Chapter 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 the 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. Readers should look elsewhere for formal justifications.

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.

Emphasis is placed on demonstrating the central theories of equilibrium and not on providing the most comprehensive or general results. For that 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 who have jointly committed a crime, are being interrogated in separate quarters. Unfortunately, the interrogators are unable to prosecute either prisoner without a confession. Each prisoner is offered the following deal: If she confesses and their accomplice does not, she will be released and her accomplice will serve the full sentence of ten years in prison. If they both confess, they will

> share the sentence and serve five years each. If neither confesses, they will both be prosecuted for a minimal offense and each receive a year of prison.

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 prisoners' dilemma. Suppose the row player is silent. What should the column player do? Remaining silent as well results in one year of prison while confessing results in immediate release. Clearly confessing is better. Now suppose that the row player confesses. Now what should the column player 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 the row player does, the column player is better of by confessing. The prisoners dilemma is hardly a dilemma at all: the *strategy profile* (confess, confess) is a *dominant strategy equilibrium*.

Definition 2.1. *A* dominant strategy equilibrium *(DSE) in a complete information game is a strategy profile in which each player's strategy is as least as good as all other strategies regardless of the strategies of all other players.*

DSE 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

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

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 1, 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.*

Definition 2.2. *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.*

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 $\mathbf{t} = (t_1, \ldots, t_n)$.

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 $s_i(\cdot)$ the strategy of agent i and $\mathbf{s} = (s_1, \ldots, s_n)$ 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., $t_i = v_i$. As we described, strategies in the English auction were $s_i(v_i) =$ "drop out when the price exceeds v_i" and strategies in the second-price auction were $s_i(v_i) =$ "bid $b_i = v_i$." We refer to this latter strategy as *truthtelling.* Both of these strategy profiles are in *dominant strategy equilibrium* for their respective games.

Definition 2.3. *A* dominant strategy equilibrium *(DSE) is a strategy profile* $\mathbf{s}$ *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* $s_i(t_i)$*.*

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. Meaning, 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 additionally on i's beliefs on the types of the other agents. The standard economic treatment addresses this by assuming a common prior.

Definition 2.4. *Under the* common prior assumption*, the agent types* $\mathbf{t}$ *are drawn at random from a* prior distribution $\mathbf{F}$ *(a joint probability distribution over type profiles) and this prior distribution is* common knowledge.

The distribution $\mathbf{F}$ 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 $\mathbf{F}_{-i}\big|_{t_i}$. Of course, when the distribution of types is independent, i.e., $\mathbf{F}$ is the *product distribution* $F_1 \times \cdots \times F_n$, then $\mathbf{F}_{-i}\big|_{t_i} = \mathbf{F}_{-i}$.

Notice that a prior $\mathbf{F}$ and strategies $\mathbf{s}$ 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 2.5. *A* Bayes-Nash equilibrium (BNE) *for a game G and common prior* $\mathbf{F}$ *is a strategy profile* $\mathbf{s}$ *such that for all i and t_i, $s_i(t_i)$ is a best response when other agents play* $\mathbf{s}_{-i}(\mathbf{t}_{-i})$ *when* $\mathbf{t}_{-i} \sim \mathbf{F}_{-i}\big|_{t_i}$.

To illustrate BNE, consider using the first-price auction to sell a single item to one of two agents, each with valuation drawn independently and identically from $U[0,1]$, i.e., $\mathbf{F} = F \times F$ 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 $s_i(z) = z/2$ for $i \in \{1,2\}$. Second, we calculate agent 1's expected utility with value v_1 and bid b_1 under the standard assumption that the agent's utility u_i is her value less her payment (when she wins).

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

Calculate $\mathbf{Pr}[1 \text{ wins}] = \mathbf{Pr}[b_2 \leq b_1] = \mathbf{Pr}[v_2/2 \leq b_1] = \mathbf{Pr}[v_2 \leq 2b_1] = \mathbf{Pr}[F(2b_1)] = 2b_1$, so,

$$\begin{aligned}\mathbf{E}[u_1] &= (v_1 - b_1) \times 2b_1 \\ &= 2v_1 b_1 - 2b_1^2.\end{aligned}$$

Third, we optimize agent 1's bid. Agent 1 with value v_1 should maximize this quantity as a function of b_1, and to do so, can differentiate the function and set its derivative equal to zero. The result is $\frac{d}{db_1}(2v_1b_1 - 2b_1^2) = 2v_1 - 4b_1 = 0$ and we can conclude that the optimal bid is $b_1 = v_1/2$. 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 the agents learn their types, but before playing in the game. In the interim, an agent knows her own type and 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 actions of all agents are known.

2.4 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., $t_i = v_i$. A game has an outcome $\mathbf{x} = (x_1, \ldots, x_n)$ and payments $\mathbf{p} = (p_1, \ldots, p_n)$ where x_i is an indicator for whether agent i indeed received their desired service, i.e., $x_i = 1$ 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 $u_i = v_i x_i - p_i$. Agents are risk-neutral expected utility maximizers.

A game G maps actions $\mathbf{b}$ of agents to an outcome and payment. Formally we will specify these outcomes and payments as:

- $x_i^G(\mathbf{b})$ = outcome to i when actions are $\mathbf{b}$, and
- $p_i^G(\mathbf{b})$ = payment from i when actions are $\mathbf{b}$.

Given a game G and a strategy profile $\mathbf{s}$ 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 the game outcome is much more relevant. Define

- $x_i(\mathbf{v}) = x_i^G(\mathbf{s}(\mathbf{v}))$, and
- $p_i(\mathbf{v}) = p_i^G(\mathbf{s}(\mathbf{v}))$.

We refer to the former as the *allocation rule* and the latter as the *payment rule* for G and $\mathbf{s}$ (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 $\mathbf{F}$ (conditioned on her value). For G, $\mathbf{s}$, and $\mathbf{F}$ taken implicitly we can specify agent i's interim allocation and payment rules as functions of v_i.

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

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

- $u_i(v_i) = v_i x_i(v_i) - p_i(v_i)$.

Finally, we say that a strategy $s_i(\cdot)$ is *onto* if every action b_i agent i could play in the game is prescribed by s_i for some value v_i, i.e., $\forall b_i \exists v_i \, s_i(v_i) = b_i$. 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.

Fact 2.6. *For single-dimensional game G and common prior $\mathbf{F}$, an onto strategy profile $\mathbf{s}$ is in BNE if and only if for all i, v_i, and z,*

$$v_i x_i(v_i) - p_i(v_i) \geq v_i x_i(z) - p_i(z),$$

where G, $\mathbf{F}$, and $\mathbf{s}$ are implicit in the definition of $x_i(\cdot)$ and $p_i(\cdot)$.

It is easy to see that the restriction to onto strategies is only required for the "if" direction of Fact 2.6; the "only if" direction holds for all strategy profiles.

2.5 Characterization of Bayes-Nash equilibrium

We now discuss what Bayes-Nash equilibria look like. For instance, when given G, $\mathbf{s}$, and $\mathbf{F}$ we can calculate the interim allocation and payment rules $x_i(v_i)$ and $p_i(v_i)$ of each agent. We want to succinctly describe properties of these allocation and payment rules that can arise as BNE.

Theorem 2.7. *When values are drawn from a continuous joint distribution $\mathbf{F}$; G, $\mathbf{s}$, and $\mathbf{F}$ are in BNE only if for all i,*

1. *(monotonicity) $x_i(v_i)$ is monotone non-decreasing, and*
2. *(payment identity) $p_i(v_i) = v_i x_i(v_i) - \int_0^{v_i} x_i(z)\,dz + p_i(0)$,*

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.

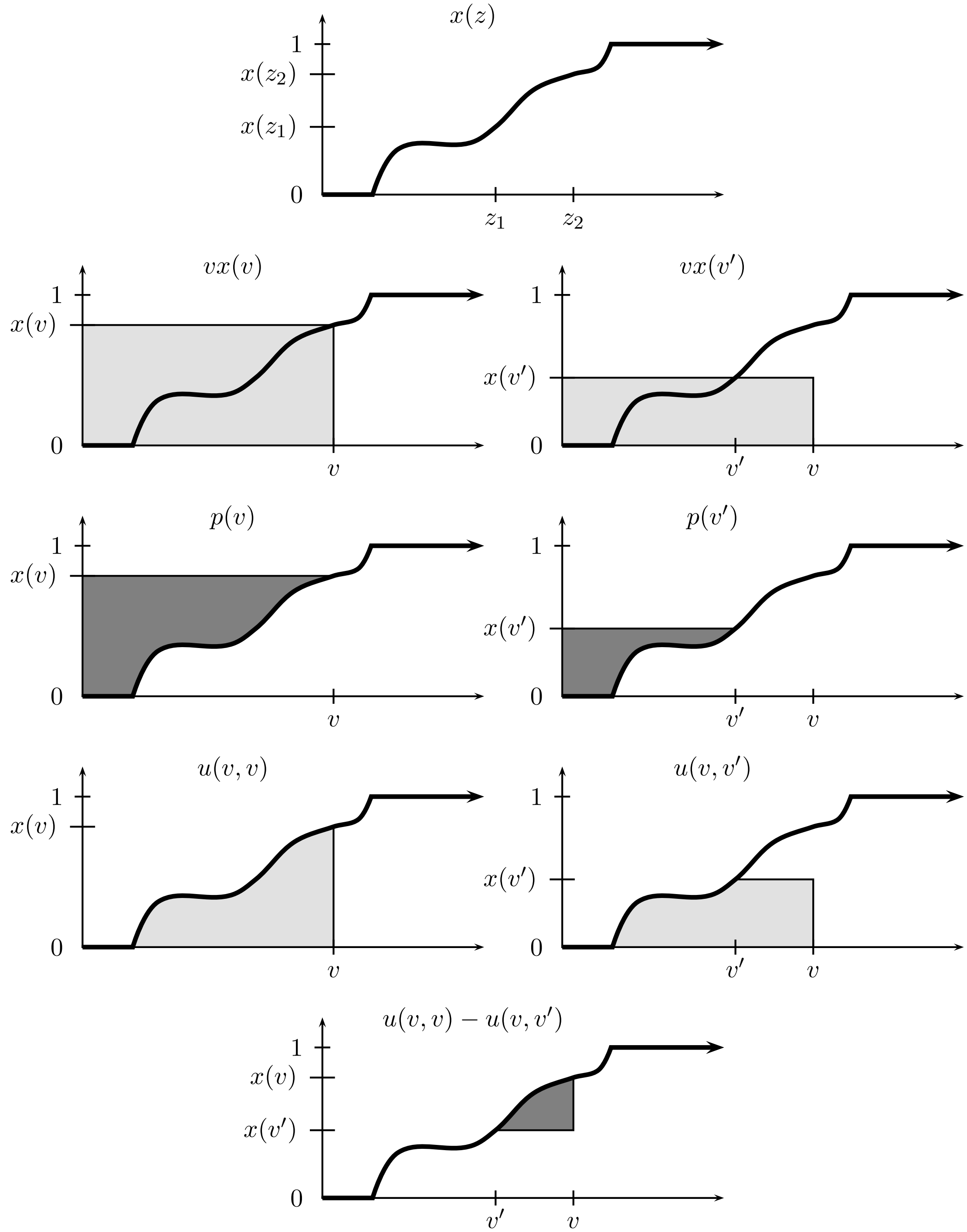

Figure 2.1: The left column shows (shaded) the surplus, payment, and utility of Alice playing action $s(v = z_2)$. The right column shows (shaded) the same for Alice playing action $s(v' = z_1)$. The final diagram shows (shaded) the difference between Alice's utility for these strategies. Monotonicity implies this difference is non-negative.

Note that if Alice with value v deviates from the equilibrium and takes action $s(v')$ instead of $s(v)$ then she will receive outcome and payment $x(v')$ and $p(v')$. This motivates the definition,

$$u(v, v') = vx(v') - p(v'),$$

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 v', $u(v, v) \geq u(v, v')$, i.e., Alice derives no increased utility by deviating. The strategy profile **s** 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 v' such that $s(v')$ is that action.)

1. G, **s**, and **F** are in BNE if **s** 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 $z_1 < z_2$. Supposing Alice has the high value, $v = z_2$, we argue that Alice does not benefit by simulating her strategy for the lower value, $v' = z_1$, i.e., by playing $s(v')$ to obtain outcome $x(v')$ and payment $p(v')$. We leave the proof of the opposite, that when $v = z_1$ and Alice is considering simulating the higher strategy $v' = z_2$, as an exercise for the reader.

   To start with this proof, we assume that $x(v)$ is monotone and that $p(v) = vx(v) - \int_0^v x(z)\,dz$.

   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 show Alice's surplus, $vx(v)$; payment, $p(v)$, and utility, $u(v) = vx(v) - p(v)$, assuming that she follow the BNE strategy $s(v = z_2)$. The column on the right shows the analogous quantities when Alice follows strategy $s(v' = z_1)$ but has value $v = z_2$. 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, with value v, (weakly) prefers to play the BNE strategy $s(v)$.

2. G, **s**, and **F** are in BNE only if the allocation rule is monotone.

   If we are in BNE then for all valuations, v and v', $u(v, v) \geq u(v, v')$. Expanding we require

   $$vx(v) - p(v) \geq vx(v') - p(v').$$

   We now consider z_1 and z_2 and take turns setting $v = z_1$, $v' = z_2$, and $v' = z_1$, $v = z_2$. This yields the following two inequalities:

   $$v = z_2, v' = z_1 \Longrightarrow z_2x(z_2) - p(z_2) \geq z_2x(z_1) - p(z_1), \text{ and} \tag{2.1}$$

   $$v = z_1, v' = z_2 \Longrightarrow z_1x(z_1) - p(z_1) \geq z_1x(z_2) - p(z_2). \tag{2.2}$$

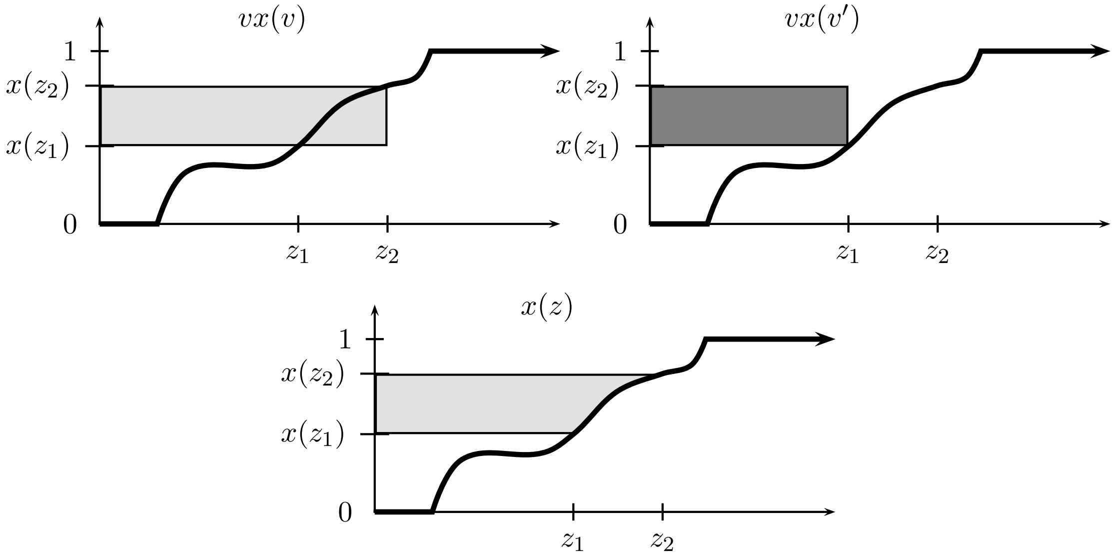

Figure 2.2: Upper (top, left) and lower bounds (top, right) for the difference in payments for two strategies z_1 and z_2 imply that the difference in payments (bottom) must satisfy the payment identity.

Adding these inequalities and canceling the payment terms we have,

$$z_2x(z_2) + z_1x(z_1) \geq z_2x(z_1) + z_1x(z_2).$$

Rearranging,

$$(z_2 - z_1)(x(z_2) - x(z_1)) \geq 0.$$

For $z_2 - z_1 > 0$ it must be that $x(z_2) - x(z_1) \geq 0$, i.e., $x(\cdot)$ is monotone non-decreasing.

3. G, $\mathbf{s}$, and $\mathbf{F}$ are in BNE only if the payment rule satisfies the payment identity.

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

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

$$u'(v,v) = vx'(v) - p'(v) = 0.$$

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

$$zx'(z) - p'(z) = 0.$$

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

$$\begin{aligned} p'(z) &= zx'(z), \text{ so} \\ \int_0^v p'(z) dz &= \int_0^v zx'(z)\, dz, \text{ so} \\ p(v) - p(0) &= \Big[zx(z)\Big]_0^v - \int_0^v x(z)\, dz \\ &= vx(v) - \int_0^v x(z)\, dz. \end{aligned}$$

Adding $p(0)$ from both sides of the equality, we conclude that the payment identity must hold.

Picture-based proof: Consider equations (2.1) and (2.2) and solve in each for $p(z_2) - p(z_1)$ in each:

$$z_2(x(z_2) - x(z_1)) \geq p(z_2) - p(z_1) \geq z_1(x(z_2) - x(z_1)).$$

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 $x(z_1)$ and $x(z_2)$. See Figure 2.2. The payment identity follows by taking $z_1 = 0$ and $z_2 = v$. □

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 them regardless of what other agents are doing. The DSE characterization theorem below follows from the BNE characterization theorem.

Theorem 2.8. *G and $\mathbf{s}$ are in DSE only if for all i,*

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

where $(\mathbf{v}_{-i}, z)$ denotes the valuation profile with the ith 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 $x_i(\mathbf{v}_{-i}, v_i)$ 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., τ_i at which $x_i(\mathbf{v}_{-i}, \cdot)$ changes from 0 to 1. This gives the following corollary.

Corollary 2.9. *A deterministic game G and strategies $\mathbf{s}$ are in DSE only if for all i,*

1. *(step-function) $x_i(\mathbf{v}_{-i}, v_i)$ steps from 0 to 1 at $\tau_i = \inf\{z : x_i(\mathbf{v}_{-i}, z) = 1\}$, and*

2. *(critical value) for* $p_i(\mathbf{v}_{-i}, v_i) = \begin{cases} \tau_i & \text{if } x_i(\mathbf{v}_{-i}, v_i) = 1 \\ 0 & \text{otherwise} \end{cases} + p_i(\mathbf{v}_{-i}, 0)$.

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, but, 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 $t_1 = v_2$ and $t_2 = v_1$. We will usually prefer the randomized tie-breaking rule because of its symmetry.

2.7 Revenue Equivalence

We are now ready to make one of the most significant observations in auction theory. Namely, mechanisms with the same outcome 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 in each mechanism. We state this result as a corollary of Theorem 2.7 which is intuitively clear. The payment identity means that the payment rule is precisely determined by the allocation rule and the payment of the lowest type, i.e., $p_i(0)$.

Corollary 2.10. *Consider any two mechanisms where 0-valued agents pay nothing, if the mechanisms have the same BNE outcome then they have same expected revenue.*

We can now quantitatively compare the second-price and first-price auctions from a revenue standpoint. Consider the case where the agent's values are distributed independently and identically. 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 $s_i = s_{i'}$ for all i and i'. Furthermore, it is reasonable to expect that the strategies are 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, we can conclude that the two auctions obtain 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 $\mathbf{E}\big[v_{(2)}\big]$. In Section 2.3 we saw that the symmetric strategy of the first-price auction in this environment is for each agent to bid half their value. The expected revenue of first-price auction is therefore $\mathbf{E}\big[v_{(1)}/2\big]$. An important fact about uniform random variables is that in expectation they evenly divide the interval they are over, i.e., $\mathbf{E}\big[v_{(1)}\big] = 2/3$ and $\mathbf{E}\big[v_{(2)}\big] = 1/3$. How do the revenues of these two auctions compare? Their revenues are identically 1/3.

Corollary 2.11. *When agent's values are independent and identically distributed, the second-price and first-price auction have the same expected revenue.*

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 bid; is revenue equivalent to the first- and second-price auctions.

2.8 Solving for Bayes-Nash Equilibrium

While it is quite important to know what outcomes are possible in BNE, it is also often important to be able to solve for the BNE strategies. For instance, suppose you were a bidder bidding in an auction. How would you bid? In this section we describe an elegant technique for calculating BNE strategies in symmetric environments using revenue equivalence. Actually, we use something a little stronger than revenue equivalence: *interim payment equivalence*. This is the fact that if two mechanisms have the same allocation rule, they must have the same payment rule (because the payment rules satisfy the payment identity). As described previously, the interim payment of agent i with value v_i is $p_i(v_i)$.

Suppose we are to solve for the BNE strategies of mechanism M. The approach is to express an agent's payment in M as a function of the agent's action, then to calculate the agent's expected payment in a strategically-simple mechanism M' that is revenue equivalent to M (usually a "second-price implementation" of M). Setting these terms equal and solving for the agents action gives the equilibrium strategy.

We give the high level the procedure below. As a running example we will calculate the equilibrium strategies in the first-price auction with two $U[0,1]$ agents, in doing so we will use a calculation of expected payments in the strategically-simple second-price auction in the same environment.

1. *Guess* what the outcome might be in Bayes-Nash equilibrium.

   E.g., for the first-price auction with two agents with values $U[0,1]$, in BNE, we expect the agent with the highest value to win. Thus, guess that the highest-valued agent always wins.

2. *Calculate* the expected payment of an agent with fixed value in a strategically-simple auction with the same equilibrium outcome.

   E.g., recall that it is a dominant strategy equilibrium (a special case of Bayes-Nash equilibrium) in the second-price auction for each agent to bid their value. I.e., $b_1 = v_1$ and $b_2 = v_2$. Thus, the second-price auction also selects the agent with the highest value to win. So, let us calculate the expected payment of player 1 when their value is v_1.

   $$\mathbf{E}[p_1(v_1)] = \mathbf{E}[p_1(v_1) \mid 1 \text{ wins}]\,\mathbf{Pr}[1 \text{ wins}] + \mathbf{E}[p_1(v_1) \mid 1 \text{ loses}]\,\mathbf{Pr}[1 \text{ loses}]\,.$$

   Calculate each of these components for the second-price auction:

   $$\begin{aligned}\mathbf{E}[p_1(v_1) \mid 1 \text{ wins}] &= \mathbf{E}[v_2 \mid v_2 < v_1] \\ &= v_1/2.\end{aligned}$$

   The first step follows by the definition of the second-price auction and its dominant strategy equilibrium (i.e., $b_2 = v_2$). The second step follows because in expectation a uniform random variable evenly divides the interval it is over, and once we condition on $v_2 < v_1$, v_2 is $U[0, v_1]$.

   $$\begin{aligned}\mathbf{Pr}[1 \text{ wins}] &= \mathbf{Pr}[v_2 < v_1] \\ &= v_1.\end{aligned}$$

   The first step follows from the definition of the second-price auction and its dominant strategy equilibrium. The second step is because v_1 is uniform on $[0,1]$.

   $$\mathbf{E}[p_1(v_1) \mid 1 \text{ loses}] = 0.$$

   This is because a loser pays nothing in the second-price auction. This means that we do not need to calculate $\mathbf{Pr}[1 \text{ loses}]$. Plug these into the equation above to obtain:

   $$\mathbf{E}[p_1(v_1)] = v_1^2/2.$$

3. *Solve* for bidding strategies from expected payments.

   E.g., by revenue equivalence the expected payment of player 1 with value v_1 is $v_1^2/2$ in both the first-price and second-price auction. We can recalculate this expected payment in the first-price auction using the bid of the player as a variable and then solve for what that bid must be.

   Again,

$$\mathbf{E}[p_1(v_1)] = \mathbf{E}[p_1(v_1) \mid 1 \text{ wins}]\,\mathbf{Pr}[1 \text{ wins}] + \mathbf{E}[p_1(v_1) \mid 1 \text{ loses}]\,\mathbf{Pr}[1 \text{ loses}]\,.$$

   Calculate each of these components for the first-price auction where agent 1 follows strategy $s_1(v_1)$:

$$\mathbf{E}[p_i(v_i) \mid 1 \text{ wins}] = s_1(v_1).$$

   This by the definition of the first-price auction: if you win you pay your bid.

$$\mathbf{Pr}[1 \text{ wins}] = \mathbf{Pr}[v_2 < v_1] = v_1.$$

   This is the same as above.

$$\mathbf{E}[p_1(v_1) \mid 1 \text{ loses}] = 0.$$

   This is because a loser pays nothing in the first-price auction. This means that we do not need to calculate **Pr**[1 loses]. Plug these into the equation above to obtain:

$$\mathbf{E}[p_1(v_1)] = s_1(v_1) \cdot v_1.$$

   This must equal the expected payment calculated in the previous step for the second-price auction, implying:

$$v_1^2/2 = s_1(v_1) \cdot v_1.$$

   We can solve for $s_1(v_1)$ and get

$$s_1(v_1) = v_1/2.$$

4. Verify initial guess was correct.

   Indeed, if agents follow symmetric strategies $s_1(z) = s_2(z) = z/2$ then the agent with the highest value wins.

In the above first-price auction example it should be readily apparent that we did slightly more work than we had to. In this case it would have been enough to note that in both the first- and second-price auction a loser pays nothing. We could therefore simply equate the expected payments conditioned on winning:

$$\mathbf{E}[p_1(v_1) \mid 1 \text{ wins}] = \underbrace{v_1/2}_{\text{second-price}} = \underbrace{s_1(v_i)}_{\text{first-price}} .$$

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. The all-pay auction is also is revenue equivalent to the second-price auction. However, now we compare the total expected payment (regardless of winning) to conclude:

$$\mathbf{E}[p_1(v_1)] = \underbrace{v_1^2/2}_{\text{second-price}} = \underbrace{s_1(v_i)}_{\text{all-pay}} .$$

I.e., the BNE strategies for the all-pay auction are $s_i(z) = z^2/2$. 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 difference.

2.9 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 2.12. *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 value they might have)*

Definition 2.13. *A direct revelation mechanism is* Bayesian incentive compatible *(BIC) if truthtelling is a Bayes-Nash equilibrium.*

Definition 2.14. *A direct revelation mechanism is* dominant strategy incentive compatible *(DSIC) if truthtelling is a dominant strategy equilibrium.*

Theorem 2.15. *Any mechanism M with good BNE (resp. DSE) can be converted into a BIC (resp. DSIC) mechanism M' with the same BNE (resp. DSE) outcome.*

Proof. We will prove the BNE variant of the theorem. Let $\mathbf{s}$, $\mathbf{F}$, and M be in BNE. Define single-round, sealed-bid mechanism M' as follows:

1. Accept sealed bids $\mathbf{b}$.

2. Simulate $\mathbf{s}(\mathbf{b})$ in M.

3. Output the outcome of the simulation.

We now claim that $\mathbf{s}$ being a BNE of M implies truthtelling is a BNE of M' (for distribution $\mathbf{F}$). Let $\mathbf{s}'$ denote the truthtelling strategy. In M', consider agent i and suppose all other agents are truthtelling. This means that the actions of the other players in M are distributed as $\mathbf{s}_{-i}(\mathbf{s}'_{-i}(\mathbf{v}_{-i})) = \mathbf{s}_{-i}(\mathbf{v}_{-i})$ for $\mathbf{v}_{-i} \sim \mathbf{F}_{-i}\big|_{v_i}$. Of course, in M if other players are playing $\mathbf{s}_{-i}(\mathbf{v}_{-i})$ then since $\mathbf{s}$ is a BNE, i's best response is to play $s_i(v_i)$ as well. Agent i can play this action in the simulation of M is by playing the truthtelling strategy $s'_i(v_i) = v_i$ in M'. □

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

We defined the allocation and payment rules $\mathbf{x}(\cdot)$ and $\mathbf{p}(\cdot)$ as functions of the valuation profile for an implicit game G and strategy profile $\mathbf{s}$. When the strategy profile is truthtelling, the allocation and payment rules are identical the original mappings of the game from actions to allocations and prices, denoted $\mathbf{x}^G(\cdot)$ and $\mathbf{p}^G(\cdot)$. Additionally, let $x_i^G(v_i) = \mathbf{E}\big[x_i^G(\mathbf{v}) \mid v_i\big]$ and $p_i^G(v_i) = \mathbf{E}\big[p_i^G(\mathbf{v}) \mid v_i\big]$ for $\mathbf{v} \sim \mathbf{F}$. Furthermore, the truthtelling strategy profile in a direct-revelation game is onto.

Corollary 2.16. *A direct mechanism M is BIC for distribution $\mathbf{F}$ if and only if for all i,*

1. *(monotonicity) $x_i^M(v_i)$ is monotone non-decreasing, and*
2. *(payment identity) $p_i^M(v_i) = v_i x_i^M(v_i) - \int_0^{v_i} x_i^M(z)\,dz + p_i^M(0)$.*

Corollary 2.17. *A direct mechanism M is DSIC if and only if for all i,*

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

Corollary 2.18. *A direct, deterministic mechanism M is DSIC if and only if for all i,*

1. *(step-function) $x_i^M(\mathbf{v}_{-i}, v_i)$ steps from 0 to 1 at $\tau_i = \inf\{z : x_i^M(\mathbf{v}_{-i}, z) = 1\}$, and*
2. *(critical value) for* $p_i^M(\mathbf{v}_{-i}, v_i) = \begin{cases} \tau_i & \text{if } x_i^M(\mathbf{v}_{-i}, v_i) = 1 \\ 0 & \text{otherwise} \end{cases} + p_i^M(\mathbf{v}_{-i}, 0).$

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 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, $\mathbf{v} = (v_1, \ldots, v_n)$, an allocation vector, $\mathbf{x} = (x_1, \ldots, x_n)$, and payments, $\mathbf{p} = (p_1, \ldots, p_n)$. 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,

$$v_i x_i - p_i \geq v_i x_j - p_j.$$

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.)

2.3 AdWords is Google product in which the search engine sells at auction advertisements that appear along side search results on the 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. Assume that the slots are ordered from highest click-through rate to lowest, i.e., $w_j \geq w_{j+1}$ for all j.

(a) Find the envy-free (See Exercise 2.2) outcome and payments with the maximum social 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 $\mathbf{p}'$ from $\mathbf{p}$.

(c) The real AdWords problem is solved by auction. Design an auction that maximizes the social 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.

2.4 Consider the first-price auction for selling k units of an item to n unit-demand agents. This auction solicits bids and allocates one units to each of the k highest-bidding agents. These winners are charged their bids. This auction is revenue equivalent to the k-unit "second-price" auction where the winners are charged the $k+1$st highest bid, $b_{(k+1)}$. Solve for the symmetric Bayes-Nash equilibrium strategies in the first-price auction when the agent values are i.i.d. $U[0, 1]$.

2.5 Consider the position auction environment with $n = m = 2$ (See Exercise 2.3). Consider running the following first-price auction: The advertisers submit bids $\mathbf{b} = (b_1, b_2)$. The advertisers are assigned to slots in order of their bids. Advertisers pay their bid when clicked. Use revenue equivalence to solve for BNE strategies $\mathbf{s}$ when the values of the advertisers are drawn independent and identically from $U[0, 1]$.

Chapter Notes

The characterization of Bayes-Nash equilibrium, revenue equivalence, and the revelation principle come from Myerson (1981). The BNE characterization proof presented here comes from Archer and Tardos (2001).