implies w \geq 91$. However, when this is the case, the low-ability workers also prefer $(e'_H = 1, w'_H)$ since $w'_H - 21 \geq 70$ (> 60). Then the firm obtains $(100 - w'_H) * 20\% + (60 - w'_H) * 80\% = 68 - w'_H < 0$. And thus this is not a profitable deviation.
2. Another possible deviation is to offer $(e = 0, w)$ and attract both types of workers. The low-ability workers would accept this contract as long as $w \geq 60$ and the high-ability workers would accept this contract is $w \geq 100 - 9 * 2 = 82$. Thus, the lowest wage is 82 and the firm obtains $68 - 82 < 0$.

Therefore, it is an equilibrium that both firms offer the two contracts $(e_H = 2, w_H = 100)$ and $(e_L = 0, w_L = 60)$.

Remark 3.8. Separating equilibria do not always exist. For example, if we change the proportion of high-ability workers to 80%, then there will be no separating equilibria.