2
Equilibrium

The theory of *equilibrium* attempts to predict what happens in a game when players behave strategically. This is a central concept to this text as, in mechanism design, we are optimizing over games to find games with good equilibria. Here, we review the most fundamental notions of equilibrium. They will all be static notions in that players are assumed to understand the game and will play once in the game. While such foreknowledge is certainly questionable, some justification can be derived from imagining the game in a dynamic setting where players can learn from past play.

This chapter reviews equilibrium in both complete and incomplete information games. As games of incomplete information are the most central to mechanism design, special attention will be paid to them. In particular, we will characterize equilibrium when the private information of each agent is single-dimensional and corresponds, for instance, to a value for receiving a good or service. We will show that auctions with the same equilibrium outcome have the same expected revenue. Using this so-called *revenue equivalence* we will describe how to solve for the equilibrium strategies of standard auctions in symmetric environments.

Our emphasis will be on demonstrating the central theories of equilibrium and not on providing the most comprehensive or general treatment. For more general discussion of equilibrium, readers are recommended to consult a game theory textbook.

2.1 Complete Information Games

In games of compete information all players are assumed to know precisely the payoff structure of all other players for all possible outcomes of the game. A classic example of such a game is the *prisoner's dilemma*, the story for which is as follows.

Two prisoners, Bonnie and Clyde, have jointly committed a crime and are being interrogated in separate quarters. Unfortunately, the interrogators are unable to prosecute either prisoner without a confession. Bonnie is offered the following deal: If she confesses and Clyde does not, she will be released and Clyde will serve the full sentence of ten years in prison. If they both confess, she will share the sentence and serve five years. If neither confesses, she will be prosecuted for a minimal offense and receive a year of prison. Clyde is offered the same deal.

This story can be expressed as the following *bimatrix game* where entry (a, b) represents row player's payoff a and column player's payoff b.

	silent	confess
silent	(-1,-1)	(-10,0)
confess	(0,-10)	(-5,-5)

A simple thought experiment enables prediction of what will happen in the prisoner's dilemma. Suppose the Clyde is silent. What should Bonnie do? Remaining silent as well results in one year of prison while confessing results in immediate release. Clearly confessing is better. Now suppose that Clyde confesses. Now what should Bonnie do? Remaining silent results in ten years of prison while confessing as well results in only five. Clearly confessing is better. In other words, no matter what Clyde does, Bonnie is better of by confessing. The prisoner's dilemma is hardly a dilemma at all: the *strategy profile* (confess, confess) is a *dominant strategy equilibrium*.

Definition 2.1.1. *A* dominant strategy equilibrium *(DSE) in a complete information game is a strategy profile in which each player's strat-*

Chapter 2: Topics Covered.
- equilibrium: Nash, Bayes-Nash, dominant strategy;
- game forms: complete information, bimatrix, incomplete information, single-dimensional;
- characterization of equilibria in single-dimensional games;
- revenue equivalence;
- solving for Bayes-Nash equilibria;
- uniqueness of Bayes-Nash equilibria; and
- incentive compatibility and the revelation principle.

egy is as least as good as all other strategies regardless of the strategies of all other players.

Dominant strategy equilibrium is a strong notion of equilibrium and is therefore unsurprisingly rare. For an equilibrium notion to be complete it should identify equilibrium in every game. Another well studied game is *chicken*.

James Dean and Buzz (in the movie *Rebel without a Cause*) face off at opposite ends of the street. On the signal they race their cars on a collision course towards each other. The options each have are to swerve or to stay their course. Clearly if they both stay their course they crash. If they both swerve (opposite directions) they escape with their lives but the match is a draw. Finally, if one swerves and the other stays, the one that stays is the victor and the other the loses.[1]

A reasonable bimatrix game depicting this story is the following.

	stay	swerve
stay	(-10,-10)	(1,-1)
swerve	(-1,1)	(0,0)

Again, a simple thought experiment enables us to predict how the players might play. Suppose James Dean is going to stay, what should Buzz do? If Buzz stays they crash and Buzz's payoff is -10, but if Buzz swerves his payoff is only -1. Clearly, of these two options Buzz prefers to swerve. Suppose now that Buzz is going to swerve, what should James Dean do? If James Dean stays he wins and his payoff is one, but if he swerves it is a draw and his payoff is zero. Clearly, of these two options James Dean prefers to stay. What we have shown is that the strategy profile (stay, swerve) is a mutual best response, a.k.a., a *Nash equilibrium*. Of course, the game is symmetric so the opposite strategy profile (swerve, stay) is also an equilibrium.

Definition 2.1.2. *A* Nash equilibrium *in a game of complete information is a strategy profile where each players strategy is a best response to the strategies of the other players as given by the strategy profile.*

In the examples above, the strategies of the players correspond directly to actions in the game, a.k.a., *pure strategies*. In general, Nash equilibrium strategies can be randomizations over actions in the game, a.k.a., *mixed strategies* (see Exercise 2.1).

[1] The actual chicken game depicted in *Rebel without a Cause* is slightly different from the one described here.

2.2 Incomplete Information Games

Now we turn to the case where the payoff structure of the game is not completely known. We will assume that each agent has some private information and this information affects the payoff of this agent in the game. We will refer to this information as the agent's type and denote it by t_i for agent i. The profile of types for the n agents in the game is $\mathbf{t} = (\mathsf{t}_1, \ldots, \mathsf{t}_n)$.

A *strategy* in a game of incomplete information is a function that maps an agent's type to any of the agent's possible actions in the game (or a distribution over actions for mixed strategies). We will denote by $b_i(\cdot)$ the strategy of agent i and $\boldsymbol{b} = (b_1, \ldots, b_n)$ a *strategy profile*.

The auctions described in Chapter 1 were games of incomplete information where an agent's private type was her value for receiving the item, i.e., $\mathsf{t}_i = \mathsf{v}_i$. As we described, strategies in the ascending-price auction were $b_i(\mathsf{v}_i) =$ "drop out when the price exceeds v_i" and strategies in the second-price auction were $b_i(\mathsf{v}_i) =$ "bid $\mathsf{b}_i = \mathsf{v}_i$." We refer to this latter strategy as *truthtelling*. Both of these strategy profiles are in *dominant strategy equilibrium* for their respective games.

Definition 2.2.1. *A* dominant strategy equilibrium *(DSE) is a strategy profile* $\boldsymbol{b}$ *such that for all* i, t_i, *and* $\mathbf{b}_{-i}$ *(where* $\mathbf{b}_{-i}$ *generically refers to the actions of all players but* i*), agent* i*'s utility is maximized by following strategy* $b_i(\mathsf{t}_i)$.

Notice that aside from strategies being defined as a map from types to actions, this definition of DSE is identical to the definition of DSE for games of complete information.

2.3 Bayes-Nash Equilibrium

Naturally, many games of incomplete information do not have dominant strategy equilibria. Therefore, we will also need to generalize Nash equilibrium to this setting. Recall that equilibrium is a property of a strategy profile. It is in equilibrium if each agent does not want to change her strategy given the other agents' strategies. For an agent i, we want to the fix other agent strategies and let i optimize her strategy (meaning: calculate her best response for all possible types t_i she may have). This is an ill specified optimization as just knowing the other agents' strategies is not enough to calculate a best response. Additionally, i's best response depends on i's beliefs on the types of the other agents. The standard economic treatment addresses this ill specification by assuming a common prior.

Definition 2.3.1. *Under the* common prior assumption*, the agent types* $\mathbf{t}$ *are drawn at random from a* prior distribution $\boldsymbol{F}$ *(a joint probability distribution over type profiles) and this prior distribution is* common knowledge*.*

The distribution $\boldsymbol{F}$ over $\mathbf{t}$ may generally be correlated. Which means that an agent with knowledge of her own type must do *Bayesian updating* to determine the distribution over the types of the remaining bidders. We denote this conditional distribution as $\boldsymbol{F}_{-i}\big|_{t_i}$. Of course, when the distribution of types is independent, i.e., $\boldsymbol{F}$ is the *product distribution* $F_1 \times \cdots \times F_n$, then $\boldsymbol{F}_{-i}\big|_{t_i} = \boldsymbol{F}_{-i}$.

Notice that a prior $\boldsymbol{F}$ and strategies $\boldsymbol{b}$ induces a distribution over the actions of each of the agents. With such a distribution over actions, the problem each agent faces of optimizing her own action is fully specified.

Definition 2.3.2. *A* Bayes-Nash equilibrium (BNE) *for a game G and common prior $\boldsymbol{F}$ is a strategy profile $\boldsymbol{b}$ such that for all i and t_i, $b_i(t_i)$ is a best response when other agents play $\boldsymbol{b}_{-i}(\mathbf{t}_{-i})$ when $\mathbf{t}_{-i} \sim \boldsymbol{F}_{-i}\big|_{t_i}$.*

To illustrate Bayes-Nash equilibrium, consider using the first-price auction to sell a single item to one of two agents, each with valuation drawn independently and identically from the uniform distribution on $[0, 1]$, i.e., the common prior distribution is $\boldsymbol{F} = F \times F$ with $F(z) = \mathbf{Pr}_{v \sim F}[v < z] = z$. Here each agent's type is her valuation. We will calculate the BNE of this game by the "guess and verify" technique. First, we guess that there is a symmetric BNE with $b_i(z) = z/2$ for $i \in \{1, 2\}$. Second, we calculate agent 1's expected utility with any value v_1 and any bid b_1 under the standard assumption that the agent's utility u_i is her value less her payment (when she wins). In this calculation v_1 and b_1 are fixed and $b_2 = v_2/2$ is random. By the definition of the first-price auction:

$$\mathbf{E}[u_1] = (v_1 - b_1) \times \mathbf{Pr}[1 \text{ wins with bid } b_1] \,.$$

Calculate $\mathbf{Pr}[1 \text{ wins with } b_1]$ as

$$\begin{aligned}\mathbf{Pr}[b_2 \leq b_1] = \mathbf{Pr}[v_2/2 \leq b_1] = \mathbf{Pr}[v_2 \leq 2b_1] = F(2b_1) \\ = 2b_1.\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{E}[\mathsf{u}_1] &= (\mathsf{v}_1 - \mathsf{b}_1) \times 2\mathsf{b}_1 \\ &= 2\mathsf{v}_1\mathsf{b}_1 - 2\mathsf{b}_1^2.\end{aligned}$$

Third, we optimize agent 1's bid. Agent 1 with value v_1 should maximize $2\mathsf{v}_1\mathsf{b}_1 - 2\mathsf{b}_1^2$ as a function of b_1, and to do so, can differentiate the function and set its derivative equal to zero. The result is $\frac{\mathrm{d}}{\mathrm{d}\mathsf{b}_1}(2\mathsf{v}_1\mathsf{b}_1 - 2\mathsf{b}_1^2) = 2\mathsf{v}_1 - 4\mathsf{b}_1 = 0$ and we can conclude that the optimal bid is $\mathsf{b}_1 = \mathsf{v}_1/2$. This proves that agent 1 should bid as prescribed if agent 2 does; and vice versa. Thus, we conclude that the guessed strategy profile is in BNE.

In Bayesian games it is useful to distinguish between stages of the game in terms of the knowledge sets of the agents. The three stages of a Bayesian game are *ex ante*, *interim*, and *ex post*. The ex ante stage is before values are drawn from the distribution. Ex ante, the agents know this distribution but not their own types. The interim stage is immediately after each agent learns her own type, but before playing in the game. In the interim, an agent assumes the other agent types are drawn from the prior distribution conditioned on her own type, i.e., via *Bayesian updating*. In the ex post stage, the game is played and the types (and actions) of all agents are known.

2.4 Independent Single-dimensional Games

We will focus on a conceptually simple class of single-dimensional games that is relevant to the auction problems we have already discussed. In a single-dimensional game, each agent's private type is her value for receiving an abstract service, i.e., $\mathsf{t}_i = \mathsf{v}_i$. The distribution over types is independent (i.e., a product distribution $\boldsymbol{F} = F_1 \times \cdots \times F_n$). A game has an outcome $\mathbf{x} = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$ and payments $\mathbf{p} = (\mathsf{p}_1, \ldots, \mathsf{p}_n)$ where x_i is an indicator for whether agent i indeed received their desired service, i.e., $\mathsf{x}_i = 1$ if i is served and 0 otherwise. Price p_i will denote the payment i makes to the mechanism. An agent's value can be positive or negative and an agent's payment can be positive or negative. An agent's utility is linear in her value and payment and specified by $\mathsf{u}_i = \mathsf{v}_i\mathsf{x}_i - \mathsf{p}_i$. Agents are risk-neutral expected utility maximizers.

Definition 2.4.1. *A* single-dimensional linear utility *is defined as having utility* $\mathsf{u} = \mathsf{v}\mathsf{x} - \mathsf{p}$ *for service-payment outcomes* (x, p) *and private* *value* v*; a* single-dimensional linear agent *possesses such a utility function.*

A game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ maps actions $\mathbf{b}$ of agents to an allocation and payments. Formally we will specify these allocations and payments on bids $\mathbf{b}$ as $\tilde{\boldsymbol{x}}(\mathbf{b})$ and $\tilde{\boldsymbol{p}}(\mathbf{b})$ with

- $\tilde{x}_i(\mathbf{b})$ = outcome to i when actions are $\mathbf{b}$, and
- $\tilde{p}_i(\mathbf{b})$ = payment from i when actions are $\mathbf{b}$.

Given a game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ and a strategy profile $\boldsymbol{b}$ we can express the outcome and payments of the game as a function of the valuation profile. From the point of view of analysis this description of the game outcome is much more relevant. Define

- $x_i(\mathbf{v}) = \tilde{x}_i(\boldsymbol{b}(\mathbf{v}))$, and
- $p_i(\mathbf{v}) = \tilde{p}_i(\boldsymbol{b}(\mathbf{v}))$.

We refer to the former as the *ex post allocation rule* and the latter as the *ex post payment rule* for game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ and strategy $\boldsymbol{b}$ (implicit). Consider an agent i's interim perspective. She knows her own value v_i and believes the other agents' values to be drawn from the distribution $\boldsymbol{F}$ (conditioned on her value). For $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$, $\boldsymbol{b}$, and $\boldsymbol{F}$ taken implicitly we can specify agent i's *interim allocation and payment rules* as functions of v_i.

- $x_i(\mathsf{v}_i) = \mathbf{Pr}[x_i(\mathbf{v}) = 1 \mid \mathsf{v}_i] = \mathbf{E}[x_i(\mathbf{v}) \mid \mathsf{v}_i]$, and
- $p_i(\mathsf{v}_i) = \mathbf{E}[p_i(\mathbf{v}) \mid \mathsf{v}_i]$.

Technical Note. The notation developed in this section will be used systematically throughout the text. Scalar values pertaining to an agent, e.g., the valuation of an agent or an allocation probability will be denoted as v and x; when an agent i is designated they will be subscripted, e.g., v_i and x_i; and for profile of multiple agents they will be in bold, e.g., $\mathbf{v}$ and $\mathbf{x}$. Interim functions, e.g., allocation and payment rules, will be denoted as x and p for values and as $\tilde{x}$ and $\tilde{p}$ for bids. An agent i's cumulative distribution function and strategy are an interim functions and thus denote b_i and F_i, respectively. Ex post functions which map profiles to outcomes will be denoted, e.g., x and p for values and $\tilde{x}$ and $\tilde{p}$ for bids. Importantly, a profile of interim allocation rules $\boldsymbol{x}$ is in $(\mathbb{R} \to [0,1])^n$ while the profile of ex post allocation rules $\boldsymbol{x}$ is in $\mathbb{R}^n \to [0,1]^n$.

With linearity of expectation we can combine these with the agent's utility function to write

- $u_i(\mathsf{v}_i) = \mathsf{v}_i\, x_i(\mathsf{v}_i) - p_i(\mathsf{v}_i)$.

Finally, we say that a strategy $b_i(\cdot)$ is *onto* if every action b_i agent i could play in the game is prescribed by b_i for some value v_i, i.e., $\forall \mathsf{b}_i \exists \mathsf{v}_i\, b_i(\mathsf{v}_i) = \mathsf{b}_i$. We say that a strategy profile is *onto* if the strategy of every agent is onto. For instance, the truthtelling strategy in the second-price auction is onto. When the strategies of the agents are onto, the interim allocation and payment rules defined above completely specify whether the strategies are in equilibrium or not. In particular, BNE requires that each agent (weakly) prefers playing the action corresponding (via their strategy) to her value than the action corresponding to any other value.

Proposition 2.4.1. *When values are drawn from a product distribution* $\boldsymbol{F}$*; single-dimensional game* $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ *and strategy profile* $\boldsymbol{b}$ *are in Bayes-Nash equilibrium only if for all* i, v_i, *and* z,

$$\mathsf{v}_i\, x_i(\mathsf{v}_i) - p_i(\mathsf{v}_i) \geq \mathsf{v}_i\, x_i(\mathsf{z}) - p_i(\mathsf{z}).$$

If the strategy profile is onto then the converse also holds.

Notice that in Proposition 2.4.1 the distribution $\boldsymbol{F}$ is required to be a product distribution. If $\boldsymbol{F}$ is not a product distribution, then when agent i's value is v_i then $x_i(\mathsf{z})$ is not generally the probability that she will win when she follows her designated strategy for value z. This distinction arises because the conditional distribution of the other agents' values need not be the same when i's value is v_i or z.

2.5 Characterization of Bayes-Nash Equilibrium

We now discuss what Bayes-Nash equilibria look like. For instance, when given $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$, $\boldsymbol{b}$, and $\boldsymbol{F}$ we can calculate the interim allocation and payment rules $x_i(\mathsf{v}_i)$ and $p_i(\mathsf{v}_i)$ of each agent. We want to succinctly describe properties of these allocation and payment rules that can arise as BNE.

Theorem 2.5.1. *When values are drawn from a continuous product distribution* $\boldsymbol{F}$*; single dimensional* $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ *and strategy profile* $\boldsymbol{b}$ *are in BNE only if for all* i,

(i) *(monotonicity)* $x_i(\mathsf{v}_i)$ *is monotone non-decreasing, and*

(ii) *(payment identity)* $p_i(\mathsf{v}_i) = \mathsf{v}_i\, x_i(\mathsf{v}_i) - \int_0^{\mathsf{v}_i} x_i(\mathsf{z})\, \mathrm{d}\mathsf{z} + p_i(0)$,

where often $p_i(0) = 0$. *If the strategy profile is onto then the converse also holds.*

Proof. We will prove the theorem in the special case where the support of each agent i's distribution is $[0, \infty]$. Focusing on a single agent i, who we will refer to as Alice, we drop subscripts i from all notations.

We break this proof into three pieces. First, we show, by picture, that the game is in BNE if the characterization holds and the strategy profile is onto. Next, we will prove that a game is in BNE only if the monotonicity condition holds. Finally, we will prove that a game is in BNE only if the payment identity holds.

Note that if Alice with value v deviates from the equilibrium and takes action $b(\mathsf{v}^\dagger)$ instead of $b(\mathsf{v})$ then she will receive outcome and payment $x(\mathsf{v}^\dagger)$ and $p(\mathsf{v}^\dagger)$. This motivates the definition,

$$u(\mathsf{v}, \mathsf{v}^\dagger) = \mathsf{v}\, x(\mathsf{v}^\dagger) - p(\mathsf{v}^\dagger),$$

which corresponds to Alice utility when she makes this deviation. For Alice's strategy to be in equilibrium it must be that for all v, and $\mathsf{v}^\dagger$, $u(\mathsf{v}, \mathsf{v}) \geq u(\mathsf{v}, \mathsf{v}^\dagger)$, i.e., Alice derives no increased utility by deviating. The strategy profile $\boldsymbol{b}$ is in equilibrium if and only if the same condition holds for all agents. (The "if" direction here follows from the assumption that strategies map values onto actions. Meaning: for any action in the game there exists a value $\mathsf{v}^\dagger$ such that $b(\mathsf{v}^\dagger)$ is that action.)

(i) Game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$, strategies $\boldsymbol{b}$, and distributions $\boldsymbol{F}$ are in BNE if $\boldsymbol{b}$ is onto and monotonicity and the payment identity hold.

We prove this by picture. Though the formulaic proof is simple, the pictures provide useful intuition. We consider two possible values z_1 and z_2 with $\mathsf{z}_1 < \mathsf{z}_2$. Supposing Alice has the high value, $\mathsf{v} = \mathsf{z}_2$, we argue that Alice does not benefit by simulating her strategy for the lower value, $\mathsf{v}^\dagger = \mathsf{z}_1$, i.e., by playing $b(\mathsf{v}^\dagger)$ to obtain outcome $x(\mathsf{v}^\dagger)$ and payment $p(\mathsf{v}^\dagger)$. We leave the proof of the opposite, that when $\mathsf{v} = \mathsf{z}_1$ and Alice is considering simulating the higher strategy $\mathsf{v}^\dagger = \mathsf{z}_2$, as an exercise for the reader.

To start with this proof, we assume that $x(\mathsf{v})$ is monotone and that $p(\mathsf{v}) = \mathsf{v}\, x(\mathsf{v}) - \int_0^{\mathsf{v}} x(\mathsf{z})\, \mathrm{d}\mathsf{z}$.

Consider the diagrams in Figure 2.1. The first diagram (top, center) shows $x(\cdot)$ which is indeed monotone as per our assumption. The column on the left shows Alice's surplus, $\mathsf{v}\, x(\mathsf{v})$; payment, $p(\mathsf{v})$, and

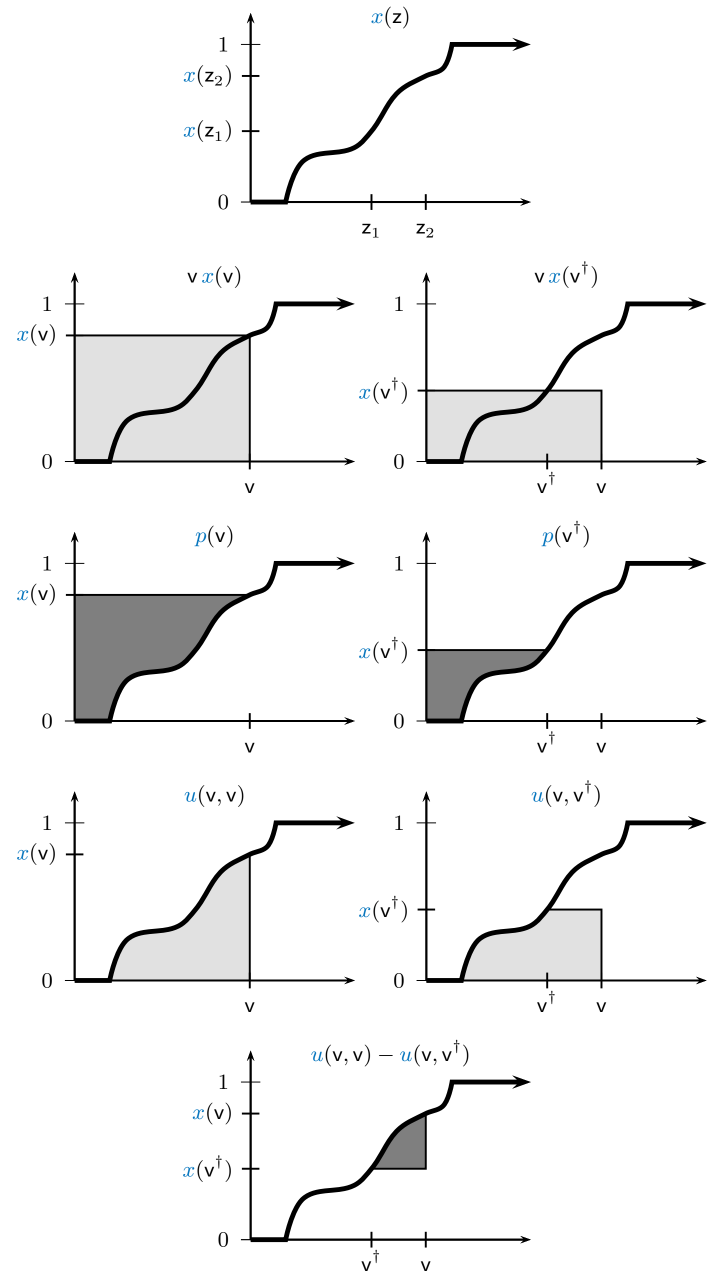


Figure 2.1. The left column shows (shaded) the surplus, payment, and utility of Alice playing action $b(\mathsf{v} = \mathsf{z}_2)$. The right column shows (shaded) the same for Alice playing action $b(\mathsf{v}^\dagger = \mathsf{z}_1)$. The final diagram shows (shaded) the difference between Alice's utility for these strategies. Monotonicity implies this difference is non-negative.

utility, $u(\mathsf{v}) = \mathsf{v}\, x(\mathsf{v}) - p(\mathsf{v})$, assuming that she follow the BNE strategy $b(\mathsf{v} = \mathsf{z}_2)$. The column on the right shows the analogous quantities when Alice follows strategy $b(\mathsf{v}^\dagger = \mathsf{z}_1)$ but has value $\mathsf{v} = \mathsf{z}_2$. The final diagram (bottom, center) shows the difference in the Alice's utility for the outcome and payments of these two strategies. Note that as the picture shows, the monotonicity of the allocation function implies that this difference is always non-negative. Therefore, there is no incentive for Alice to simulate the strategy of a lower value.

As mentioned, a similar proof shows that Alice has no incentive to simulate her strategy for a higher value. We conclude that she (weakly) prefers to play the action given by the BNE $b(\cdot)$ over any other action in the range of her strategy function; since $b(\cdot)$ is onto this range includes all actions.

(ii) Game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$, strategies $\boldsymbol{b}$, and distributions $\boldsymbol{F}$ are in BNE only if the allocation rule is monotone.

If we are in BNE then for all valuations, v and $\mathsf{v}^\dagger$, $u(\mathsf{v}, \mathsf{v}) \geq u(\mathsf{v}, \mathsf{v}^\dagger)$. Expanding we require

$$\mathsf{v}\, x(\mathsf{v}) - p(\mathsf{v}) \geq \mathsf{v}\, x(\mathsf{v}^\dagger) - p(\mathsf{v}^\dagger).$$

We now consider z_1 and z_2 with $\mathsf{z}_1 < \mathsf{z}_2$ and take turns setting $\mathsf{v} = \mathsf{z}_1$, $\mathsf{v}^\dagger = \mathsf{z}_2$, and $\mathsf{v}^\dagger = \mathsf{z}_1$, $\mathsf{v} = \mathsf{z}_2$. This yields the following two inequalities:

$$\mathsf{v} = \mathsf{z}_2, \mathsf{v}^\dagger = \mathsf{z}_1 : \quad \mathsf{z}_2\, x(\mathsf{z}_2) - p(\mathsf{z}_2) \geq \mathsf{z}_2\, x(\mathsf{z}_1) - p(\mathsf{z}_1), \text{ and} \tag{2.5.1}$$

$$\mathsf{v} = \mathsf{z}_1, \mathsf{v}^\dagger = \mathsf{z}_2 : \quad \mathsf{z}_1\, x(\mathsf{z}_1) - p(\mathsf{z}_1) \geq \mathsf{z}_1\, x(\mathsf{z}_2) - p(\mathsf{z}_2). \tag{2.5.2}$$

Adding these inequalities and canceling the payment terms we have,

$$\mathsf{z}_2\, x(\mathsf{z}_2) + \mathsf{z}_1\, x(\mathsf{z}_1) \geq \mathsf{z}_2\, x(\mathsf{z}_1) + \mathsf{z}_1\, x(\mathsf{z}_2).$$

Rearranging,

$$(\mathsf{z}_2 - \mathsf{z}_1)\,(x(\mathsf{z}_2) - x(\mathsf{z}_1)) \geq 0.$$

For $\mathsf{z}_2 - \mathsf{z}_1 > 0$ it must be that $x(\mathsf{z}_2) - x(\mathsf{z}_1) \geq 0$, i.e., $x(\cdot)$ is monotone non-decreasing.

(iii) Game $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$, strategies $\boldsymbol{b}$, and distributions $\boldsymbol{F}$ are in BNE only if the payment rule satisfies the payment identity.

We will give two proofs that payment rule must satisfy $p(\mathsf{v}) = \mathsf{v}\, x(\mathsf{v}) - \int_0^{\mathsf{v}} x(\mathsf{z})\, \mathsf{dz} + p(0)$; the first is a calculus-based proof under the assumption that and each of $x(\cdot)$ and $p(\cdot)$ are continuous and differentiable and the second is a picture-based proof that requires no assumption.

Calculus-based proof: Fix v and recall that $u(\mathsf{v},\mathsf{z}) = \mathsf{v}\,x(\mathsf{z}) - p(\mathsf{z})$. Let $u'(\mathsf{v},\mathsf{z})$ be the partial derivative of $u(\mathsf{v},\mathsf{z})$ with respect to z. Thus, $u'(\mathsf{v},\mathsf{z}) = \mathsf{v}\,x'(\mathsf{z}) - p'(\mathsf{z})$, where x' and p' are the derivatives of p and x, respectively. Since BNE implies that $u(\mathsf{v},\mathsf{z})$ is maximized at $\mathsf{z} = \mathsf{v}$. It must be that

$$u'(\mathsf{v},\mathsf{v}) = \mathsf{v}\,x'(\mathsf{v}) - p'(\mathsf{v}) = 0.$$

This formula must hold true for all values of v. For remainder of the proof, we treat this identity formulaically. To emphasize this, substitute $\mathsf{z} = \mathsf{v}$:

$$\mathsf{z}\,x'(\mathsf{z}) - p'(\mathsf{z}) = 0.$$

Solving for $p'(\mathsf{z})$ and then integrating both sides of the equality from 0 to v we have,

$$p'(\mathsf{z}) = \mathsf{z}\,x'(\mathsf{z}), \text{ so}$$
$$\int_0^{\mathsf{v}} p'(\mathsf{z})\,\mathrm{d}\mathsf{z} = \int_0^{\mathsf{v}} \mathsf{z}\,x'(\mathsf{z})\,\mathrm{d}\mathsf{z}.$$

Simplifying the left-hand side and adding $p(0)$ to both sides,

$$p(\mathsf{v}) = \int_0^{\mathsf{v}} \mathsf{z}\,x'(\mathsf{z})\,\mathrm{d}\mathsf{z} + p(0).$$

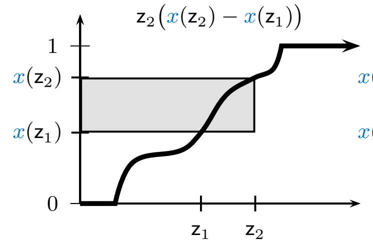
Finally, we obtained the desired formula by integrating the right-hand side by parts,

$$\begin{aligned} p(\mathsf{v}) &= \Big[\mathsf{z}\,x(\mathsf{z})\Big]_0^{\mathsf{v}} - \int_0^{\mathsf{v}} x(\mathsf{z})\,\mathrm{d}\mathsf{z} + p(0) \\ &= \mathsf{v}\,x(\mathsf{v}) - \int_0^{\mathsf{v}} x(\mathsf{z})\,\mathrm{d}\mathsf{z} + p(0). \end{aligned}$$

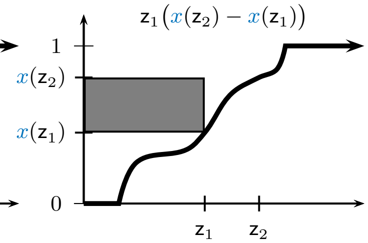
Picture-based proof: Consider equations (2.5.1) and (2.5.2) and solve for $p(\mathsf{z}_2) - p(\mathsf{z}_1)$ in each:

$$\mathsf{z}_2\,(x(\mathsf{z}_2) - x(\mathsf{z}_1)) \geq p(\mathsf{z}_2) - p(\mathsf{z}_1) \geq \mathsf{z}_1\,(x(\mathsf{z}_2) - x(\mathsf{z}_1)).$$

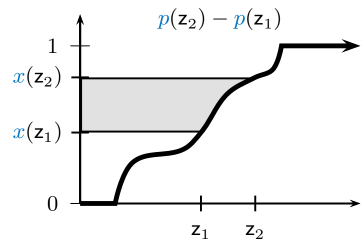
The first inequality gives an upper bound on the difference in payments for two types z_2 and z_1 and the second inequality gives a lower bound. It is easy to see that the only payment rule that satisfies these upper and lower bounds for all pairs of types z_2 and z_1 has payment difference exactly equal to the area to the left of the allocation rule between $x(\mathsf{z}_1)$ and $x(\mathsf{z}_2)$. See Figure 2.2. The payment identity follows by taking $\mathsf{z}_1 = 0$ and $\mathsf{z}_2 = \mathsf{v}$. □



(a) Payment upper bound



(b) Payment lower bound



(c) Payment identity

Figure 2.2. Upper (top, left) and lower bounds (top, right) for the difference in payments for two strategies z_1 and z_2 imply that the difference in payments (bottom) must satisfy the payment identity.

As we conclude the proof of the BNE characterization theorem, it is important to note how little we have assumed of the underlying game. We did not assume it was a single-round, sealed-bid auction. We did not assume that only a winner will make payments. Therefore, we conclude for any potentially wacky, multi-round game the outcomes of all Bayes-Nash equilibria have a nice form.

2.6 Characterization of Dominant Strategy Equilibrium

Dominant strategy equilibrium is a stronger equilibrium concept than Bayes-Nash equilibrium. All dominant strategy equilibria are Bayes-Nash equilibria, but as we have seen, the opposite is not true; for instance, there is no DSE in the first-price auction. Recall that a strategy profile is in DSE if each agent's strategy is optimal for her regardless of what other agents are doing. The DSE characterization theorem below follows from the BNE characterization theorem.

Theorem 2.6.1. *Game* $(\tilde{x}, \tilde{p})$ *and strategies* $\boldsymbol{b}$ *are in DSE only if for all* i *and* $\mathbf{v}$,

(i) *(monotonicity)* $x_i(v_i, \mathbf{v}_{-i})$ *is monotone non-decreasing in* v_i, *and*
(ii) *(payment identity)* $p_i(\mathbf{v}) = v_i\, x_i(\mathbf{v}) - \int_0^{v_i} x_i(z, \mathbf{v}_{-i})\, dz + p_i(0, \mathbf{v}_{-i})$,

where $(z, \mathbf{v}_{-i})$ *denotes the valuation profile with the* i*th coordinate replaced with* z. *If the strategy profile is onto then the converse also holds.*

It was important when discussing BNE to explicitly refer to $x_i(v_i)$ and $p_i(v_i)$ as the probability of allocation and the expected payments because a game played by agents with values drawn from a distribution will inherently, from agent i's perspective, have a randomized outcome and payment. In contrast, for games with DSE we can consider outcomes and payments in a non-probabilistic sense. A deterministic game, i.e., one with no internal randomization, will result in deterministic outcomes and payments. For our single-dimensional game where an agent is either served or not served we will have $x_i(\mathbf{v}) \in \{0, 1\}$. This specification along with the monotonicity condition implied by DSE implies that the function $x_i(v_i, \mathbf{v}_{-i})$ is a step function in v_i. The reader can easily verify that the payment required for such a step function is exactly the critical value, i.e., the $\hat{v}_i$ at which $x_i(\cdot, \mathbf{v}_{-i})$ changes from 0 to 1. This gives the following corollary.

Corollary 2.6.2. *A deterministic game* (x, p) *and deterministic strategies* $\boldsymbol{b}$ *are in DSE only if for all* i *and* $\mathbf{v}$,

(i) *(step-function)* $x_i(v_i, \mathbf{v}_{-i})$ *steps from 0 to 1 at some* $\hat{v}_i(\mathbf{v}_{-i})$, *and*
(ii) *(critical value)* $p_i(\mathbf{v}) = \begin{cases} \hat{v}_i(\mathbf{v}_{-i}) & \text{if } x_i(\mathbf{v}) = 1 \\ 0 & \text{otherwise} \end{cases} \; + \; p_i(0, \mathbf{v}_{-i})$.

If the strategy profile is onto then the converse also holds.

Notice that the above theorem deliberately skirts around a subtle tie-breaking issue. Consider the truthtelling DSE of the second-price auction on two agents. What happens when $v_1 = v_2$? One agent should win and pay the other's value. As this results in a utility of zero, from the perspective of utility maximization, both agents are indifferent as to which of them it is. One natural tie-breaking rule is the lexicographical one, i.e., in favor of agent 1 winning. For this rule, agent 1 wins when

$v_1 \in [v_2, \infty)$ and agent 2 wins when $v_2 \in (v_1, \infty)$. The critical values are $\hat{v}_1 = v_2$ and $\hat{v}_2 = v_1$. We will usually prefer the randomized tie-breaking rule because of its symmetry.

2.7 The Method of Revenue Equivalence

We are now ready to make one of the most significant observations in auction theory. Specifically, mechanisms with the same allocation in BNE have the same expected revenue. In fact, not only do they have the same expected revenue, but each agent has the same expected payment. This result is in fact a direct corollary of Theorem 2.5.1. The payment identity implies that an agent's expected payment is precisely determined by the allocation rule and the expected payment of the lowest type, i.e., $p_i(0)$.

Corollary 2.7.1. *For any two mechanisms where 0-valued agents have zero expected payment, if the mechanisms have the same BNE allocation then they have same expected BNE revenue.*

We can now quantitatively compare the second-price and first-price auctions from a revenue standpoint. Consider the case where the agents' values are distributed independently and identically from a continuous distribution. What is the equilibrium outcome of the second-price auction? The agent with the highest valuation wins. What is the equilibrium outcome of the first-price auction? This question requires a little more thought. Since the distributions are identical, it is reasonable to expect that there is a symmetric equilibrium, i.e., one where the bid strategy $b_i = b$ for all i. Furthermore, it is reasonable to expect that the strategies are strictly monotone, i.e., an agent with a higher value will out bid an agent with a lower value. Under these assumptions, the agent with the highest value wins. Of course, in both auctions a 0-valued agent will pay nothing. Therefore, Corollary 2.7.1 implies that the two auctions obtain the same expected revenue.

Corollary 2.7.2. *When agents' values are independent and identically distributed according to a continuous distribution, the second-price and first-price auctions have the same expected revenue.*

As an example of revenue equivalence consider first-price and second-price auctions for selling a single item to two agents with values drawn from $U[0, 1]$. The expected revenue of the second-price auction is $\mathbf{E}[v_{(2)}]$.

In Section 2.3 we saw that it is a Bayes-Nash equilibrium of the first-price auction in this environment for each agent to bid half her value. The expected revenue of first-price auction is therefore $1/2\,\mathbf{E}[v_{(1)}]$. An important fact about uniform random variables is that in expectation they evenly divide the interval they are over, i.e., $\mathbf{E}[v_{(1)}] = 2/3$ and $\mathbf{E}[v_{(2)}] = 1/3$. How do the revenues of these two auctions compare? Their expected revenues are identically $1/3$.

Of course, much more bizarre auctions are governed by revenue equivalence. As an exercise the reader is encourage to verify that the *all-pay auction* – where agents submit bids, the highest bidder wins, and all agents pay their bids – is revenue equivalent to the first- and second-price auctions.

Revenue equivalance gives a method for analyzing equilibria of auctions. The method is based on equating two equations for interim payments in an auction:

(i) *Interim payments from payment format:* Typically the expected payment of an agent i can be easily calculated from the equilibrium allocation rule x_i, equilibrium bid strategy b_i, and the payment format of the auction. For winner-pays-bids mechanisms like the first-price auctions, the agent with value v_i pays her bid $b_i(v_i)$ with probability given by the allocation rule $x_i(v_i)$. Thus, her expected payment is:

$$p_i(v_i) = b_i(v_i)\, x_i(v_i). \tag{2.7.1}$$

For (single-item) second-price auctions, an agent's payment can be calculated by:

$$p_i(v_i) = \mathbf{E}_{v_{-i}}[v_{(2)} \mid v_{(2)} < v_i]\, x_i(v_i). \tag{2.7.2}$$

This calculation follows from the definition of the second-price auction and the law of conditional expectation. The conditional expectation is simply the expected second-highest value, denoted $v_{(2)}$, given that agent i wins. Note that $x_i(v_i) = \mathbf{Pr}_{v_{-i}}\big[v_{(2)} < v_i\big]$.

(ii) *Interim payments from the payment identity:* The payment identity gives the expected payment of the agent with value v_i and allocation rule x_i as:

$$p_i(v_i) = v_i\, x_i(v_i) - \int_0^{v_i} x_i(z)\, dz + p_i(0) \tag{2.7.3}$$

where often $p_i(0) = 0$.

The *method of revenue equivalance* refers to analyses derived from the equation of two formulas for interim payments.

Example 2.7.1. *Revenue equivalence can be used to simplify the calculation of the payment identity (2.7.3). Consider the second-price auction with two agents and uniformly distributed values on* $[0,1]$*. The equation for payments from the definition of the auction (2.7.2), suggests that to identify* p_1 *we musts only identify*

$$x_1(\mathsf{v}_1) = \mathbf{Pr}_{\mathsf{v}_2}[\mathsf{v}_2 \leq \mathsf{v}_1] = \mathsf{v}_1$$

and

$$\mathbf{E}_{\mathsf{v}_2}[\mathsf{v}_2 \mid \mathsf{v}_2 < \mathsf{v}_1] = {}^{\mathsf{v}_1}/_2.$$

Both identities follow because v_2 *is a uniform random variable on* $[0,1]$*; the latter follows additionally because conditioned on* $\mathsf{v}_2 < \mathsf{v}_1$*, value* v_2 *is uniform on* $[0, \mathsf{v}_1]$*. Combining, we have:*

$$p_1(\mathsf{v}_1) = {}^{\mathsf{v}_1^2}/_2.$$

Of course, symmetry across agents implies $p_i(\mathsf{v}_i) = {}^{\mathsf{v}_i^2}/_2$*. We could have, of course, evaluated* $p_i(\mathsf{v}_i)$ *directly from the payment identity (2.7.3) with* $x_i(\mathsf{v}_i) = \mathsf{v}_i$ *as* $\mathsf{v}_i^2 - \int_0^{\mathsf{v}_i} \mathsf{z}\, \mathsf{dz} = {}^{\mathsf{v}_i^2}/_2$*.*

In the next two sections we use the revenue equivalance method to solve for the Bayes-Nash equilibrium of the first-price auction and to prove the uniqueness of this Bayes-Nash equilibrium.

2.8 Solving for Bayes-Nash Equilibrium

Solving for Bayes-Nash equilibrium is an important step of the analysis of auctions. In this section we describe an elegant technique for calculating BNE strategies in symmetric environments using the method of revenue equivalence.

To solve for the Bayes-Nash equilibrium strategies in a given target mechanism $(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{p}})$ under value distribution $\boldsymbol{F}$:

(i) guess the interim allocation rule $\boldsymbol{x}$ of the equilibrium.
(ii) calculate the interim payment according to the payment format.
(iii) calculate the interim payment according to the payment identity.
(iv) solve for the bid strategy of the target mechanism by equating the interim payment equations; and

(v) verify that the solved bid strategy is an equilibrium by showing:
   (a) the solved strategy profile composed with the mechanism in expectation over the distribution gives the guessed interim allocation rule, and
   (b) (if the solved strategy function is not onto) the actions out of the range of the strategy function are dominated by actions in its range.

Example 2.8.1. *Consider the first-price auction with two agents and uniformly distributed values on* $[0, 1]$.

(i) *Guess that the highest-valued agent wins; thus,* $x_i(\mathsf{v}_i) = \mathsf{v}_i$.
(ii) *By the payment format equation* (2.7.1)*,* $p_i(\mathsf{v}_i) = b_i(\mathsf{v}_i)\, x_i(\mathsf{v}_i)$.
(iii) *By the payment identity* (2.7.3)*,* $p_i(\mathsf{v}_i) = \mathsf{v}_i^2 - \int_0^{\mathsf{v}_i} \mathsf{z}\, \mathsf{dz} = \mathsf{v}_i^2/2$.
(iv) *Equating and solving,* $b_i(\mathsf{v}_i) = p_i(\mathsf{v}_i)/x_i(\mathsf{v}_i) = \mathsf{v}_i/2$.
(v) *Verifying,*
   (a) *strategy* b_i *is monotonic and increasing so the equilibrium allocation rule is highest-value-wins; and*
   (b) *strategy* b_i *maps the domain* $[0, 1]$ *onto the range* $[0, 1/2]$ *while the action space of the first price auction is* $[0, \infty)$*, i.e., it is not onto. We check that bid* $1/2$ *dominates all bids* $\mathsf{b}_i > 1/2$*. Indeed, with all bids* $\mathsf{b}_i \geq 1/2$ *under strategies* $\boldsymbol{b}$ *the agent wins with certainty; but with the bid of* $1/2$ *the payment (which equals the bid) is the lowest. Thus,* $\mathsf{b}_i = 1/2$ *which is in the range of* b_i *dominates all bids outside the range.*

In Example 2.8.1 were the two equations for payments are equated, we could have observed that under the guess that the Bayes-Nash equilibrium of the first-price auction is highest-value-wins; it is revenue equivalent to the second-price auction and, thus, the payment-format equations (2.7.1) and (2.7.2) for winner-pays-bid and second-price auctions could be directly equated as

$$b_i(\mathsf{v}_i)\, x_i(\mathsf{v}_i) = p_i^{\mathrm{FP}}(\mathsf{v}_i) = p_i^{\mathrm{SP}}(\mathsf{v}_i) = \mathbf{E}_{\mathsf{v}_{-i}}[\mathsf{v}_{(2)} \mid \mathsf{v}_{(2)} < \mathsf{v}_i]\, x_i(\mathsf{v}_i)$$

Solving for $b_i(\mathsf{v}_i)$ by dividing through by $x_i(\mathsf{v}_i)$ gives

$$b_i(\mathsf{v}_i) = \mathbf{E}_{\mathsf{v}_{-i}}[\mathsf{v}_{(2)} \mid \mathsf{v}_{(2)} < \mathsf{v}_i] = \mathsf{v}_i/2$$

with the final equality calculated for the case of two agents with uniform values on $[0, 1]$. This calculation is slightly simpler than the previous calculation as the functional form of the allocation rule does not need to be calculated.

We can also work through the above framework for the *all-pay* auction where the agents submit bids, the highest bid wins, but all agents pay their bid. For all-pay mechanisms, the interim payments from the payment format are

$$p_i(\mathsf{v}_i) = b_i(\mathsf{v}_i). \tag{2.8.1}$$

We guess and verify that the all-pay auction has Bayes-Nash allocation of highest-value-wins. For example, for two agents with uniform random values on $[0, 1]$, the payment identity (2.7.3) is $p_i(\mathsf{v}_i) = \mathsf{v}_i^2/2$; thus, equating with the interim payment from the payment format we have $b_i(\mathsf{v}_i) = \mathsf{v}_i^2/2$.

Remember, of course, that the equilibrium strategies solved for above are for single-item auctions and two agents with values uniform on $[0, 1]$. For different distributions or numbers of agents the equilibrium strategies will generally be different.

We conclude by observing that if we fail to exhibit a Bayes-Nash equilibrium via this approach then our original guess is contracted and there is no equilibrium of the given mechanism that corresponds to the guess. Conversely, if the approach succeeds then the equilibrium found is the only equilibrium consistent with the guess. As an example, we can conclude the following for first-price auctions.[2]

Proposition 2.8.1. *When agents' values are independent and identically distributed from a continuous distribution, the first-price auction has a unique Bayes-Nash equilibrium for which the highest-valued agent always wins.*

2.9 Uniqueness of Equilibria

As equilibrium attempts to make a prediction of what will happen in a game or mechanism, the uniqueness of equilibrium is important. If there are multiple equilibria then the prediction is to a set of outcomes not a single outcome. In terms of mechanism design, some of these outcomes could be good and some could be bad. There are also questions of how the players coordinate on an equilibrium.

As an example, in the second-price auction for two agents with values

[2] In the next section we will strengthen Proposition 2.8.1 and show that for the first-price auction (with independent, identical, and continuous distributions) there are no equilibria where the highest-valued agent does not win. Thus, the equilibrium solved for is the unique Bayes-Nash equilibrium.

uniformly distributed on $[0, 1]$ there is the dominant strategy equilibrium where agents truthfully report their values. This outcome is good from the perspective of surplus in that the item is awarded to the highest-valued agent. There are, however, other Bayes-Nash equilibria. For instance, it is also a BNE for agent 1 (Alice) to bid one and agent 2 (Bob) to bid zero (regardless of their values). Alice is happy to win and pay zero (Bob's bid); Bob, with any value $v_2 \leq 1$, is at least as happy to lose and pay zero versus winning and paying one (Alice's bid). Via examples like this the surplus of the worst BNE in the second-price auction can be arbitrarily worse than the surplus of the best BNE (Exercise 2.9).

In contrast, the first-price auction for independent and identical prior distributions does not suffer from multiplicity of Bayes-Nash equilibria. Specifically, the method described in the previous section for solving for the symmetric equilibrium in the first-price auction gives a unique BNE. We describe this result as two parts. First, we exclude the possibility of multiple symmetric equilibria. Second, we exclude the existence of asymmetric equilibria.

Lemma 2.9.1. *For agents with values drawn independently and identically from a continuous distribution, the first-price auction admits exactly one symmetric Bayes-Nash equilibrium.*

Proof. Consider a symmetric strategy profile $\boldsymbol{b} = (b, \ldots, b)$. First, the common strategy $b(\cdot)$ must be non-decreasing (otherwise BNE is contradicted by Theorem 2.5.1).

Second, if the strategy is only weakly increasing then there is a point mass at some bid b in the bid distribution. Symmetry with respect to this strategy implies that all agents will make a bid equal to this point mass with some measurable (i.e., strictly positive) probability. All but one of these bidders must lose (perhaps via random tie-breaking). Winning, however, must be strictly preferred to losing for some of the values in the interval (as an agent with value v is only indifferent to winning or losing when $\mathsf{v} = \mathsf{b}$). Such a losing agent has a deviation of bidding $\mathsf{b} + \epsilon$, and for ϵ approaching zero this deviation is strictly better than bidding b. This deviation is a contradiction to the existence of such a weakly increasing equilibrium.

Finally, for a strictly increasing strategy b the highest-valued agent must always win; therefore, Proposition 2.8.1 implies that there is only one such equilibrium. □

Next, we disprove the existence of asymmetric equilibria in the first-price auction. The theorem is straightforward from the following lemma which uses revenue equivalance to contradict the existence of asymmetric equilibria for two-agents first-price auctions with a random reserve bid.

Lemma 2.9.2. *For $n = 2$ agents with values drawn independently and identically from a continuous distribution F, the first-price auction with an unknown random reserve from known distribution G admits no asymmetric Bayes-Nash equilibrium.*

Theorem 2.9.3. *For $n \geq 2$ agents with values drawn independently and identically from a continuous distribution F, the first-price auction admits a unique Bayes-Nash equilibrium and it is symmetric.*

Proof. By Lemma 2.9.1 there is exactly one symmetric Bayes-Nash equilibrium of an n-agent first-price auction. If there is an asymmetric equilibrium there must be two agents whose strategies are distinct. We can view the n-agent first-price auction in BNE, from the perspective of this pair of agents, as a two-agent first-price auction with a random reserve drawn from the distribution of BNE bids of the other $n - 2$ agents. Lemma 2.9.2 then contradicts the distinctness of these two strategies. □

As usual, we will refer to agents 1 and 2 by the names Alice and Bob. The proof of Lemma 2.9.2 will follow by contradiction. Assume that, for values in a given interval, Alice's bid exceeds Bob's. A first lemma will combine with the payment identity (2.7.3) to show that a high-bidding Alice has strictly higher utility than a low-bidding Bob. A second lemma is based on the payment format (2.7.1) and shows that a low-bidding Bob has at least the utility of a high-bidding Alice. These lemmas are then combined in a subsequent proof by contradiction to obtain Lemma 2.9.2.

Lemma 2.9.4. *In a two-agent first-price auction with random reserve with strictly-increasing continuous strategies, the strategies satisfy $b_1(\mathsf{v}) > b_2(\mathsf{v})$ at some v if any only if the interim allocation probabilities satisfy $x_1(\mathsf{v}) > x_2(\mathsf{v})$.*

Proof. See Figure 2.3 for a graphical representation of the following argument. Since the strategies are continuous and strictly increasing, the inverses of the strategies are well defined. Calculate Alice's interim allocation probability x_1 at value v, for Bob's value $\mathsf{v}_2 \sim F$ and reserve bid

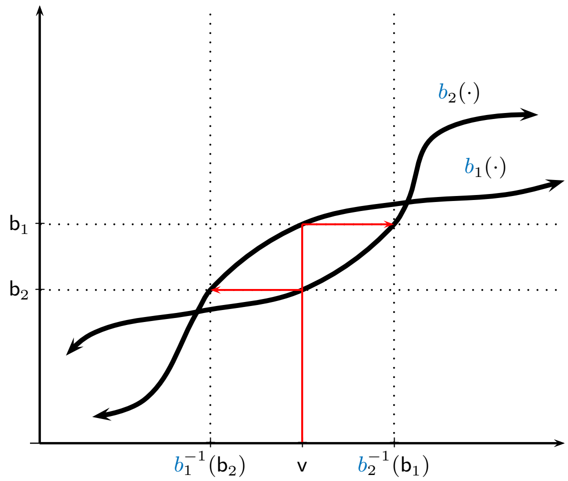


Figure 2.3. Graphical depiction of the first claim in the proof of Lemma 2.9.2 with $\mathsf{b}_i = b_i(\mathsf{v})$. Clearly, $b_2^{-1}(\mathsf{b}_1) > b_1^{-1}(\mathsf{b}_2)$. Strict monotonicity of the distribution function $F(\cdot)$ then implies that $F(b_2^{-1}(\mathsf{b}_1)) > F(b_1^{-1}(\mathsf{b}_2))$.

$\hat{\mathsf{b}} \sim G$, as:

$$\begin{aligned} x_1(\mathsf{v}) &= \mathbf{Pr}[b_1(\mathsf{v}) > b_2(\mathsf{v}_2) \ \&\ b_1(\mathsf{v}) > \hat{\mathsf{b}}] \\ &= \mathbf{Pr}[b_2^{-1}(b_1(\mathsf{v})) > \mathsf{v}_2 \ \&\ b_1(\mathsf{v}) > \hat{\mathsf{b}}] \\ &= F(b_2^{-1}(b_1(\mathsf{v}))) \cdot G(b_1(\mathsf{v})). \end{aligned}$$

Likewise, Bob's interim allocation probability is

$$x_2(\mathsf{v}) = F(b_1^{-1}(b_2(\mathsf{v}))) \cdot G(b_2(\mathsf{v})).$$

For $b_1(\mathsf{v}) \geq b_2(\mathsf{v})$ then the last term in the allocation probabilities satisfies $G(b_1(\mathsf{v})) \geq G(b_2(\mathsf{v}))$ (as the distribution function $G(\cdot)$ is non-decreasing). Similarly, strict monotonicity of the strategy functions and distribution function imply that for $b_1(\mathsf{v}) \geq b_2(\mathsf{v})$ the first term in the allocation probabilities satisfies $F(b_2^{-1}(b_1(\mathsf{v}))) \geq F(b_1^{-1}(b_2(\mathsf{v})))$; moreover, either both inequalities are strict or both are equality. □

Lemma 2.9.5. *In a two-agent first-price auction with random reserve with strictly-increasing continuous strategies in Bayes-Nash equilibrium, for maximal interval* $I = (\mathsf{v}^\dagger, \mathsf{v}^\ddagger)$ *where agent 1 strictly out bids agent 2,* *i.e., with* $b_1(\mathsf{v}) > b_2(\mathsf{v})$ *for* $\mathsf{v} \in I$*, the low-bidding agent 2 obtains (weakly) at most the utility of high-bidding agent 1 at the low endpoint* $\mathsf{v}^\dagger$ *and (weakly) at least the utility of the high-bidding agent 1 at the high endpoint* $\mathsf{v}^\ddagger$*, i.e.,* $u_2(\mathsf{v}^\dagger) \leq u_1(\mathsf{v}^\dagger)$ *and* $u_2(\mathsf{v}^\ddagger) \geq u_1(\mathsf{v}^\ddagger)$*.*

Proof. We argue the claim for the upper endpoint of the interval $\mathsf{v}^\ddagger$; the case of lower endpoint $\mathsf{v}^\dagger$ is analogous.

The key to this claim is that there are not higher values $\mathsf{v} > \mathsf{v}^\ddagger$ where $b_2(\mathsf{v}) < b_1(\mathsf{v}^\ddagger)$. This is either because $b_1(\mathsf{v}^\ddagger) = b_2(\mathsf{v}^\ddagger)$ (and the strategies are monotonically increasing) or because $\mathsf{v}^\ddagger$ is the maximum value in the support of the value distribution F. In the first case, by Lemma 2.9.4, $x_1(\mathsf{v}^\ddagger) = x_2(\mathsf{v}^\ddagger)$ so by the payment format (2.7.1) the agents' utilities are equal, i.e., $(\mathsf{v}^\ddagger - b_1(\mathsf{v}^\ddagger))\, x_1(\mathsf{v}^\ddagger) = (\mathsf{v}^\ddagger - b_2(\mathsf{v}^\ddagger))\, x_2(\mathsf{v}^\ddagger)$

In the second case, Bob with value $\mathsf{v}^\ddagger$ could deviate by copying Alice's strategy at value $\mathsf{v}^\ddagger$, i.e., bidding $b_1(\mathsf{v}^\ddagger)$, and obtain the same allocation probability as Alice would with that bid (argued below). By the payment format (2.7.1), then, such a deviation would give Bob (with value $\mathsf{v}^\ddagger$) the same utility as Alice (with value $\mathsf{v}^\ddagger$). Existence of such a deviation gives a lower bound on Bob's utility under the original Bayes-Nash equilibrium strategies. Thus, the statement of the lemma holds.

We now show that Bob with value $\mathsf{v}^\ddagger$ at the upper endpoint of the support of the distribution can deviate to bid identically to Alice with this value, i.e., to bidding $b_1(\mathsf{v}^\ddagger)$, and guarantee the same allocation probability as Alice with this bid. Notice that there is a measure zero probability Alice has value $\mathsf{v}^\ddagger$ and consequently, since the strategies are continuous and strictly increasing, there is a measure zero probability that Alice bids $b_1(\mathsf{v}^\ddagger)$ or above. Thus, even with this deviation, there is measure zero probability that either Alice's or Bob's bids are $b_1(\mathsf{v}^\ddagger)$ or higher. With value $\mathsf{v}^\ddagger$, the interim allocation probability of both Alice and Bob with the deviating strategy is $G(\mathsf{v}^\ddagger)$, the probability that this bid beats the random reserve, as the other agent (Bob with the deviating strategy or Alice, respectively) is out bid with probability one. □

Proof of Lemma 2.9.2. We will prove this lemma for the special case of strictly-increasing and continuous strategies (for the general argument, see Exercise 2.13).

Suppose the bid strategies of Alice and Bob are distinct at some value v and without loss of generality Alice's bid exceeds Bob's bid, i.e., $b_1(\mathsf{v}) > b_2(\mathsf{v})$. By the continuity of the bid strategies, there is a non-trivial maximal interval $I = (\mathsf{v}^\dagger, \mathsf{v}^\ddagger)$ that contains value v and for which $b_1(\mathsf{v}) > b_2(\mathsf{v})$ for all $\mathsf{v} \in I$.

The lemma follows from writing the difference in utilities of each of Alice and Bob with values $\mathsf{v}^\dagger$ and $\mathsf{v}^\ddagger$ using the payment format (2.7.1) and the payment identity (2.7.3) and deriving a contradiction. By Lemma 2.9.5, which follows from the payment format (2.7.1), this difference is (weakly) greater for Bob than Alice (specifically, relative to Alice's utility, Bob's utility is no higher at $\mathsf{v}^\dagger$ and no lower at $\mathsf{v}^\ddagger$).

$$u_1(\mathsf{v}^\ddagger) - u_1(\mathsf{v}^\dagger) \leq u_2(\mathsf{v}^\ddagger) - u_2(\mathsf{v}^\dagger) \tag{2.9.1}$$

However, by Lemma 2.9.4, Alice has a strictly higher allocation rule on interval $I = (\mathsf{v}^\dagger, highval)$ and therefore, by the payment identity (2.7.3), strictly higher change in utility. Recall, $u_i(\mathsf{v}_i) = \mathsf{v}_i\, x_i(\mathsf{v}_i) - p_i(\mathsf{v}_i) = \int_0^{\mathsf{v}_i} x(\mathsf{z})\, \mathrm{d}\mathsf{z}$.

$$u_1(\mathsf{v}^\ddagger) - u_1(\mathsf{v}^\dagger) = \int_{\mathsf{v}^\dagger}^{\mathsf{v}^\ddagger} x_1(\mathsf{z})\, \mathrm{d}\mathsf{z} > \int_{\mathsf{v}^\dagger}^{\mathsf{v}^\ddagger} x_2(\mathsf{z})\, \mathrm{d}\mathsf{z} = u_2(\mathsf{v}^\ddagger) - u_2(\mathsf{v}^\dagger) \tag{2.9.2}$$

Equations (2.9.1) and (2.9.2) combine to give a contradiction. □

2.10 The Revelation Principle

We are interested in designing mechanisms and, while the characterization of Bayes-Nash equilibrium is elegant, solving for equilibrium is still generally quite challenging. The final piece of the puzzle, and the one that has enabled much of modern mechanism design is the *revelation principle*. The revelation principle states, informally, that if we are searching among mechanisms for one with a desirable equilibrium we may restrict our search to single-round, sealed-bid mechanisms in which truthtelling is an equilibrium.

Definition 2.10.1. *A* direct revelation *mechanism is single-round, sealed bid, and has action space equal to the type space, (i.e., an agent can bid any type she might have)*

Definition 2.10.2. *A direct revelation mechanism is* Bayesian incentive compatible *(BIC) if truthtelling is a Bayes-Nash equilibrium.*

Definition 2.10.3. *A direct revelation mechanism is* dominant strategy incentive compatible *(DSIC) if truthtelling is a dominant strategy equilibrium.*

Theorem 2.10.1. *Any mechanism $(\tilde{x}, \tilde{p})$ with good BNE (resp. DSE) can be converted into a BIC (resp. DSIC) mechanism $(x^\dagger, p^\dagger)$ with the same BNE (resp. DSE) outcome.*

Proof. We will prove the BNE variant of the theorem. Let b, F, and $(\tilde{x}, \tilde{p})$ be in BNE. Define the single-round, sealed-bid *revelation mechanism* $(x^\dagger, p^\dagger)$ as follows:

(i) Accept sealed bids $\mathbf{b}^\dagger$.
(ii) Simulate $b(\mathbf{b}^\dagger)$ in $(\tilde{x}, \tilde{p})$.
(iii) Output the outcome of the simulation.

We now claim that b being a BNE of $(\tilde{x}, \tilde{p})$ implies truthtelling is a BNE of $(x^\dagger, p^\dagger)$ (for distribution F). Let $b^\dagger$ denote the truthtelling strategy. In $(x^\dagger, p^\dagger)$, consider agent i and suppose all other agents are truthtelling, i.e., $\mathbf{b}^\dagger_{-i} = \mathbf{v}_{-i}$. Thus, the actions of the other players in $(\tilde{x}, \tilde{p})$ are distributed as $b_{-i}(b^\dagger_{-i}(\mathbf{v}_{-i})) = b_{-i}(\mathbf{v}_{-i})$ for $\mathbf{v}_{-i} \sim F_{-i}$. Of course, in $(\tilde{x}, \tilde{p})$ if other players are playing $b_{-i}(\mathbf{v}_{-i})$ then since b is a BNE, i's best response is to play $b_i(\mathsf{v}_i)$ as well. Agent i can play this action in the simulation of $(\tilde{x}, \tilde{p})$ by playing the truthtelling strategy $b^\dagger_i(\mathsf{v}_i) = \mathsf{v}_i$ in $(x^\dagger, p^\dagger)$. □

Notice that we already, in Chapter 1, saw the revelation principle in action. The second-price auction is the revelation principle applied to the ascending-price auction.

Because of the revelation principle, for many of the mechanism design problems we consider, we will look first for Bayesian or dominant-strategy incentive compatible mechanisms. The revelation principle guarantees that, in our search for optimal BNE mechanisms, it suffices to search only those that are BIC (and likewise for DSE and DSIC). The following are corollaries of our BNE and DSE characterization theorems.

When discussing incentive compatible mechanisms, i.e., ones where truthtelling is an equilibrium, we will define the allocation rules and payment rules in value space directly, denoted (x, p), and forgo defining the mechanism in bid space as $(\tilde{x}, \tilde{p})$ and discussing strategies b. In the truthtelling equilibrium, strategies b is the profile of identify functions and $(\tilde{x}, \tilde{p}) = (x, p)$. As previously defined for values drawn from product distribution F, the interim allocation and payment rules of agent i are $x_i(\mathsf{v}_i) = \mathbf{E}_{\mathbf{v}_{-i}}[x_i(\mathbf{v})]$ and $p_i(\mathsf{v}_i) = \mathbf{E}_{\mathbf{v}_{-i}}[p_i(\mathbf{v})]$.

We proceed by characterizing incentive compatibility, i.e., specializing the Bayes-Nash equilibrium characterization (Theorem 2.5.1) to the special case of mechanisms with a truthtelling equilibrium. Note that the truthtelling strategy profile in a direct-revelation mechanism is onto and, thus, the statement of the characterization of Bayesian incentive compatibility can be simplified relative to the characterization of Bayes-Nash equilibrium.

Corollary 2.10.2. *A direct mechanism* $(\boldsymbol{x}, \boldsymbol{p})$ *is BIC for distribution* $\boldsymbol{F}$ *if and only if for all agents* i*,*

(i) *(monotonicity)* $x_i(\mathsf{v}_i)$ *is monotone non-decreasing, and*
(ii) *(payment identity)* $p_i(\mathsf{v}_i) = \mathsf{v}_i\, x_i(\mathsf{v}_i) - \int_0^{\mathsf{v}_i} x_i(\mathsf{z})\,\mathrm{d}\mathsf{z} + p_i(0)$.

Corollary 2.10.3. *A direct mechanism* $(\boldsymbol{x}, \boldsymbol{p})$ *is DSIC if and only if for all agents* i *and valuation profiles* $\mathbf{v}$*,*

(i) *(monotonicity)* $x_i(\mathsf{v}_i, \mathbf{v}_{-i})$ *is monotone non-decreasing in* v_i*, and*
(ii) *(payment identity)* $p_i(\mathsf{v}_i, \mathbf{v}_{-i}) = \mathsf{v}_i\, x_i(\mathsf{v}_i, \mathbf{v}_{-i}) - \int_0^{\mathsf{v}_i} x_i(\mathsf{z}, \mathbf{v}_{-i})\,\mathrm{d}\mathsf{z} + p_i(0, \mathbf{v}_{-i})$.

Corollary 2.10.4. *A direct, deterministic mechanism* $(\boldsymbol{x}, \boldsymbol{p})$ *is DSIC if and only if for all agents* i *and valuation profiles* $\mathbf{v}$*,*

(i) *(step-function)* $x_i(\mathsf{v}_i, \mathbf{v}_{-i})$ *steps from 0 to 1 at some* $\hat{v}_i(\mathbf{v}_{-i})$*, and*
(ii) *(critical value)* $p_i(\mathsf{v}_i, \mathbf{v}_{-i}) = \begin{cases} \hat{v}_i(\mathbf{v}_{-i}) & \text{if } x_i(\mathsf{v}_i, \mathbf{v}_{-i}) = 1 \\ 0 & \text{otherwise} \end{cases} + p_i(0, \mathbf{v}_{-i})$.

When we construct mechanisms we will use the “if” directions of these theorems. When discussing incentive compatible mechanisms we will assume that agents follow their equilibrium strategies and, therefore, each agent’s bid is equal to her valuation.

Between DSIC and BIC clearly DSIC is a stronger incentive constraint and we should prefer it over BIC if possible. Importantly, DSIC requires fewer assumptions on the agents. For a DSIC mechanisms, each agent must only know her own value; while for a BIC mechanism, each agent must also know the distribution over other agent values. Unfortunately, there will be some environments where we derive BIC mechanisms where no analogous DSIC mechanism is known.

The revelation principle fails to hold in some environments of interest. We will take special care to point these out. Two such environments, for instance, are where agents only learn their values over time, or where the designer does not know the prior distribution (and hence cannot simulate the agent strategies).

Exercises

2.1 Find a symmetric mixed strategy equilibrium in the chicken game described in Section 2.1. I.e., find a probability ρ such that if James Dean stays with probability ρ and swerves with probability $1-\rho$ then Buzz is happy to do the same.

2.2 Give a characterization of Bayes-Nash equilibrium for discrete single-dimensional type spaces for agents with linear utility. Assume that $\mathscr{T} = \{\mathsf{v}^0, \ldots, \mathsf{v}^N\}$ with the probability that an agent's value is $\mathsf{v} \in \mathscr{T}$ given by probability mass function $f(\mathsf{v})$. Assume $\mathsf{v}^0 = 0$. You will not get a payment identity; instead characterize for any BNE allocation rule, the maximum payments.

(a) Give a characterization for the special case where the values are uniform, i.e., $\mathsf{v}^j = j$ for all j.

(b) Give a characterization for the special case where the probabilities are uniform, i.e., $f(\mathsf{v}^j) = 1/N$ for all j.

(c) Give a characterization for the general case.

(Hint: You should end up with a very similar characterization to that for continuous type spaces.)

2.3 In Section 2.3 we characterized outcomes and payments for BNE in single-dimensional games. This characterization explains what happens when agents behave strategically.

Suppose instead of strategic interaction, we care about fairness. Consider a valuation profile, $\mathbf{v} = (\mathsf{v}_1, \ldots, \mathsf{v}_n)$, an allocation vector, $\mathbf{x} = (\mathsf{x}_1, \ldots, \mathsf{x}_n)$, and payments, $\mathbf{p} = (\mathsf{p}_1, \ldots, \mathsf{p}_n)$. Here x_i is the probability that i is served and p_i is the expected payment of i regardless of whether i is served or not.

Allocation $\mathbf{x}$ and payments $\mathbf{p}$ are *envy-free* for valuation profile $\mathbf{v}$ if no agent wants to unilaterally swap allocation and payment with another agent. I.e., for all i and j,

$$\mathsf{v}_i\mathsf{x}_i - \mathsf{p}_i \geq \mathsf{v}_i\mathsf{x}_j - \mathsf{p}_j.$$

Characterize envy-free allocations and payments (and prove your characterization correct). Unlike the BNE characterization, your characterization of payments will not be unique. Instead, characterize the minimum payments that are envy-free. Draw a diagram illustrating your payment characterization. (Hint: You should end up with a very similar characterization to that of BNE.)

2.4 AdWords is a Google Inc. product in which the company sells the placement of advertisements along side the search results on its search results page. Consider the following *position auction* environment which provides a simplified model of AdWords. There are m advertisement slots that appear along side search results and n advertisers. Advertiser i has value v_i for a click. Slot j has *click-through rate* w_j, meaning, if an advertiser is assigned slot j the advertiser will receive a click with probability w_j. Each advertiser can be assigned at most one slot and each slot can be assigned at most one advertiser. If a slot is left empty, all subsequent slots must be left empty, i.e., slots cannot be skipped. Assume that the slots are ordered from highest click-through rate to lowest, i.e., $\mathsf{w}_j \geq \mathsf{w}_{j+1}$ for all j.

(a) Find the envy-free (see Exercise 2.3) outcome and payments with the maximum surplus. Give a description and formula for the envy-free outcome and payments for each advertiser. (Feel free to specify your payment formula with a comprehensive picture.)

(b) In the real AdWords problem, advertisers only pay if they receive a click, whereas the payments calculated, i.e., $\mathbf{p}$, are in expected over all outcomes, click or no click. If we are going to charge advertisers only if they are clicked on, give a formula for calculating these payments $\mathbf{p}^{\dagger}$ from $\mathbf{p}$.

(c) The real AdWords problem is solved by auction. Design an auction that maximizes the surplus in dominant strategy equilibrium. Give a formula for the payment rule of your auction (again, a comprehensive picture is fine). Compare your DSE payment rule to the envy-free payment rule. Draw some informal conclusions.

2.5 Consider the first-price auction for selling a single item to two agents whose values are independent but not identical. In each of the settings below prove or disprove the claim that there is a Bayes-Nash equilibrium wherein the item is always allocated to the agent with the highest value.

(a) Agent 1 has value $U[0, 1]$ and agent 2 has value $U[0, 1/2]$.

(b) Agent 1 has value $U[0, 1]$ and agent 2 has value $U[1/2, 1]$.

2.6 Refine the Bayes-Nash equilibrium characterization (Theorem 2.5.1) to the special case of winner-pays-bid mechanisms that (a) collect bids, (b) map the bid profile to a partition of the agents to winners and non-winners, and (c) charge all winners their bids. Hint: use revenue equivalence.

2.7 Consider a first-price position auction (see Exercise 2.4) with n agents, n positions, and decreasing weights $\mathbf{w} = (\mathsf{w}_1, \mathsf{w}_2)$. In such an auction agents are assigned to positions in the order of their bids. The agent assigned to position i is is served with probability w_i and she pays her bid when served. Use revenue equivalence to solve for symmetric Bayes-Nash equilibrium strategies $\boldsymbol{b}$ when the values of the agents are drawn independent and identically from $U[0, 1]$.

2.8 Consider a first-price position auction (see Exercise 2.4) with n agents, n positions, and position weights $\mathbf{w}$ defined by $\mathsf{w}_i = 1 - {}^{i-1}/_{n-1}$ for $i \in \{1, \ldots, n\}$. In such an auction agents are assigned to positions in the order of their bids. The agent assigned to position i is is served with probability w_i and she pays her bid when served. Use revenue equivalence to prove that the symmetric strategy profile $\boldsymbol{b} = (b, \ldots, b)$ with $b(\mathsf{v}) = {}^{\mathsf{v}}/_{2}$ is a Bayes-Nash equilibrium when the values of the agents are drawn independent and identically from $U[0, 1]$.

2.9 Prove that in a two-agent second-price auction for a single-item, that the best Bayes-Nash equilibrium can have a surplus (i.e., the expected value of the winner) that is arbitrarily larger than the worst Bayes-Nash equilibrium. (Hint: Show that for any fixed β that there is a value distribution $\boldsymbol{F}$ and two BNE where the surplus in one BNE is strictly larger than a β fraction of the surplus of the other BNE.)

2.10 Show that with independent, identical, and continuously distributed values, the two-agent all-pay auction (where agents bid, the highest-bidder wins, and all agents pay their bids) admits exactly one strictly continuous Bayes-Nash equilibrium.

2.11 Show that with independent, identical, and continuously distributed values, the two-agent first-price position auction (cf. Exercise 2.4: where agents bid, the highest bidder is served with given probability w_1, the second-highest bidder is served with given probability $\mathsf{w}_2 \leq \mathsf{w}_1$, and all agents pay their bids when they are served) admits exactly one strictly continuous Bayes-Nash equilibrium.

2.12 Consider the following auction with first-price payment semantics. Agents bid, any agent whose bid is (weakly) higher than all other bids wins, all winners are charged their bids. Notice that in the case of a tie in the highest bid, all of the tied agents win. Prove that there are multiple Bayes-Nash equilibria when agents have values that are independently, identically, and continuously distributed.

2.13 Prove Lemma 2.9.2: For two agents with values drawn independently and identically from a continuous distribution F with support $[0, 1]$, the first-price auction with an unknown random reserve from known distribution G admits no asymmetric Bayes-Nash equilibrium. I.e., remove the assumption of strictly-increasing and continuous strategies from the proof given in the text.

Chapter Notes

The formulation of Bayesian games is due to Harsanyi (1967). The characterization of Bayes-Nash equilibrium, revenue equivalence, and the revelation principle come from Myerson (1981). Parts of the BNE characterization proof presented here come from Archer and Tardos (2001). Amann and Leininger (1996), Bajari (2001), Maskin and Riley (2003), and Lebrun (2006) studied the uniqueness of equilibrium in the first-price and all-pay auctions. The revenue-equivalence-based uniqueness proof presented here is from Chawla and Hartline (2013). The refinement of the BNE characterization for winner-pays-bid mechanisms is due to Hartline et al. (2019).

The position auction was formulated by Edelman et al. (2007) and Varian (2007); see Jansen and Mullen (2008) for the history of auctions for advertisements on search engines. Envy freedom has been considered in algorithmic (e.g., Guruswami et al., 2005) and economic (e.g., Jackson and Kremer, 2007) contexts. Hartline and Yan (2011) characterized envy-free outcomes for single-dimensional agents.