B. Incentives in Teams

Rarely do the employees of a firm work as individuals on separate tasks. Salespeople working in distinct assigned regions come closest to being so separate, although even in their case, the performance of an individual salesperson is affected by the support of others in the office. Usually people work in teams, and the outcome for the team, and for each of its members, depends on the efforts of all. A firm's profit as a whole, for example, depends on the performance of all of its workers and managers. This interaction creates special incentive-design problems.

When one worker's earnings depend on the profit of the firm as a whole, that worker will see only a weak link between his effort and the aggregate profit. Because each worker in a large firm gets only a small fractional share of the aggregate profit, each has a very weak incentive to exert extra effort. Even in a smaller team, each member will be tempted to shirk and become a free rider on the efforts of the others. This outcome mirrors the prisoners' dilemma of collective action that we saw in the street-garden game of Chapters 3 and 4, and throughout Chapter 10. If the team is small and stable over a sufficiently long time, we can expect its members to resolve the dilemma by devising internal and perhaps nonmonetary schemes of rewards and punishments like the ones we saw in Chapter 10, Section 3.

In another context, the existence of many workers on a team can sharpen incentives. Suppose a firm has many workers performing similar tasks, perhaps selling different items from the firm's product line. If there is a common (positively correlated) random component to all of the workers' sales, perhaps based on the strength of the underlying economy, then the sales of one worker relative to those of another worker are a good indicator of their relative effort levels. For example, the efforts of workers 1 and 2, denoted by x_1 and x_2, might be related to their sales, y_1 and y_2, according to the formulas $y_1 = x_1 + r$ and $y_2 = x_2 + r$, where r represents the common random component. In this case, it follows that $y_2 - y_1 = x_2 - x_1$ with no randomness; that is, the difference in observed sales will exactly equal the difference in exerted effort between workers 1 and 2. The firm employing these workers can then reward them according to their relative outcomes. This payment scheme entails no risk for the workers. The trade-off we considered in Section 5, between evoking optimal effort and sharing the profits of the firm, vanishes. Now, if the first worker has a poor sales record and tries to blame it on bad luck, the firm can respond, "Then how come this other worker achieved so much more? Luck was common to the two of you, so you must have made less effort." Of course, if the two workers can collude, they can defeat the firm's purpose, but otherwise the firm can implement a powerful incentive scheme by setting workers in competition with one another. An extreme example of such a scheme is a tournament in which the best performer gets a prize.

Tournaments also help mitigate another potential moral-hazard problem. In reality, whether a worker has met the criteria of success may itself not be easily or publicly observable. Then the owner of the firm may be tempted to claim that no one has performed well enough and that no one should be paid a bonus. A tournament with a prize that must be awarded to someone, or a given aggregate bonus pool that must be distributed among the workers, eliminates this moral hazard on the part of the principal.

C. Multiple Tasks and Outcomes

Employees usually perform several tasks for their employers. These various tasks lead to several measurable outcomes of employee effort. Incentives for providing effort to these different tasks then interact. This interaction makes mechanism design more complex for the firm.

The outcome of each of an agent’s tasks depends partly on the agent’s effort and partly on chance. That is why an outcome-based incentive scheme generally exposes the agent to some risk. If the element of chance is small, then the risk to the agent is small, and the incentive to exert effort can be made more powerful. Of course, the outcomes of different tasks are likely to be affected by chance to different extents. So if the principal considers the tasks one at a time, he will use stronger incentives for effort on the tasks that have smaller elements of chance and weaker incentives for effort on the tasks where outcomes are more uncertain indicators of the agent’s effort. But a powerful incentive for one task will divert the agent’s effort away from the other task, further weakening the agent’s performance on that task. To avoid this diversion of effort toward the task with the stronger incentive, the principal has to weaken the incentive on that task, too.

An example of this problem can be found in our own lives. Professors are supposed to do research as well as teaching. There are many accurate indicators of good research: publications in and appointments to editorial positions for prestigious journals, elections to scientific academies, and so on. By contrast, indicators of good teaching can be observed only imprecisely and with long lags. Students often need years of experience to recognize the value of what they learned in college; in the short term, they may be more impressed by showmanship than by scholarship. If these two tasks required of faculty members were considered in isolation, university administrators would attach powerful incentives to research and weaker incentives to teaching. But if they did so, professors would divert their efforts away from teaching and toward research (even more so than they would otherwise do). Therefore, the imprecise observation of teaching outcomes forces deans and presidents to offer only weak incentives for research as well.

The most cited example of a situation involving multiple tasks and outcomes occurs in school teaching. Some outcomes of teaching, such as test scores, are precisely observable, whereas other valuable indicators of education, such as the ability to work in teams or speak in public, are less accurately measurable. If teachers are rewarded on the basis of their students' test scores, they will "teach to the test," and the other dimensions of their students' education will be ignored. Such "gaming" of an incentive scheme also extends to sports. In baseball, if a hitter is rewarded for hitting home runs, he will neglect other aspects of batting (taking pitches, sacrifice bunts, etc.) that can sometimes contribute more to his team's chances of winning a game. Similarly, salespeople may sacrifice long-term customer relationships in favor of driving home a sale to meet a short-term sales quota.

If the dysfunctional effects of some incentives on other tasks are too severe, other systems of rewarding tasks may be needed. A more holistic but more subjective measure of performance—for example, an overall evaluation by the worker's boss, may be used. This alternative is not without its own problems; workers may then divert their effort into activities that find favor with the boss!

D. Incentives over Time

Many employment relationships last for a long time, and that opens up opportunities for the firm to devise incentive schemes where an employee's performance at one time is rewarded at a later time. Firms regularly use promotions, seniority-based salaries, and other forms of deferred compensation for this purpose. In effect, workers are underpaid relative to their performance in the earlier stages of their careers with the firm and overpaid in later years. The prospect of future rewards motivates younger workers to exert good effort and also induces them to stay with the firm, thus reducing job turnover. Of course, the firm may be tempted to renege on its implicit promise of overpayment in later years; therefore, such schemes must be credible if they are to be effective. They are more likely to be used effectively in firms that have a long record of stability and a reputation for treating their senior workers well.

A different way that the prospect of future compensation can keep workers motivated is through the use of an efficiency wage. The firm pays a worker more than the going wage, and the excess is a surplus, or economic rent, for the worker. So long as the worker makes good effort, he will go on earning this surplus. But if he shirks, he may be detected, at which point he will be fired and will have to go back to the general labor market, where he can earn only the going wage.

The firm faces a mechanism-design problem when it tries to determine the appropriate efficiency wage level. Suppose the going wage is w_0, and the firm's efficiency wage is $w > w_0$. Let the monetary equivalent of the worker's subjective cost of making good effort be e. In each pay period, the worker has the choice of whether to make this effort. If the worker shirks, he saves e. But with probability p, the shirking will be detected. If it is discovered that he has been shirking, the worker will lose the surplus $(w - w_0)$, starting in the next pay period and continuing indefinitely. Let r be the rate of interest from one period to the next. Then, if the worker shirks today, the expected discounted present value of the worker's loss in the next pay period is $p(w - w_0)/(1 + r)$. And the worker loses $w - w_0$ with probability p in all future pay periods. A calculation similar to the ones we performed for repeated games in Chapter 10 and its appendix shows that the total expected discounted present value of the future loss to the worker is

$$p\left[\frac{w-w_0}{1+r} + \frac{w-w_0}{(1+r)^2} + \cdots\right] = p(w-w_0)\,\frac{1/(1+r)}{1-1/(1+r)} = \frac{p(w-w_0)}{r}.$$

To deter shirking, the firm needs to make sure that this expected loss is at least as high as the worker's immediate gain from shirking, e. Therefore, the firm must pay an efficiency wage that satisfies

$$\frac{p(w-w_0)}{r} \geq e \quad \text{or} \quad w - w_0 \geq \frac{er}{p} \quad \text{or} \quad w \geq w_0 + \frac{er}{p}.$$

The smallest efficiency wage is the one that makes this expression hold with equality. And the more accurately the firm can detect shirking (that is, the higher is p), the smaller its excess over the going wage needs to be.

A repeated relationship may also enable the firm to design a sharper incentive scheme in another way. In any one period, as we explained above, the worker's observed outcome is a combination of the worker's effort and an element of chance. But if the outcome is poor year after year, the worker cannot credibly blame bad luck year after year. Therefore, the average outcome over a long period can, by the law of large numbers, be used as an accurate measure of the worker's average effort, and the worker can be rewarded or punished accordingly.

Endnotes

- Canice Prendergast, “The Provision of Incentives in Firms,” *Journal of Economic Literature*, vol. 37, no. 1 (March 1999), pp. 7–63, is an excellent survey of the theory and practice of incentive mechanisms. Prendergast gives references to the original research literature from which many of the findings and anecdotes mentioned in this section are taken, so we will not repeat the specific citations. James N. Baron and David M. Kreps, *Strategic Human Resources: Frameworks for General Managers* (New York: Wiley, 1999), is a wider-ranging book on personnel management, combining perspectives from economics, sociology, and social psychology; Chapters 8, 11, and 16 and appendixes C and D are closest to the concerns of this chapter and this book. Return to reference 14

Glossary

efficiency wage

A higher-than-market wage paid to a worker as a means of incentivizing him to exert effort. If the worker shirks and is detected, he will be fired and will have to get a lower-wage job in the general labor market.

SUMMARY

The study of *mechanism design* can be summed up as learning "how to deal with someone who knows more than you do." Such situations occur in numerous contexts, usually in interactions involving a more informed player, called the *agent*, and a less informed player, called the *principal*, who wants to design a mechanism to give the agent the correct incentives to help the principal attain his goal.

Mechanism-design problems are of two types. In the first type, the principal creates a scheme that requires the agent to reveal information. In the second type, which involves moral hazard, the principal creates a scheme to elicit the optimal level of an unobservable action by the agent. In all cases, the principal attempts to maximize his own benefit (or payoff) function subject to the incentive-compatibility and participation constraints imposed by the agent.

Firms use information-revelation schemes in creating pricing structures that separate customers by their willingness to pay for the firm's product. Procurement contracts are also often designed to separate projects, or contractors, according to various levels of cost. Evidence of both *price discrimination* and screening with procurement contracts can be seen in actual markets.

When facing moral hazard, employers must devise incentives that encourage their employees to provide optimal effort. Similarly, insurance companies must write policies that give their clients the right incentives to take action to reduce the probability of loss. In some simple situations, optimal contracts will be linear schemes, but in the presence of more complex relationships, nonlinear schemes may be more beneficial. Incentive schemes designed for workers in teams,

or whose relationships continue over time, are correspondingly more complex than those designed for simpler situations.

KEY TERMS

agent (556)

efficiency wage (573)

incentive design (551)

mechanism design (551)

price discrimination (552)

principal (556)

principal–agent (agency) problem (556)

Glossary

mechanism design

Same as **incentive design**.

incentive design

The process that a *principal* uses to devise the best possible incentive scheme (or mechanism) in a *principal–agent problem* to motivate the agent to take actions that benefit the principal. By design, such incentive schemes take into account that the agent knows something (about the world or about herself) that the principal does not know. Also called **mechanism design**.

price discrimination

Perfect, or first-degree, price discrimination occurs when a firm charges each customer an individualized price based on willingness to pay. In general, price discrimination refers to situations in which a firm charges different prices to different customers for the same product.

principal

The principal is the less-informed player in a principal–agent game of asymmetric information. The principal in such games wants to design a mechanism that creates incentives for the more-informed player (agent) to take actions beneficial to the principal.

agent

The agent is the more-informed player in a principal–agent game of asymmetric information. The principal (less-informed) player in such games attempts to design a mechanism that aligns the agent’s incentives with his own.

principal–agent (agency) problem

A situation in which the less-informed player (principal) wants to design a mechanism that creates incentives for the more-informed player (agent) to take actions beneficial to himself (the principal).

efficiency wage

A higher-than-market wage paid to a worker as a means of incentivizing him to exert effort. If the worker shirks and is detected, he will be fired and will have to get a lower-wage job in the general labor market.

SOLVED EXERCISES

1. Firms that provide insurance to clients to protect them from the costs associated with theft or accident must necessarily be interested in the behavior of their policyholders. Sketch some ideas for the creation of an incentive scheme that such a firm might use to deter and detect fraud or lack of care on the part of its policyholders.
2. Some firms sell goods and services either singly at a fairly high price or in larger quantities at a discount in order to increase their own profit by separating consumers with different preferences.
   1. List three examples of quantity discounts offered by firms.
   2. How do quantity discounts allow firms to screen consumers by their preferences?
3. Omniscient Wireless Limited (OWL) is planning to roll out a new nationwide broadband wireless telephone service next month. The firm has conducted market research indicating that its 10 million potential customers can be divided into two types, which they call Light users and Regular users. Light users have less demand for wireless phone service, and in particular, they seem unlikely to place any value on more than 300 minutes of calls per month. Regular users have more demand for wireless phone service generally and place a high value on more than 300 minutes of calls per month. OWL analysts have determined that the best service plans to offer to customers entail 300 minutes per month and 600 minutes per month, respectively. They estimate that 50% of customers are Light users and 50% are Regular users, and that the two types have the following willingness to pay for each type of service:

	300 minutes	600 minutes

	300 minutes	600 minutes
Light user (50%)	$20	$30
Regular user (50%)	$25	$70

OWL's cost per additional minute of wireless service is negligible, so the cost to the company of providing service is $10 per user, no matter which plan the user chooses.

Each potential customer calculates the net payoff (benefit minus price) that she would get from each of the service plans and buys the plan that would give her the higher net payoff, so long as this payoff is not negative. If both plans give her equal, nonnegative net payoffs, she goes for 600 minutes; if both plans give her negative net payoffs, she does not purchase. OWL wants to maximize its expected profit per potential customer.

1. Suppose the firm were to offer only the 300-minute plan, but not the 600-minute plan. What would be the optimal price to charge, and what would be the average profit per potential customer?
2. Suppose instead that the firm were to offer only the 600-minute plan. What would be the optimal price, and what would be the average profit per potential customer?
3. Suppose the firm wanted to offer both plans. Suppose further that it wanted the Light users to purchase the 300-minute plan and the Regular users to purchase the 600-minute plan. Write down the incentive-compatibility constraint for the Light users.
4. Similarly, write down the incentive-compatibility constraint for the Regular users.
5. Use the results from parts (c) and (d) to calculate the optimal pair of prices to charge for the 300-minute and 600-minute plans, so that each user type

will purchase its intended service plan. What would be the average profit per potential customer?

6. Consider the outcomes described in parts (a), (b), and (e). For each of the three situations, describe whether it is a separating outcome, a pooling outcome, or a semiseparating outcome.

4. Mictel Corporation has a world monopoly on the production of personal computers. It can make two kinds of computers: low-end and high-end. One-fifth of the potential buyers are casual users, and the rest are intensive users.

The costs of production of the two kinds of machines, as well as the benefits gained from the two by the two types of prospective buyers, are given in the following table (all figures are in thousands of dollars):

		COST	BENEFIT FOR USER TYPE	
			Casual	Intensive
PC TYPE	Low-end	1	4	5
	High-end	3	5	8

You may need to scroll left and right to see the full figure.

Each type of buyer calculates the net payoff (benefit minus price) that he would get from each kind of machine and buys the kind that would give him the higher net payoff, so long as this payoff is not negative. If both kinds give him equal, nonnegative net payoffs, he goes for the high-end machine; if both kinds give him negative net payoffs, he does not purchase.

Mictel wants to maximize its expected profit.

1. If Mictel were omniscient, then, when a prospective customer came along, knowing his type, the company could offer to sell him just one kind of machine at a stated price, on a take-it-or-leave-it basis. What kind of machine would Mictel offer, and at what price, to each type of buyer?

   In fact, Mictel does not know the type of any particular buyer. It just makes its catalog available for all buyers to choose from.

2. First, suppose the company produces only the low-end machines and sells them for price *x*. What value of *x* will maximize its profit? Why?
3. Next, suppose Mictel produces only the high-end machines and sells them for price *y*. What value of *y* will maximize its profit? Why?
4. Finally, suppose the company produces both kinds of machines, selling the low-end machines for price *x* and the high-end machines for price *y*. What incentive-compatibility constraints on *x* and *y* must the company satisfy if it wants the casual users to buy the low-end machines and the intensive users to buy the high-end machines?
5. What participation constraints must *x* and *y* satisfy for the casual users to be willing to buy the low-end machines and for the intensive users to be willing to buy the high-end machines?
6. Given the constraints in parts (d) and (e), what values of *x* and *y* will maximize the expected profit when the company sells both kinds of machines? What is the company’s expected profit from this policy?
7. Putting it all together, decide what production and pricing policy the company should pursue.

5. Redo Exercise S4, assuming that one-half of Mictel’s customers are casual users.

6. Using the insights gained in Exercises S4 and S5, solve Exercise S4 for the general case in which the proportion of casual users is c and the proportion of intensive users is $(1 - c)$. The answers to some parts will depend on the value of c. In these instances, list all relevant cases and how they depend on c.
7. Sticky Shoe, the discount movie theater, sells popcorn and soda at its concession counter. Cameron, Jessica, and Sean are regular patrons of Sticky Shoe, and the valuations of each for popcorn and soda are as follows:

	Popcorn	Soda
Cameron	\$3.50	\$3.00
Jessica	\$4.00	\$2.50
Sean	\$1.50	\$3.50

There are 2,997 other residents of Harkinsville who see movies at Sticky Shoe. One-third of them have valuations identical to Cameron's, one-third to Jessica's, and one-third to Sean's. If a customer is indifferent between buying and not, she buys. It costs Sticky Shoe essentially nothing to produce each additional order of popcorn or soda.

1. If Sticky Shoe sets separate prices for popcorn and soda, what price should it set for each concession to maximize its profit? How much profit does Sticky Shoe make selling concessions separately?
2. What does each type of customer (Cameron, Jessica, Sean) buy when Sticky Shoe sets separate profit-maximizing prices for popcorn and soda?
3. Instead of selling the concessions separately, Sticky Shoe decides always to sell the popcorn and soda together in a combo, charging a single price for both. What single combo price would maximize its

profit? How much profit does Sticky Shoe make selling only combos?
4. What does each type of customer buy when Sticky Shoe sets a single profit-maximizing price for a popcorn and soda combo? How does this answer compare with your answer in part (b)?
5. Which pricing scheme does each customer type prefer? Why?
6. If Sticky Shoe sold the concessions both as a combo and separately, which products (popcorn, soda, or the combo) does it want to sell to each customer type? How can Sticky Shoe make sure that each customer type purchases exactly the product that it intends for them to purchase?
7. What prices—for the popcorn, soda, and the combo—would Sticky Shoe set to maximize its profit? How much profit does Sticky Shoe make selling the concessions at these three prices?
8. How do your answers to parts (a), (c), and (g) differ? Explain why.

8. Section 5.A of this chapter discusses the principal-agent problem in the context of a company deciding whether and how to induce a manager to put in extra effort to increase the chances that the project he will be managing succeeds. The value of a successful project is $1 million; the probability of success given extra effort is 0.5; the probability of success without extra effort is 0.25. The manager's wage in his current job is now $120,000, and the money equivalent of the extra effort is $60,000. The manager is loss averse, and gives money losses relative to $120,000 twice the weight of gains relative to that amount.
   1. What contract does the company offer if it does not want extra effort from the manager?
   2. What is the expected profit to the company when it does not induce extra managerial effort?

3. What contract pair (y, x)—where y is the salary paid for a successful project and x is the salary paid for a failed project—should the company offer the manager to induce extra effort?
4. What is the company's expected profit when it induces extra effort?
5. Which level of effort does the company want to induce from its manager? Why?

9. A company has purchased fire insurance for its main factory. The probability of a fire in the factory without a fire-prevention program is 0.01. The probability of a fire in a factory with a fire-protection program is 0.001. If a fire occurred, the value of the loss would be \$300,000. A fire-prevention program would cost \$80 to run, but the insurance company cannot observe whether or not the prevention program has been implemented without incurring extra costs.
   1. Why does moral hazard arise in this situation? What is its source?
   2. Can the insurance company eliminate the moral-hazard problem? If so, how? If not, explain why not.
10. Mozart moved from Salzburg to Vienna in 1781, hoping for a position at the Habsburg court. Instead of applying for a position, he waited for the emperor to call him, because "if one makes any move oneself, one receives less pay." Discuss this situation using the theory of games with asymmetric information, including theories of signaling and screening.
11. (**Optional**, requires calculus) You are Oceania's Minister for Peace, and it is your job to purchase war materials for your country. The net benefit, measured in Oceanic dollars, from quantity Q of these materials is $2Q^{1/2} - M$, where M is the amount of money paid for the materials.

    There is just one supplier—Baron Myerson's Armaments (BMA). You do not know BMA's cost of production. Everyone knows that BMA's cost per unit of output is constant, and that it is equal to 0.10 with probability p

= 0.4 and equal to 0.16 with probability $1 - p$. Call BMA "low cost" if its per-unit cost is 0.10 and "high cost" if it is 0.16. Only BMA knows its true cost type with certainty.

In the past, your ministry has used two kinds of purchase contracts: cost plus and fixed price. But cost-plus contracts create an incentive for BMA to overstate its cost, and fixed-price contracts may compensate the firm more than is necessary. You decide to offer a menu of two possibilities:

Contract 1: Supply us quantity Q_1, and we will pay you money M_1.

Contract 2: Supply us quantity Q_2, and we will pay you money M_2.

The idea is to set Q_1, M_1, Q_2, and M_2 such that a low-cost BMA will find contract 1 more profitable, and a high-cost BMA will find contract 2 more profitable. If another contract is exactly as profitable, a low-cost BMA will choose contract 1, and a high-cost BMA will choose contract 2. Further, regardless of its cost, BMA will need to receive at least zero economic profit in any contract it accepts.

1. Write expressions for the profit of a low-cost BMA and a high-cost BMA when it supplies quantity Q and is paid M.
2. Write the incentive-compatibility constraints to induce a low-cost BMA to select contract 1 and a high-cost BMA to select contract 2.
3. Give the participation constraints for each type of BMA.
4. Assuming that each of the BMA types chooses the contract designed for it, write the expression for

Oceania's expected net benefit.

Now your problem is to choose Q_1, M_1, Q_2, and M_2 to maximize the expected net benefit found in part (d) subject to the incentive-compatibility (IC) and participation constraints (PC).

5. Assume that $Q_1 > Q_2$, and further assume that constraints IC_1 and PC_2 bind—that is, they will hold with equalities instead of weak inequalities. Use these constraints to derive lower bounds on your feasible choices of M_1 and M_2 in terms of Q_1 and Q_2.
6. Show that when IC_1 and PC_2 bind, IC_2 and PC_1 are automatically satisfied.
7. Substitute out for M_1 and M_2, using the expressions found in part (e) to express your objective function in terms of Q_1 and Q_2.
8. Write the first-order conditions for maximization, and solve them for Q_1 and Q_2.
9. Solve for M_1 and M_2.
10. What is Oceania's expected net benefit from offering this menu of contracts?
11. What general principles of screening are illustrated in the menu of contracts you found?

12. (**Optional**) Revisit Oceania's problem in Exercise S11 to see how the optimal menu of contracts found in that problem compares with some alternative contracts:
    1. If you decided to offer a single fixed-price contract that was intended to attract only the low-cost BMA, what would it be? That is, what single (Q, M) pair would be optimal if you knew BMA had a low cost?
    2. Would a high-cost BMA want to accept the contract offered in part (a)? Why or why not?
    3. Given the probability that BMA has a low cost, what would the expected net benefit to Oceania be from offering the contract in part (a)? How does your

answer compare with the expected net benefit from offering the menu of contracts you found in part (j) of Exercise S11?

4. What single fixed-price contract would you offer to a high-cost BMA?
5. Would a low-cost BMA want to accept the contract offered in part (d)? What would its profit be if it did?
6. Given your answer in part (e), what would be the expected net benefit to Oceania from offering the contract in part (d)? How does your answer compare with the expected net benefit from offering the menu of contracts you found in part (j) of Exercise S11?
7. Consider the case in which an industrial spy within BMA has promised to divulge the true per-unit cost, so that Oceania can offer the optimal single, fixed-price contract geared toward BMA’ s true type. What would Oceania’ s expected net benefit be if it knew that it was going to learn BMA’ s true type? How does your answer compare with those you found in parts (c) and (f) of this exercise and in part (j) of Exercise S11?

UNSOLVED EXERCISES

1. What problems of moral hazard and/or adverse selection arise in your dealings with each of the following? In each case, outline some appropriate incentive schemes and/or signaling and screening strategies to cope with these problems. No mathematical analysis is expected, but you should state clearly the economic reasoning underlying your suggested strategies.
   1. Your financial adviser tells you what stocks to buy or sell.
   2. You consult a real-estate agent when you are selling your house.
   3. You visit your doctor, whether for routine checkups or treatments.
2. MicroStuff is a software company that sells two popular applications, WordStuff and ExcelStuff. It doesn' t cost anything for MicroStuff to make each additional copy of its applications. MicroStuff has three types of potential customers, represented by Ingrid, Javiera, and Kathy. There are 100 million potential customers of each type, whose valuations for each application are as follows:

	WordStuff	ExcelStuff
Ingrid	100	20
Javiera	30	100
Kathy	80	0

---

   1. If MicroStuff sets separate prices for WordStuff and ExcelStuff, what price should it set for each application to maximize its profit? How much profit does MicroStuff earn with these prices?

2. What does each type of customer (Ingrid, Javiera, Kathy) buy when MicroStuff sets separate profit-maximizing prices for WordStuff and ExcelStuff?
3. Instead of selling the applications separately, MicroStuff decides always to sell WordStuff and ExcelStuff together in a bundle, charging a single price for both. What single price for the bundle would maximize its profit? How much profit does MicroStuff make selling its software only in bundles?
4. What does each type of customer buy when MicroStuff sets a single profit-maximizing price for a bundle containing WordStuff and ExcelStuff? How does this answer compare with your answer in part (b)?
5. Which pricing scheme does each customer type prefer? Why?
6. If MicroStuff sold the applications both as a bundle and separately, which products (WordStuff, ExcelStuff, or the bundle) would it want to sell to each customer type? How can MicroStuff make sure that each customer type purchases exactly the product that it intends for them to purchase?
7. What prices—for WordStuff, ExcelStuff, and the bundle—would MicroStuff set to maximize its profit? How much profit does MicroStuff make selling the products at these three prices?
8. How do your answers to parts (a), (c), and (g) differ? Explain why.

3. Consider a managerial effort example similar to the one in Section 5. The value of a successful project is $420,000; the probabilities of success are 0.5 with good supervision and 0.25 without. The manager's expected payoff equals his expected income minus the cost of his effort. His current job pays $90,000, and his cost for exerting the extra effort for good supervision on your project is $100,000.
    1. Show that inducing extra effort would require the firm to offer a compensation scheme with a negative

base salary; that is, if the project fails, the manager pays the firm an amount stipulated in the scheme.

2. How might a negative base salary be implemented in reality?
3. Show that if a negative base salary is not feasible, then the firm does better to settle for low pay and no extra effort.

4. Cheapskates is a very minor-league professional hockey team. Its facilities are large enough to accommodate all of the 1,000 fans who might want to watch its home games. It can provide two kinds of seats—ordinary and luxury. There are also two types of fans: 60% of the fans are blue-collar fans, and the rest are white-collar fans. The costs of providing each kind of seat and the fans' willingness to pay for each kind of seat are given in the following table (measured in dollars):

		Cost	WILLINGNESS TO PAY	
			Blue-Collar	White-Collar
SEAT TYPE	Ordinary	4	12	14
	Luxury	8	15	22

You may need to scroll left and right to see the full figure.

Each fan will buy at most one seat, depending on the consumer surplus he would get (maximum willingness to pay minus the actual price paid) from each kind. If the surplus for both kinds is negative, then he won't buy any. If at least one kind gives him a nonnegative surplus, then he will buy the kind that gives him the larger surplus. If the two kinds give him an equal nonnegative surplus, then the blue-collar fan will buy an ordinary seat, and the white-collar fan will buy a luxury seat.

The team owners provide and price their seating to maximize profit, measured in thousands of dollars per game. They set a price for each kind of seat, sell as many tickets as are demanded at these prices, and then provide the numbers of seats of each kind for which the tickets have sold.

1. First, suppose the team owners can identify the type of each individual fan who arrives at the ticket window (presumably by the color of his collar) and can offer him just one kind of seat at a stated price, on a take-it-or-leave-it basis. What is the owners' maximum profit, π^*, under this system?
2. Now, suppose that the owners cannot identify the type of any individual fan, but they still know the proportion of blue-collar fans. Let the price of an ordinary seat be X and the price of a luxury seat be Y. What are the incentive-compatibility constraints that will ensure that the blue-collar fans buy the ordinary seats and the white-collar fans buy the luxury seats? Graph these constraints on an X-Y coordinate plane.
3. What are the participation constraints for the fans' decisions on whether to buy tickets at all? Add these constraints to the graph in part (b).
4. Given the constraints you found in parts (b) and (c), what prices X and Y maximize the owners' profit, π_2, under this price system? What is π_2?
5. The owners are considering whether to set prices so that only the white-collar fans will buy tickets. What is their profit, π_w, if they decide to cater to only the white-collar fans?
6. Comparing π_2 and π_w, determine the pricing policy that the owners will set. How does their profit achieved from this policy compare with the case with full information, where they earn π^*?

    7. What is the “cost of coping with the information asymmetry” in part (f)? Who bears this cost? Why?
5. Redo Exercise U4 above, assuming that 10% of the fans are blue-collar fans.
6. Using the insights you gained in Exercises U4 and U5, solve Exercise U4 for the general case where a fraction B of the fans are blue-collar and a fraction $(1 - B)$ are white-collar. The answers to some parts will depend on the value of B. In these instances, list all relevant cases and how they depend on B.
7. In many situations, agents exert extra effort in order to get promoted to a better-paid position, where the reward for that position is fixed and where agents compete among themselves for those positions. Tournament theory considers a group of agents competing for a fixed set of prizes. In this case, all that matters for winning is one’s performance relative to that of the others, rather than one’s absolute level of performance.
    1. Discuss the reasons why a firm might wish to employ the tournament scheme described above. Consider the effects on the incentives of both the firm and its employees.
    2. Discuss the reasons why a firm might *not* wish to employ the tournament scheme described above.
    3. State one specific prediction of tournament theory and provide an example of empirical evidence in support of that prediction.
8. Repeat Exercise S8 with the following adjustments: Due to the departure of some of the company’s brightest engineers, the probability of the project’s success with extra managerial effort is only 0.4, and the probability of its success without extra managerial effort is reduced to 0.24.
9. (**Optional**) A teacher wants to find out how confident his students are about their own abilities. He proposes the following scheme: “After you answer this question, state your estimate of the probability that you are right. I

will then check your answer to the question. Suppose you have given the probability estimate x. If your answer is actually correct, your grade will be $\log(x)$. If your answer is incorrect, it will be $\log(1 - x)$." Show that this scheme will elicit the students' own truthful estimates—that is, if the truth is p, show that a student's stated estimate $x = p$.

10. **(Optional)** Repeat Exercise S11, but assume that the probability that BMA has a low cost is 0.6.
11. **(Optional)** Repeat Exercise S11, but assume that a low-cost BMA has a per-unit cost of 0.2, and a high-cost BMA has a per-unit cost of 0.38. Let the probability that BMA has a low cost be 0.4.
12. **(Optional)** Revisit the situation in which Oceania is procuring arms from BMA (see Exercise S11). Now consider the case in which BMA has three possible cost types: c_1, c_2, and c_3, where $c_3 > c_2 > c_1$. BMA has cost c_1 with probability p_1, cost c_2 with probability p_2, and cost c_3 with probability p_3, where $p_1 + p_2 + p_3 = 1$. In what follows, we will say that BMA is of type i if its cost is c_i, for $i = 1, 2, 3$.

    You offer a menu of three possibilities: "Supply us quantity Q_i, and we will pay you M_i," for $i = 1, 2$, and 3. Assume that more than one contract is equally profitable, so that a BMA of type i will choose contract i. To meet the participation constraint, contract i should give a BMA of type i nonnegative profit.

    1. Write an expression for the profit of a type i BMA when it supplies quantity Q and is paid M.
    2. Give the participation constraints for each BMA type.
    3. Write the six incentive-compatibility constraints. That is, for each type i, give separate expressions that state that the profit that BMA receives under contract i is greater than or equal to the profit it would receive under the other two contracts.

4. Write down the expression for Oceania’s expected net benefit, B. This is the objective function (what you want to maximize).

   Now your problem is to choose the three Q_i and the three M_i to maximize expected net benefit, subject to the incentive-compatibility (IC) and participation constraints (PC).

5. Begin with just three constraints: the IC constraint for type 2 to prefer contract 2 over contract 3, the IC constraint for type 1 to prefer contract 1 over contract 2, and the participation constraint for type 3. Assume that $Q_1 > Q_2 > Q_3$. Use these constraints to derive lower bounds on your feasible choices of M_1, M_2, and M_3 in terms of c_1, c_2, and c_3 and Q_1, Q_2, and Q_3. (Note that two or more of the cs and Qs may appear in the expression for the lower bound for each of the Ms.)
6. Prove that these three constraints—the two ICs and one PC in part (e)—will bind at the optimum.
7. Now prove that when the three constraints in part (e) are binding, the other six constraints (the remaining four ICs and two PCs) are automatically satisfied.
8. Substitute out for the M_i to express your objective function in terms of the three Q_i only.
9. Write the first-order conditions for maximization, and solve for each of the Q_i. That is, take the three partial derivatives $\partial Q_i/\partial B$, set them equal to zero, and solve for Q_i.
10. Show that the assumption made above, $Q_1 > Q_2 > Q_3$, will be true at the optimum if

$$\frac{c_3 - c_2}{c_2 - c_1} > \frac{p_1 p_3}{p_2}.$$

15■ Auctions, Bidding Strategy, and Auction Design

THE BIDDING was hot at Christie’s auction house in midtown Manhattan when Andy Warhol’s *Colored Mona Lisa* (1963) came up for sale in May 2015. The painting had been expected to sell “in the region of $35 million,” but the bidding quickly blew past that mark to a final price of $50 million, as at least three bidders outdid one another to secure this iconic masterwork.[1] Elsewhere, around the same time, a different sort of bidding contest was under way, as college students enjoyed what the *Wall Street Journal* called “one of the strongest graduate hiring seasons in recent memory.”[2] More and more graduating students were receiving multiple job offers, which gave them the opportunity to pit potential employers against one another in auctions of their own.

Anything and everything can be (and usually has been) auctioned, even gold medals honoring world-class achievement. In 2014, *USA Today* reported that Olympic gold medals “are hitting the market in record number,” most selling for about $10,000 and a famous few commanding prices well over a million.[3] Even closer to our hearts as game theorists, the gold medal celebrating John Nash’s 1994 Nobel Prize in Economics came up for auction in October 2019; the winner paid $735,000.[4]

In this chapter, we explore auctions from many angles and from the perspective of both bidders, who are competing to secure a scare resource, and auction designers, those who are in control of the scarce resource. We first clarify what is meant by the term *auction* and provide examples that illustrate some of the wide variety of interactions that can be viewed as auctions. We also consider the importance of information in auctions and some interesting auction-related phenomena, such as the *winner's curse*. Finally, we delve into both bidding strategy and auction design, considering optimal behavior across a range of different auction types.

Endnotes

- See Judd Tully, “Christie’s Scores 4th Biggest Art Auction Ever, Pulls in $1.36 Billion in 2 Nights,” Blouin Art Info International, May 14, 2015, available at https://www.blouinartinfo.com/news/story/1157180/christies-scores-4th-biggest-art-auction-ever-pulls-in-136. A video of the live bidding for *Colored Mona Lisa* is available on Christie’s Web site at http://www.christies.com/lotfinder/paintings/andy-warhol-colored-mona-lisa-5896014-details.aspx#features-videos. Return to reference 1
- Lindsay Gellman, “The Workers Who Say ‘Thanks, but No Thanks’ to Jobs: Recruiters Complain College Hires Are Leaving Them in the Lurch,” *Wall Street Journal*, July 14, 2015. Return to reference 2
- According to Ingrid O’Neil, an auctioneer at RR Auction in Boston who auctioned 14 gold medals in September 2014, “[A] bronze medal from the Summer Olympics should bring $5,000–$6,000, a silver $8,000, and a gold $10,000.” On the other hand, one of Jesse Owens’s gold medals from 1936 sold in 2013 for $1.46 million. See Karen Rosen, “Olympic Medals Hit the Market in Record Number,” *USA Today: Sports*, September 16, 2014. Return to reference 3
- The three authors had the privilege of writing a brief essay on Nash’s Nobel-winning contributions for publication in Christie’s catalog for this auction. See “Gold Medals for Games” by Avinash Dixit, David McAdams, and Susan Skeath in Christie’s *Fine Printed Books & Manuscripts including Americana: New York 25*, October 2019 [Auction catalog] (New York: Christie’s Auction House, 2019), pp. 38–39. Return to reference 4

1 WHAT ARE AUCTIONS?

An auction is a game between players competing for a scarce resource. The players in an auction are called bidders, the item they compete for is the *object*, and the action that a bidder takes to try to win the object is his *bid*. The benefit that a bidder gets when winning the object is his valuation (also called *value* or *willingness to pay*) for the object. Each bidder's strategy in the auction game is referred to as his *bidding strategy*. The cost that each bidder incurs within the auction game is referred to as his *payment*. In many, but not all, auctions, there is another player, referred to as the auction designer, who sets the rules of the game that the bidders play. There is also sometimes yet another player, the *auctioneer*, who runs the auction according to the rules laid down by the auction designer.

Real-world auctions come in so many forms and formats that studying them can be a bit bewildering for those new to the subject. We are used to thinking about auctions as mechanisms by which an object is sold, but there are also auctions in which objects are purchased (procurement auctions). Sometimes a single object is being bought or sold (single-object auctions). In other cases (multi-unit auctions), multiple identical objects are bought or sold (e.g., Treasury-bond auctions used to fund U.S. debt), and in still others (combinatorial auctions), combinations of non-identical objects are bought or sold (e.g., the consumer-goods company Procter & Gamble uses a combinatorial procurement auction to meet some of its complex procurement needs[5]).

Even if we restrict our attention to auctions of a single object—as we will do throughout most of this chapter—auctions can take seemingly countless forms in practice. Some are noisy affairs with yelling bidders[6] and fast-talking auctioneers (e.g., classic-car auctions), while others are completely silent, with bidders dropping off their bids in sealed envelopes. Some happen in places where bidders are able to inspect the merchandise (e.g., art auctions) while in others, bids are submitted over the Internet without any need for bidders to congregate. Some last for several days (e.g., eBay auctions) while others are over in seconds (e.g., fish auctions) or even milliseconds (e.g., Google Adwords auctions to determine which ads are placed next to Google search results).

The variety of games that can be interpreted as auctions is enormous, but not every game played to allocate a scarce resource is an auction. The first key feature of any auction is that bidders must be able to do something to increase their likelihood of winning the resource. When scientists compete for federal funding, they submit research proposals explaining the importance of their work, but cannot do anything to influence whether they win funding. The contest among scientists for funding is therefore not an auction—there is *no bidding*. The second key feature of any auction is that bidders must compete directly against one another. In real-estate markets without a lot of buyers, sellers typically interact with potential buyers one at a time, deciding whether to accept the current buyer's terms or wait for the next buyer to come along. Such home sales are therefore not auctions—there is *no direct competition*. As long as there is bidding and direct competition among bidders, the game being played will qualify as an auction. In the remainder of this section, we describe examples of interactions that can be analyzed as auctions, the formats that auctions can take, and the role of asymmetric information in auctions.

A. More Than Just Buying and Selling

The auction concept is extraordinarily broad. It includes situations like the previous examples in which several bidders would each like to buy or sell something—what people typically refer to as "auctions" in everyday language—but also many other sorts of interactions in which there is some conflict or competition over a scarce resource.

Consider the game that firms play when competing to hire a new employee named Anne. Suppose that Anne has been successful in her job search, so much so that three firms plan to make her an offer. Anne tells the hiring manager at each company that she would love to work there, but two other firms with equally attractive jobs are also trying to hire her. Faced with this tough decision, Anne commits to sign with whichever firm offers the highest salary. (Anne also commits not to renegotiate. Each firm's salary offer will have to be its best-and-final offer.)

This hiring contest can be interpreted as an auction in which the object up for bid is Anne herself, a future employee, and she is the auction designer. The bidders are the firms looking to hire her. Because firms make their bids without knowing what others have bid, this auction game has simultaneous moves. Auctions with simultaneous moves are called *sealed-bid auctions*. Several sorts of sealed-bid auctions are commonly used in practice, with different rules for how much the winning bidder needs to pay. The auction format that Anne is using is known as a *first-price auction* because the winning firm pays its own bid. We examine first-price auctions, along with auctions that use different rules for winning payments, in more detail later in this section.

Consider next a sales contest. Adam and Bengt are salespeople in an office that pays the employee with the most sales a bonus at the end of each year. How much each of them sells depends on how much time he spends pursuing new clients. Whoever puts in the most time will take home the bonus, but there's no way for either of them to observe how hard the other is working.

This sales contest can be interpreted as an auction where the bidders are the salespeople, the object is the bonus, and bids correspond to the amount of time spent selling. (The employer is the auction designer here, using the bonus system to encourage each salesperson to pursue new clients.) As in the previous example, bids are unobservable to other bidders, so this is a sealed-bid auction. However, in this case, everyone pays his bid, even the losers. This type of auction is known as an *all-pay auction*. We study all-pay auctions in more detail in Section 3.E.

B. Auction Formats

Auctions are broadly divided into two main categories: sealed-bid auctions, in which each bidder must decide what to bid without being able to observe anything about others' bids, and open-outcry auctions, in which the auction game plays out publicly and bidders can decide what to do at each moment based on what has happened so far. Within the category of sealed-bid auctions, three different formats—based on different rules for how much the winning bidder must pay—are used. We saw the first-price format above in our hiring contest example and the all-pay format in our sales contest example. The third format is a *second-price auction*, in which the winning bidder pays not his own bid, but the highest of the losing bids. Within the category of open-outcry auctions, there are also three different formats: the traditional *ascending-price auction*, in which offered prices rise over time, a similar *descending-price auction*, in which offered prices fall over time, and the *war of attrition*, which is the open-outcry analogue to an all-pay auction. Figure 15.1 summarizes the most notable auction formats in each category, specifying who pays what in each format.

Who pays what?	**Sealed bid**	**Open outcry**
Winner pays own bid. Losers do not pay.	First-price auction [Section 3.C]	Descending-price (“Dutch”) auction [Section 3.B]

Winner pays (exactly or approximately) the highest losing bid. Losers do not pay.	Second-price auction [Section 3.D]	Ascending-price (“English”) auction [Section 3.A]
Every bidder pays.	All-pay auction [Section 3.E]	War of attrition [Section 3.F]

FIGURE 15.1 Notable Auction Formats

C. Information in Auctions

Despite the extraordinarily wide variety of auction formats in practice, auction theorists have identified a relatively short list of key features of auction environments that play important roles in determining how sellers should design the auction and how bidders should approach their bidding problem. The most fundamental and important feature of an auction is one that you might not expect: *information*. In particular, what is most important is whether individual bidders possess private information (something they know that others do not know) about the object that is up for auction. Although the levels and types of information available to bidders in auctions can vary widely, three specific cases are of particular interest: (i) the object has an *objective value* known to all bidders; (ii) bidders have *private values* for the object; and (iii) bidders have *common values* for the object.

I. OBJECTIVE VALUEAn item being auctioned has an objective value if each bidder assigns the same value to the object and this value is known to all bidders. For instance, imagine that a game-theory professor holds an auction in class, with students as the bidders, but with a strange twist—she is going to auction off \$20. She announces that she will be conducting an auction to sell a single \$20 bill (definitely not counterfeit). Anyone can yell out a bid at any time, but each bid must be in whole dollars, must be at least \$1, and must be at least \$1 more than the previous high bid.

What would happen? There would be confusion at first, probably, as the students slowly adjusted to the strange notion of bidding money to win money. However, eventually the auction would begin with some smart aleck in the back yelling out, “I’ll bid \$1!” But it wouldn’t end there. Paying just $1 for a $20 bill is such a good deal that someone else would inevitably cry out a higher bid. And on and on, until eventually the price would get up to $19 or $20—or, if there’s another smart aleck out there looking for a laugh, perhaps $21.

No matter what, we would expect the auction to fetch approximately $20 for the $20 bill being sold. This may seem obvious, but there’s a deeper point here. Imagine that you have a stamp or coin collection—something that you know must be valuable but that you cannot precisely value on your own. The beauty of holding an auction in such a situation is that, as a seller, you don’t need to know the object’s worth! As long as bidders know how much it’s worth (and don’t collude to keep the price low), bidding competition among them will tend to drive the price up to a level reflecting its true value.

II. PRIVATE VALUESBidders in an auction have private values if (1) each bidder has private information—something he knows that others do not know—about how much the object is worth to him and if (2) knowing other bidders’ private information would not change any bidder’s willingness to pay (*valuation*) for the object. For example, imagine that our game-theory professor holds a picnic at her house after the final exam, grilling steak and shrimp for all her students to enjoy. However, to make things interesting, she holds an auction to determine who will get the shrimp (and donates all proceeds to charity). Each student’s willingness to pay depends on how much they want shrimp that day (rather than steak)—this is their private information—and the auction has private values because each student’s value depends *only* on his own hunger for shrimp.

When bidders have private values, they would still like to know others’ private information. For instance, in a first-price auction for an object that you really love (and which is worth $100 to you), you would like to know how others feel about it. If others also like the object, you might need to submit a high bid (say $80 or $90) in order to have a decent chance of winning it. On the other hand, if no one else likes it, you may be able to win the object with a very low bid (say, $20). Knowing how much others want the object has no effect on how much you want it—your value is $100 no matter what they think—only on the bid you believe you need to make to win. This is a defining feature of auctions with private values: bidders’ value assessments do not depend on others’ private information.

III. COMMON VALUESBidders in an auction have common values when the object to be auctioned has the same value to all of them, but each bidder has only an imprecise estimate of that value. In an auction with common values, bidders would all agree about the object’ s value *if* they all had the same information. However, bidders in such auctions do not have the same information, since each bidder is privy only to his own estimate of the object’ s value.

In an auction with common values, the winner is typically the bidder who has most *overestimated* the value of the object. For instance, suppose that a jar full of pennies is being auctioned. Each bidder is allowed to inspect the jar to make his own best guess of its value. Even if the bidders’ guesses are correct *on average*, the highest estimate across all the bidders is often much higher than the true value. Unless bidders submit bids that are much lower than their own estimates, the winner therefore tends to pay more for the jar than it is actually worth. This tendency of bidders to overpay in common-value auctions, which we study in detail in the next section, is known as the *winner’ s curse*.

Endnotes

- In 2006, Procter & Gamble vice president for global purchases Rick Hughes and coauthors reported that P&G was spending about $3 billion each year in combinatorial procurement auctions and saving an estimated $300 million per year relative to what it had previously been paying for the same goods and services. See Tuomas Sandholm et. al., “Changing the Game in Strategic Sourcing at Procter & Gamble: Expressive Competition Enabled by Optimization,” *Interfaces*, vol. 36, no. 1, pp. 55–68. Return to reference 5
- Mark Hinog, “Car Auction Has the Loudest ‘69’ We’ve Ever Heard Someone Scream,” *sbnation.com*, May 20, 2017, available at https://www.sbnation.com/lookit/2017/5/20/15669582/mecam-auto-auctions-69-yell-nice-nice-niiiiiiiiiice. Return to reference 6

Glossary

auction
: A game in which multiple players (called bidders) compete for a scarce resource.

bidder
: A player in an auction game.

valuation
: The benefit that a bidder gets from winning the object in an auction.

auction designer
: A player who sets the rules of an auction game.

procurement auction
: An auction in which multiple bidders compete to supply an item. Bids in a procurement auction are prices that bidders are willing to receive to supply the good. The lowest bidder wins and is paid her bid.

multi-unit auction
: An auction in which multiple identical objects are sold.

combinatorial auction
: An auction of multiple dissimilar objects in which bidders are able to bid on and win combinations of objects.

sealed-bid auction
: An auction mechanism in which bids are submitted privately in advance of a specified deadline, sometimes in sealed envelopes.

open-outcry auction
: An auction mechanism in which bids are made openly for all to hear or see.

private information
: Information known by only one player.

objective value
: An auction is called an objective-value auction when the object up for sale has the same value to all bidders and

each bidder knows that value.

private value

An auction is called a private-value auction when each bidder has private information about their own valuation of the object up for sale, but knowing others' private information would not change any bidder's own willingness to pay for the object. An important special case is when each bidder knows their own valuation but others do not.

common value

An auction is called a common-value auction when the object up for sale has the same value to all bidders, but each bidder knows only an imprecise estimate of that value.

single-object auction

An auction in which a single indivisible object is sold.

2 THE WINNER’ S CURSE

The winner’ s curse is a phenomenon often experienced in common-value auctions whereby the winner of the auction pays more (on average) than the object is worth. The winner’ s curse arises when a bidder fails to take account of the fact that, when she wins, she is likely to have made an overly optimistic estimate of the object’ s value. Fortunately, the name is not entirely apt, because the “curse” can be avoided once you understand it. The goal of this section is to help you understand the winner’ s curse well enough to avoid it.

In the early 1980s, two economics professors, Max Bazerman and William Samuelson, ran an auction experiment with MBA students at Boston University.[7] In each of 12 class sessions, students (from 34 to 54 per class) participated in a first-price auction for a jar full of pennies, nickels, or other small items such as paper clips with an assigned value. Students would first write down their best guesses about the value of the items in the jar, then submit their bids. Results were revealed at the end of the term, and the winner of each auction was paid the value of the items in the jar (in cash, not pennies or paper clips) minus their winning bid. As it turned out, students systematically underestimated the value of the jars being auctioned, guessing, on average, that the jars had a value of $5.13, whereas in fact each jar was worth exactly $8. Nonetheless, the average winning bid was $10.01—$2.01 *more* than the actual value. So, at the end of the term, the winning students wound up paying the professors: a winner’ s curse!

What happened? Even though students’ guesses about the value of the jar were, on average, on the low side, each class was big enough that there were always a few students who overestimated the jar’ s value by a substantial amount.

Because these students had so overestimated the jar’s value, they were at risk of bidding, and hence paying, more than its true value. Thus, an auction’s winner typically is not an average bidder, but rather one who has *overestimated* the true value of the object at auction. This explains why winners of auctions with common values routinely overpay—they don’t account for the fact that, when they win the auction, they are likely to have overestimated the object’s value.

Consider another example from the world of business. In May 2018, PayPal paid $2.2 billion to acquire iZettle, a financial-technology company whose smartphone-connected payment systems make it easy for small businesses to securely process credit-card and debit-card payments. There’s no way for us to know whether PayPal got a good deal for its $2.2 billion, but let’s consider how the winner’s curse *could* have been at play in this transaction. (All numbers in this example other than the $2.2 billion purchase price are made up. We have no inside information about this transaction.)

iZettle made its name with innovative devices like its 2015 credit-card reader, which connects to iPhones through the audio jack, and other technologies with the potential to transform commerce—but someone could come out tomorrow with an even cleverer approach and render iZettle’s technology obsolete. The company’s value could therefore be extremely high, or close to nothing at all. In the face of such uncertainty, how should PayPal think about the acquisition and what price to offer?

To keep things simple, suppose that iZettle’s technology is either Strong (will continue to dominate the marketplace) or Weak (will be made obsolete in the near future). If the technology is Strong, iZettle’s “stand-alone value” if it doesn’t get acquired by PayPal is $6 billion, but if the technology is Weak, iZettle’s stand-alone value is $0. Moreover, due to synergies with PayPal’s existing businesses, iZettle's value to PayPal is 150% of its stand-alone value ($9 billion if the technology is Strong or $0 if it is Weak).

Suppose for a moment that PayPal and iZettle are equally uncertain about the technology's potential, each believing that it has a 50% chance of being Strong. iZettle's expected stand-alone value would then be 50% × $6 billion = $3 billion, while PayPal's expected value for the company would be 150% × $3 billion = $4.5 billion. This leaves plenty of room for negotiation (see Chapter 17 for more on negotiation games), as any price between $3 billion and $4.5 billion would make both sides better off.

Our analysis so far hinges on the assumption that iZettle and PayPal share the same belief about the technology's potential—that neither side has private information. But what if iZettle's founders (Jacob de Geer and Magnus Nilsson) have a better sense of their own technology than PayPal's negotiators? In particular, suppose that de Geer and Nilsson have a *feeling* about their technology: a "good feeling" half of the time that the technology's likelihood of being Strong is 80%, and a "bad feeling" the other half of the time that the technology's likelihood of being Strong is only 20%. (The overall likelihood of its being Strong is still 50% = ½ × 80% + ½ × 20%.)

To see how private information changes the game here and creates the potential for a winner's curse, imagine a scenario in which PayPal enters the negotiation with a lowball bid equal to just half of its estimate of iZettle's value. As discussed above, PayPal's estimate (based on its belief that the technology has a 50% chance of being Strong) is $4.5 billion, so its lowball offer is $2.25 billion. PayPal is expecting this initial offer to be rejected, but then iZettle accepts!

What this tells us is that *de Geer and Nilsson must have a bad feeling* about the technology. Why? If they had had a good feeling—if they had thought the technology had an 80% chance of being Strong—their estimate of the company's stand-alone value would have been 80% × $6 billion = $4.8 billion, and they would have laughed off an offer of $2.25 billion. But if they have a bad feeling, their estimate of the company's stand-alone value would be only 20% × $6 billion = $1.2 billion, and $2.25 billion would look pretty good to them! Of course, once PayPal owned the company and came to understand why de Geer and Nilsson had a bad feeling, they would wish that they hadn't acquired the company at all, since its expected value to them would then be only 150% × $1.2 billion = $1.8 billion. They would have made a lowball offer but wound up overpaying by $445 million—a winner's curse.

The mathematics associated with calculating optimal (equilibrium) bidding strategies in auctions with common values is complex, not least because other bidders are simultaneously making similar calculations. Although the math behind bidding equilibria is beyond our scope here, we can provide you with some general advice.[8] If you imagine yourself in the position of PayPal bidding for iZettle, you can see that the question, "Would I be willing to purchase this company for $2.2 billion, given what I know before submitting my bid?" is very different from the question, "Would I still be willing to purchase this company for $2.2 billion, given what I know before submitting my bid *and* given the knowledge that $2.2 billion is good enough to win?" Whether you find yourself in a one-on-one negotiation or in an auction with several competing bidders, it is the second question that reveals correct strategic thinking, because you win with any given bid only when all others bid less—and that happens only if all others have lower estimates of the value of the object than you do.

If you fail to use your game-theoretic training and do not take the winner's curse into account, you should expect to consistently lose substantial amounts. But at least you won't be alone in this misfortune, as the winner's curse is all around us. Winners of auctions for oil- and gas-drilling rights took substantial losses on their leases for many years.[9] Baseball players who as free agents went to new teams were found to have been overpaid in comparison with those who re-signed with their old teams.[10] On the other hand, if you understand the winner's curse and take it into account, you can avoid it. True, you may win less often, but at least you will truly "win" when you do.

Endnotes

- Max H. Bazerman and William F. Samuelson, "I Won the Auction but Don't Want the Prize," *Journal of Conflict Resolution*, vol. 27, no. 4 (December 1983), pp. 618–34. Return to reference 7
- See Steven Landsburg, *The Armchair Economist* (New York: Free Press, 1993), p. 175. Return to reference 8
- A comprehensive study estimated that in 1,000 lease auctions in the 1950s and 1960s, 12 of the 18 major bidders consistently overbid; the median firm in this group overbid by more than 30%, and one firm (Texaco) bid *seven times* what it ought to have bid. See Kenneth Hendricks, Robert H. Porter, and Bryan Boudreau, "Information, Returns, and Bidding Behavior in OCS Auctions: 1954–1969," *Journal of Industrial Economics*, vol. 35, no. 4 (June 1987), pp. 517–54. Return to reference 9
- James Cassing and Richard W. Douglas, "Implications of the Auction Mechanism in Baseball's Free Agent Draft," *Southern Economic Journal*, vol. 47, no. 1 (July 1980), pp. 110–21. Return to reference 10

Glossary

winner’s curse

A situation in a common-value auction where the winner fails to take account of the fact that when she wins, she is likely to have made an overly optimistic estimate of the object’s value. Bidders who correctly anticipate this possibility can avoid the winner’s curse by lowering their bids appropriately.

3 BIDDING IN AUCTIONS

In May 2016, 10 bottles of Chateau Mouton Rothschild 1945 Bordeaux sold at auction for $343,000 at Sotheby's New York, a record amount for a wine considered by many to be one of the finest ever bottled. Such eye-popping prices might turn some off to wine auctions, but according to *Bloomberg* magazine, "auctions can be a good way to find hard-to-get vintages and even to score bargains."[11] Better still, bidders in wine auctions these days can participate remotely over the Internet. So, you can bid to win that lot of Chateau Margaux 1952 Bordeaux you've been eyeing recently without leaving the comfort of your living-room couch.

Let's suppose that you actually do decide to bid for a lot of wine and that your valuation for the wine is V. How should you bid? The answer depends on what auction format is being used to sell the wine.

A. Ascending-Price ("English") Auction

For most people, the word *auction* evokes an image of an ascending-price auction—at a traditional auction house or perhaps at an online auction site like eBay—in which bidders "call out" ever-higher prices until finally a price is reached that no one is willing to top, at which point the object is sold to the final bidder. Real-world ascending-price auctions each have their own particular rules about exactly what happens during the auction and what bidders are allowed to do. We don't want to get bogged down in such details, and so will focus on an idealized auction game that auction theorists refer to as the "English auction"—named in honor of the English auction houses that have long dominated the world of book and art auctions.[12]

In an English auction, the price is increased continuously over time, each bidder decides when to "drop out," and the winner is whoever stays in the auction longest. More precisely, the auction begins as the seller announces a minimal acceptable price r, known as the reserve price, and potential bidders decide whether to participate. If no one participates, the auction ends, and the object is returned unsold to the seller. If only one bidder participates, that bidder wins and pays the reserve price. Finally, in the most interesting case, when multiple bidders participate, the price increases continuously over time, and each bidder decides when to drop out of the auction. (The price at which a bidder drops out is his drop-out price.) The price stops increasing when only one bidder remains active in the auction, at which point that bidder wins and pays the final price.

Should you find yourself in an English auction having value V for the object up for sale, your optimal bidding strategy is simple: Don't participate at all if the reserve price r exceeds your value V; otherwise, enter the auction and stay in the bidding until the price reaches V. Why is this your best strategy? First, consider the question of whether you should enter the auction. If the reserve price exceeds your value and you enter, you might wind up winning, which is bad news, since you will then have to pay more than the object is worth to you. On the other hand, so long as the reserve price is less than your value, there's a chance that you will be able to win the object at a price less than your value. So, you should enter the bidding only if $r < V$. Moreover, so long as the current price p is less than your value V, you stand to gain by staying in the bidding, since there's a chance that all other bidders will drop out in the near future, leaving you to win the object at a price less than your value. If, however, $p \geqslant V$ and you stay in the auction, you run the risk of winning at a price that exceeds your value. Thus, your optimal strategy is to wait until the price reaches V and then drop out. In fact, the argument here shows that *dropping out at price* $p = V$ *is a weakly dominant strategy*; this choice remains optimal no matter what bidding strategies others use.

Given that all bidders remain in the bidding until the price reaches their value, what is the final price in an English auction? For concreteness, suppose that Anna has the highest value, V_A, Ben has the second-highest value, V_B, and the *bid increment*, the minimum amount that any new bid must increase the current price, is $D > 0$. So as long as the current price p is less than $V_B - D$, both Anna and Ben are willing to increase the bid by D in order to become the current high bidder. Thus, the auction cannot possibly end at any price less than $V_B - D$. On the other hand, once the current price p exceeds $V_B - D$, Ben is no longer willing to increase the bid. Therefore, the highest that the auction price can possibly go is $V_B + D$, a possibility that could arise, for instance, if Anna was the current high bidder at price $V_B - D$, Ben raised the bid to V_B, his maximum willingness to pay, and then Anna topped that with a bid of $V_B + D$, at which point the auction would end. Note that, no matter what, the final price must be between $V_B - D$ and $V_B + D$; it will be *approximately* equal to the second-highest bidder value when the bid increment is small.

The English auction presents bidders with a single, simple decision regarding when to drop out of the bidding. However, most real-world ascending-price auctions are more complex. For instance, in a Sotheby's-style art auction, bidders can engage in jump bidding (submitting a bid that is significantly higher than the previous bid, and well beyond whatever minimum bid increment exists), as a way of signaling their interest in the object and perhaps deterring others from competing further. Bidders can also mask their participation in the auction by submitting secret instructions ahead of time to the auctioneer. And the auctioneer can employ a shill bidder (someone employed to continue to outbid a legitimate bidder) to drive up the price even after only one serious bidder remains. Similarly, in eBay auctions, bidders can wait until the last moment to submit their bids, a tactic known as sniping, or they can use eBay's "proxy bidder" to continue to bid on their behalf when they submit a maximum price that exceeds the minimum bid increment. For more on these additional features of real-world ascending-price auctions and the appropriate ways to bid in such auctions, see the references provided in Section 5.

B. Descending-Price ("Dutch") Auction

The Aalsmeer Flower Auction building in the Netherlands, known as the "Wall Street of flowers," is the second-largest building in the world (second only to the Pentagon) and consists of 128 acres filled with flowers from Ecuador, Ethiopia, and everywhere in between. About 20 million flowers are sold daily at Aalsmeer. With so many flowers to sell, the auction has to be fast, and it is; each lot goes in about four seconds.[13] Bidders sit in stadium seating facing a light display that looks like a big clock, with a button (actually three buttons: red, yellow, and green) next to their seats that they use to bid. Each point on the "clock" corresponds to a price. The light starts at a high price that no one is willing to pay, then rapidly moves around the clock to lower and lower prices until someone pushes his button. The light then stops moving, the auction ends, and the bidder who pushed his button first gets the lot at that final price.

Should you find yourself bidding at Aalsmeer, or in some other descending-price (or Dutch) auction with value V for the object up for sale, how should you bid? Figuring out the optimal price to bid in a Dutch auction involves some challenging mathematics, but we can gain some reasonable insight by working through a simple numerical example.

Imagine that you are willing to pay \$100 for a lot of flowers. You've been observing the auctions at Aalsmeer for some time now, and you've noticed that the price at which bidders jump in for similar flowers ranges from \$50 to \$150 (with every price in this range being equally likely). Now the auction has begun, at a starting price of \$200, and the price is quickly dropping. When should you jump in to win? Obviously not at any price above \$100, because then you would have to pay more than the flowers are worth to you. So you wait and, lo and behold, the price falls to $100 and no one has bid yet. You could push your button to win at $100, but since you would be paying your full value for the flowers, you would walk away from the auction with zero profit! So, as the price hits $100, you should not hit your button at that moment. You should wait at least a little longer. But how long?

Let' s step back for a moment and consider this bidding problem from your perspective *before the auction begins*. Even though the auction has not yet begun, you can think ahead to how the auction might proceed and decide, for each price that might be reached, whether you will want to jump in and win the object or wait and let the price fall further. Let P denote the first price at which you will want to jump into the bidding. If someone else jumps in before price P is reached, you will lose the auction and get zero profit; because others' jump-in prices range uniformly from 50 to 150, the probability that this happens is

$$\frac{150-P}{150-50}=\frac{150-P}{100}.$$

Alternatively, if no one jumps in before price P is reached, you will win the object, pay P, and get profit equal to your value $100 minus the price P; this happens with probability

$$\frac{P-50}{100}.$$

Overall, your expected profit [denoted by $E\Pi(P)$] equals the probability that you win multiplied by your profit when you win:

$$E\Pi(P)=\frac{P-50}{100}\times(100-P)=\frac{-P^2+150P-5{,}000}{100}.$$

Using the methods introduced in Chapter 5 to maximize a quadratic equation, we find that $E\Pi(P)$ is maximized at $P^* = 75$;[14] so, when you are willing to pay \$100, you should wait until the price reaches \$75 to jump in.

The price at which a bidder should optimally jump in to win the object depends on his belief about others' bidding strategies. To see the point, reconsider the previous numerical example and continue to suppose that you are willing to pay \$100, but now suppose that you believe that the price at which someone else will jump in is uniformly distributed from \$0 to \$100 (rather than from \$50 to \$150). The probability that no one else jumps in before price P is now $\frac{P}{100-0}=\frac{P}{100}$, higher than before. Your expected profit from jumping in at price P is now

$$E\Pi(P)=\frac{P}{100}\times(100-P)=\frac{-P^2+100P}{100},$$

which is maximized at $P^* = 50$.

Such calculations can be difficult to perform in practice, especially if you do not know others' bidding strategies. Even so, we can glean some basic insights from the logic behind these calculations. First, you should always wait until the price has fallen below your valuation before jumping in to win the object. Second, how far you should allow the price to drop below your value depends on how you assess the strength of the bidding competition—you should wait longer to jump in when the competition is weaker.

C. First-Price Auction

In a first-price auction, all bidders submit their bids simultaneously, and the highest bidder wins and pays his bid. First-price auctions are widely used for everything from selling stamps to buying fighter jets. Figuring out the optimal price to bid in a first-price auction involves some challenging mathematics, but as we did with the Dutch auction, we can gain some insight by working through a simple numerical example.

Imagine that you are willing to pay \$100 for a batch of stamps for sale in a first-price auction. You've been observing similar auctions for some time now,[15] and you've noticed that the winning price for similar batches of stamps ranges from \$50 to \$150 (with every price in this range being equally likely). What should you bid? Obviously not anything above \$100, because then you would have to pay more than the stamps are worth to you. And obviously not \$100, since then you would pay your full value when you win and have no profit. So, you should bid less than \$100. But how much less?

Suppose that you shade your bid (that is, bid below your value), bidding $P < 100$. Your chance of winning the auction depends on how likely it is that your P exceeds the next highest bid, which we know is at least \$50. The probability that you win when the range of winning prices is between \$50 and \$150 is then $\dfrac{P-50}{150-50}=\dfrac{P-50}{100}$, and when you win, your profit will be $100 - P$. Overall, then, your expected profit when bidding P equals

$$E\Pi(P) = \frac{P-50}{100} \times (100-P),$$

which is maximized at $P = 75$ (as we showed in footnote 14). So, if your value is \$100, you should shade your bid and submit a bid of \$75. As in the Dutch auction, your optimal bid will be at some price less than your true valuation that depends on how likely you are to win at that price along with the profit you would earn.

You probably noticed that what we did here in determining your optimal bid in the first-price stamp auction generated a calculation identical to the one we solved to find your optimal bid in the descending-price flower auction above. We were able to do that because the first-price auction and descending-price auction are strategically equivalent! Any game-theoretic analysis of one of these auction formats applies word for word to the other format, as long as you replace the phrase "jump-in price" with "sealed bid," or vice versa.

To see why the descending-price auction (DPA) and first-price auction (FPA) are strategically equivalent, consider how the outcome of each game depends on players' chosen strategies: The bidder with the highest jump-in price wins and pays their jump-in price in the DPA, while the bidder with the highest sealed bid wins and pays their sealed bid in the FPA. So, if we were to create a game table for each auction, they would look exactly the same—the only difference being that strategies are called "jump-in prices" in one auction format and "sealed bids" in the other.

Of course, there may be other factors outside of the game itself that cause bidders to behave differently in descending-price auctions than they do in first-price auctions. For instance, perhaps seeing the other bidders sitting nearby causes some bidders to bid more aggressively at Aalsmeer than they would in an otherwise equivalent sealed-bid auction. These factors are obviously important practical considerations for real-world auctioneers, but outside the scope of our analysis here.

D. Second-Price Auction

In a second-price auction, all bidders submit their bids simultaneously, and the highest bidder wins and pays the second-highest bid. The second-price auction is not widely used in practice,[16] but is famous nonetheless as a special case of the Vickrey auction, which won William Vickrey the Nobel Prize in economics in 1996. (The Vickrey auction in its most general form is difficult to explain in words, but when there is only a single object for sale, it reduces to the second-price auction.)

What makes the second-price (Vickrey) auction fascinating is that *bidders have a weakly dominant strategy to submit bids equal to their true valuations* (that is, to employ truthful bidding). To see why, let's work through a simple example. Suppose you are a collector of antique china, and you have discovered that a local estate auction will be selling off a nineteenth-century Meissen Blue Onion tea set in a second-price auction. As someone experienced with vintage china but lacking this particular set for your collection, you are willing to pay up to $3,000 for it. However, you have no idea what other bidders' values for the set will be. If they are inexperienced, they may not realize the considerable value of the set. On the other hand, if they have sentimental attachments to Meissen or the Blue Onion pattern, they may value it more highly than you.

According to Vickrey's famous result, bidding truthfully (for you, bidding $3,000) is your weakly dominant strategy. To see why, suppose that you were to bid some amount $B <$ $3,000 and let P denote the highest bid submitted by any other bidder in the auction. (We use notation P here to remind you that P is the *price* that you will pay if you wind up winning the auction.) If $P >$ $3,000, it doesn't matter whether you bid B or \$3,000, because you will lose the auction either way. Similarly, if $P < B$, it doesn't matter whether you bid B or \$3,000, because you will win the auction and pay P either way. However, if $B < P <$ \$3,000, you win and pay P if you bid \$3,000 (for positive profit \$3,000 − P) but lose with bid B (for zero profit). So, shading your bid always leaves you worse off than if you bid truthfully. (A similar argument shows that bidding more than your value is also weakly dominated by truthful bidding.)

Vickrey's remarkable finding that truthful bidding is a dominant strategy in second-price auctions has many other applications. For example, if each member of a group is asked what she would be willing to pay for a public project that will benefit the whole group, each has an incentive to understate her own contribution—to become a free rider on the contributions of the rest. We have already seen examples of such effects in the collective-action games of Chapter 11. A variant of the Vickrey scheme can elicit the truth in such games as well.

There is an interesting connection between the second-price auction and the English (ascending-price) auction considered in Section 3.A. In both auction formats, each bidder has an incentive to bid or stay in the bidding up to her private value, with the end result that the bidder with the highest value wins and pays the value of the highest losing bidder (or the minimum bid increment above that losing value). Consequently, when bidders have private values, second-price and English auctions have identical equilibrium outcomes. Overall, then, the four auction formats we have considered thus far fall into two basic groups. In the first group are the first-price auction and the Dutch auction. These formats are strategically equivalent and hence always generate identical equilibrium outcomes. In the second group are the second-price auction and the English auction, which generate identical equilibrium outcomes when bidders have private values. The second-price and English auctions are not strategically equivalent, however, and can generate different equilibrium outcomes when bidders have common values.

E. All-Pay Auction

The contest among firms and other special interests to influence policy in the United States and other countries can be viewed as an auction in which each firm invests (by giving money to politicians and by attempting to sway public opinion) to increase the likelihood of getting its own way. Because all participants must pay their bids, whether they win or lose, such sealed-bid auctions are called all-pay auctions.

Lobbying and bribing of public officials is the most famous example of the all-pay auction in action, but other examples can be found in any context where people compete by exerting effort. Consider the Olympic Games. As you read this, thousands of elite athletes around the world have put their lives on hold to train full-time for the Olympics. Many will fail to qualify for their national teams, and all but one in each sport will fail to win a gold medal. But whether they win or lose, they all pay the price of lost income and lost opportunities during training. The workplace tournaments we discussed in Chapter 14, Section 6.B, are similar. Once you start thinking along these lines, you will realize that all-pay auctions are, if anything, more common in real life than situations resembling the standard formal auctions where only the winner pays!

How should you bid (that is, what should your strategy be for expenditure of time, effort, and money) in an all-pay auction? Once you decide to participate, your bid is wasted unless you win, so you have a strong incentive to bid very aggressively. In experiments using this auction format, the sum of all the bids often exceeds the value of the prize by a large amount, and the auctioneer makes a handsome profit.[17] In that case, everyone submitting extremely aggressive bids cannot be the equilibrium outcome; it seems wiser to stay out of such destructive competition altogether. But if everyone else did that, then one bidder could walk away with the prize for next to nothing; thus, not bidding cannot be an equilibrium strategy either. This analysis suggests that the equilibrium lies in mixed strategies.

Consider a specific all-pay auction with n bidders. To keep the notation simple, we choose units of measurement so that the prize has objective value equal to 1. Bidding more than 1 is sure to bring a loss, so we restrict bids to those between 0 and 1. It is easier to let the bid be a continuous variable x, where x can take on any (real) value in the interval [0, 1]. Because the equilibrium will be in mixed strategies, each person's bid, x, will be a continuous random variable. Because you win the object only if all other bidders submit bids below yours, we can express your equilibrium mixing strategy as $Prob(x)$, the probability that your bid takes on a value less than x; for example, $Prob(½) = 0.25$ would mean that your equilibrium strategy entailed bids below ½ one-quarter of the time (and bids above ½ three-quarters of the time).[18]

As usual, we can find the mixed-strategy equilibrium by using an indifference condition. Each bidder must be indifferent about the choice of any particular value of x, given that the others are playing their equilibrium mixes. Suppose you, as one of the n bidders, bid x. You win if all of the remaining $(n - 1)$ bidders are bidding less than x. The probability of anyone else bidding less than x is $Prob(x)$; the probability of two others bidding less than x is $Prob(x) \times Prob(x)$, or $[Prob(x)]^2$; the probability of all $(n - 1)$ of them bidding less than x is $Prob(x) \times Prob(x) \times Prob(x)$ . . . multiplied $(n - 1)$ times, or $[Prob(x)]^{n-1}$. Thus, with a probability of $[Prob(x)]^{n-1}$, you win 1. Remember that you pay x no matter what happens. Therefore, your net expected payoff for any bid of x is $[Prob(x)]^{n-1} - x$. But you could get 0 for sure by bidding 0. Thus, because you must be indifferent about the choice of any particular x, including 0, the condition that defines the equilibrium is $[Prob(x)]^{n-1} - x = 0$. In a full mixed-strategy equilibrium, this condition must be true for all x. Therefore, the equilibrium mixed-strategy bid is $Prob(x) = x^{1/(n-1)}$.

A couple of sample calculations will illustrate what is implied here. First, consider the case in which $n = 2$; then $Prob(x) = x$ for all x. Therefore, the probability of bidding a number between two given levels x_1 and x_2 is $Prob(x_2) - Prob(x_1) = x_2 - x_1$. Because the probability that the bid lies in any range is simply the length of that range, any one bid must be just as likely as any other bid. That is, your equilibrium mixed-strategy bid should be random and uniformly distributed over the whole range from 0 to 1.

Next, let $n = 3$. Then $Prob(x) = \sqrt{x}$. For $x = \frac{1}{4}$, $Prob(x) = \frac{1}{2}$; so the probability of bidding $\frac{1}{4}$ or less is $\frac{1}{2}$. The bids are no longer uniformly distributed over the range from 0 to 1; they are more likely to be in the lower end of the range.

Further increases in n reinforce this tendency. For example, if $n = 10$, then $Prob(x) = x^{1/9}$, and $Prob(x)$ equals $\frac{1}{2}$ when $x = (\frac{1}{2})^9 = {}^1/_{512} = 0.00195$. In this situation, your bid is as likely to be smaller than 0.00195 as it is to be anywhere within the whole range from 0.00195 to 1. Thus, your bids are likely to be very close to 0.

Your average bid should correspondingly be smaller the larger the number n. In fact, a more precise mathematical calculation shows that if everyone bids according to this strategy, the average or expected bid of any one player will be just $(1/n)$.[19] With n players bidding, on average, $1/n$ each, the total expected bid is 1, and the auctioneer makes zero expected profit. This calculation provides more precise confirmation that the equilibrium strategy eliminates overbidding.

The idea that you should bid closer to 0 when the total number of bidders is larger makes excellent intuitive sense, and the finding that equilibrium bidding eliminates overbidding lends further confidence to the theoretical analysis. Unfortunately, many people in actual all-pay auctions either do not know or forget this theory and bid to excess.

Interestingly, philanthropists have figured out how to take this tendency to overbid and harness it for social benefit. Building on the historical lessons learned from prizes offered in 1919 by a New York hotelier for the first nonstop transatlantic flight (won by Charles Lindbergh in 1927) and even earlier, in 1714, by the British government for a method to precisely measure longitude for sea navigation (eventually awarded to John Harrison in the 1770s), several U.S. and international foundations have begun offering incentive prizes for various socially worthwhile innovations. One foundation in particular, the XPRIZE Foundation, has as its sole purpose the provision of incentive prizes; its first prize was awarded in 2004 for the first private space flight. More recently, the $20 million Carbon XPRIZE seeks to promote new technologies that can convert CO_2 emissions into valuable products, such as building materials or alternative fuels. Ten finalists were selected in September 2018, each with a proven new technology to convert CO_2 into everything from carbon nanotubes to methanol and concrete; winners will be announced in March 2020, after this book has gone to print. Some foundation experts estimate that as much as 40 times the amount of money that would otherwise be devoted to a particular innovation gets spent when incentive prizes are available. Thus, the tendency to overbid in all-pay auctions can actually have a beneficial effect on society (if not on the individual pursuing the prize).[20]

F. War of Attrition

A war of attrition is a contest between multiple players in which each player decides when to retreat from the contest, the victor is whoever remains the longest, and choosing to remain longer in the contest is costly for each player. A war of attrition can be viewed as an auction, with the *bidders* being the contestants, the *object* up for auction being victory, each player's *bid* being how long he remains in the contest before retreating, and each player's *payment* corresponding to the costs[21] he incurs by remaining in the contest as long as he did.

The war of attrition is an open-outcry analogue to the all-pay auction, but the connection between the war of attrition and the all-pay auction is not as close as the connection between the Dutch and first-price auction formats (which are strategically equivalent) or the connection between the English and second-price auction formats (which lead to identical equilibrium outcomes when bidders have private values). In the war of attrition, each loser pays his full bid, but the winner only has to pay the highest losing bid. By contrast, in the all-pay auction, all bidders pay their full bids. For example, suppose that two firms are seeking to bribe a corrupt official. If the firms submit their bribes all at once in sealed envelopes, then this game would be an all-pay auction, with the corrupt official pocketing all of the cash in both envelopes. On the other hand, if the firms paid the official over time in small installments, then the game would be a war of attrition, and the winning firm would only need to pay the amount paid by the losing firm. For example, suppose that Firm A "bids" \$200 and Firm B "bids" \$150 in both games. The official would pocket \$350 = \$200 + \$150 in the all-pay auction, but only \$300 = \$150 + \$150 in the war of attrition. Because of this difference in the *payment rules* of these two auctions, equilibrium bidding strategies in the all-pay auction will not be the same as those in the war of attrition.

We examined the war of attrition (also called dynamic chicken and brinkmanship with two-sided asymmetric information) in great detail earlier in the book, in Chapter 9, Section 7, and in Chapter 13. While we will not repeat those analyses here, it is worth repeating some of the main insights that emerged from them. First, as a war of attrition continues with neither side retreating, each side should infer that the other player is more likely to be Tough (with a lot to gain from victory and/or with low costs associated with continuing the contest). Second, because a war of attrition continues only if both sides are relatively tough, a war of attrition that has already lasted a long time should be expected to last a long time more.

Endnotes

- Elin McCoy, "How to Buy Wine at Auction—and Why You Should," *Bloomberg*, September 2, 2016, available at https://www.bloomberg.com/news/articles/2016-09-02/how-to-buy-wine-at-auction-and-why-you-should. Return to reference 11
- For more on the history and famous ambience of English auctions, see Marion Laffey Fox, "Inside the Great English Auction Houses," *Christian Science Monitor*, April 17, 1984. Return to reference 12
- See Martha Stewart's video *Aalsmeer Flower Auction In Amsterdam*, available at https://www.marthastewart.com/918460/aalsmeer-flower-auction-amsterdam, for an excellent overview of the auction process at Aalsmeer. Return to reference 13
- Completing the square, we can rewrite $-P^2 + 150P - 5{,}000$ as $-(P - 75)^2 + (75)^2 - 5{,}000$, which is maximized when $P - 75 = 0$, or when $P = 75$. Return to reference 14
- A simplifying assumption here is that other bidders do not change their behavior once you decide to participate in the auction. Of course, once there is another bidder (you) in the auction, their bidding incentives will naturally change. We account for this in the appendix to this chapter when deriving equilibrium bidding strategies. Return to reference 15
- Lawrence M. Ausubel and Paul Milgrom, "The Lovely but Lonely Vickrey Auction," in Peter C. Cramton, Yoav Shoham, and Richard Steinberg, *Combinatorial Auctions* (Cambridge, Mass.: MIT Press, 2006), pp.17–40. Return to reference 16
- One of us (Dixit) has auctioned $10 bills to his Games of Strategy class and made a profit of as much as $60 from a 20-student section. At Princeton, there is a tradition of giving the professor a polite round of applause at the

end of a semester. Once Dixit offered \$20 to the student who kept applauding continuously the longest. This is an example of an unusual open-outcry, all-pay auction with payments in kind (applause). Although most students dropped out after 5 to 20 minutes, three went on applauding for 4½ hours! Return to reference 17

- *Prob*(x) is called the *cumulative probability distribution function* for the random variable x. The more familiar probability density function for x is its derivative, $Prob'(x) = prob(x)$. Then $prob(x)\,dx$ denotes the probability that the variable takes on a value in a small interval from x to $x + dx$. Return to reference 18
- The expected bid of any one player is calculated as the expected value of x, by using the probability density function, $prob(x)$. In this case, $prob(x) = Prob'(x) = [1/(n - 1)]x^{(2-n)/(n-1)}$, and the expected value of x is the sum, or integral, of this from 0 to 1, namely, $\int x\, prob(x)\,dx = 1/n$. Return to reference 19
- For more on incentive prizes, see Matthew Leerberg, "Incentivizing Prizes: How Foundations Can Utilize Prizes to Generate Solutions to the Most Intractable Social Problems," Duke University Center for the Study of Philanthropy and Voluntarism Working Paper, Spring 2006. Information on the X Prize Foundation is available at www.xprize.org. Return to reference 20
- In many applications, the cost of remaining in the contest arises from the continued *risk* of some bad outcome, such as the risk of nuclear war in the U.S.-USSR confrontation studied in Chapter 13. Even though nuclear war did not occur, the United States and the Soviet Union each bore the risk that nuclear war could have erupted—a risk either side could have reduced by retreating earlier in the confrontation. Return to reference 21

Glossary

ascending-price auction
: An open-outcry auction in which prices are announced in increasing order either by an auctioneer (in the case of an English auction) or by bidders themselves (in the case of jump bidding). The last person to bid or accept the announced price wins the auction and pays that price.

English auction
: A type of *ascending-price auction* in which the auctioneer calls out a sequence of increasing prices, bidders decide when to drop out of the bidding, and the last bidder remaining pays the last announced price.

reserve price
: The minimum price set by the seller of an item up for auction; if no bids exceed the reserve, the item is not sold.

drop-out price
: In an English auction, the price at which a bidder drops out of the bidding.

jump bidding
: Submitting a bid that is significantly higher than the previous bid and well beyond whatever minimum bid increment exists.

shill bidder
: A fake bidder created by sellers at an auction to place fictitious bids for an object they are selling.

sniping
: Waiting until the last moment to make a bid.

descending-price auction
: An open-outcry auction in which the auctioneer announces possible prices in descending order. The first person to accept the announced price wins the auction and pays that price. Also called **Dutch auction**.

Dutch auction

Same as a **descending-price auction**.

auction

A game in which multiple players (called bidders) compete for a scarce resource.

first-price auction

A sealed-bid auction in which the highest bidder wins and pays the amount of her bid.

shading

A strategy in which bidders bid slightly below their true valuation of an object.

second-price auction

A sealed-bid auction in which the highest bidder wins the auction but pays a price equal to the value of the second-highest bid; a special case of the *Vickrey auction*.

Vickrey auction

An auction design proposed by William Vickrey in which truthful bidding is a weakly dominant strategy for each bidder. When a single object is sold, the Vickrey auction is the same as the *second-price auction*.

truthful bidding

A practice by which bidders in an auction bid their true valuation of an object.

all-pay auction

An auction in which each person who submits a bid must pay her highest bid amount at the end of the auction, even if she does not win the auction.

war of attrition

A contest between multiple players in which each player decides when to retreat, the victor is whoever remains the longest, and choosing to remain longer is costly for each player.

4 AUCTION DESIGN

Suppose now that, instead of being one of the many bidders at a particular auction, you are designing an auction to sell an object you already own, with the goal of maximizing your expected revenue. What auction format should you use? In this section, we focus on comparing the two most commonly used auction formats: the English auction and the first-price/Dutch auction. [We write *first-price/Dutch* to remind you that the first-price (sealed-bid) auction and the Dutch (descending-price) auction are strategically equivalent.] To provide you with some insight into the problem, we begin by exploring a simple numerical example.

Suppose there are two bidders (denoted by $i = 1, 2$) who will take part in your auction. These bidders have private values, V_i, that are uncorrelated with each other and uniformly distributed over the interval [0, 1]. (*Uniformly distributed* means that each bidder's private value is equally likely to be any number in the specified interval—here, from 0 to 1. *Uncorrelated* means that knowing one bidder's value tells you nothing about the other bidder's likely value; thus, each bidder has no idea what the other bidder's value is, and hence views it as being uniformly distributed over [0, 1].)

If you use an English auction, each bidder will stay in the bidding until the price reaches his value; the auction will end once the price hits the lower bidder's value. Because each bidder's value is uniformly distributed over [0, 1], we can show (as in the appendix to this chapter) that the lower bidder's value, on average, equals ⅓.[22] We conclude that, if you choose to use an English auction, your expected revenue will be ⅓.

If you decide to use a first-price auction (FPA), the bidders have an incentive to shade their bids below their true values, but the winner will pay his full bid. In the appendix to this chapter, we provide a complete mathematical derivation of the equilibrium bidding strategies for this case, showing that each of the two bidders submits a bid equal to exactly half of his value; that is, each bidder's bidding strategy takes the form $b_i(V_i) = V_i/2$. (A bidder with value 1 bids ½, while a bidder with value ½ bids ¼, etc.) Because the higher bidder's value is, on average, ⅔, and because each bidder submits a bid equal to half of his value, the winning bid will, on average, equal ⅓. At least in our example, then, it doesn't matter which of these auction formats you use to sell your item—you'll get the same expected revenue!

This result may seem like a coincidence, but in fact, the "revenue equivalence" that we have discovered here is the consequence of a deeper result, the so-called revenue equivalence theorem, one of the most famous results in auction theory.[23] The revenue equivalence theorem (abbreviated hereafter as RET) tells us that two different auction formats with the same set of bidders and the same reserve prices will generate the same average revenue for the seller under three conditions: (1) that bidders have uncorrelated private values drawn from the same distribution, (2) that bidders care only for expected monetary payoffs, without any aversion to loss or risk, and (3) that in each auction, bidders play a symmetric Nash equilibrium[24] in which any bidder with the lowest possible value gets zero payoff.

RET is an interesting theoretical result, but it also serves the highly practical purpose of focusing our attention on the aspects of the auction environment that matter for auction design. *By focusing on which RET assumptions fail* in a given auction environment, we can gain insight into how to design an auction for that environment. What, then, are the key RET assumptions? There are a number of them. We identify six of them below, then devote the remainder of this section to deeper analysis of the effects of these assumptions on an auction's expected revenue. (Our discussion doesn't mention every RET assumption, just the ones that are most important for you to understand. Those who want to learn more can consult the additional references provided in Section 5.)

First, RET assumes that the *same bidders* participate in the two auctions that are being compared. An auction that is more attractive to bidders, so that more participate, will (all else being equal) obviously generate greater expected revenue. Second, RET assumes that the *reserve price* is the same in the two auctions. As we discuss in Section 4.A below, changing the reserve price can affect revenue. Third, RET assumes that the bidders are *risk neutral*. If bidders are risk averse, RET does not hold, and the auction format can affect how much revenue is generated, as we discuss in Section 4.B. Fourth, RET assumes that bidders have *private values*. If bidders have common values instead, the auction format you choose can make a difference to your revenue; we discuss the case of common values in Section 4.C. Fifth, RET assumes that bidders' values are *uncorrelated*. If bidders' values are correlated, the auction format, again, can affect the revenue you earn; we address this possibility in Section 4.C. Sixth and finally, RET assumes that bidders play a *Nash equilibrium*. We consider two reasons why bidders might fail to play a Nash equilibrium in practice: incorrect beliefs about the game (due to inexperience or overconfidence, for example) in Section 4.D and collusion in Section 4.E. In both cases, the auction format can affect your expected revenue.

A. Reserve Prices

In 1997, Overture (later acquired by Yahoo!) became the first search engine to use an auction to sell the advertisements that appear next to its search results. Such "sponsored-search auctions" eventually became big money makers,[25] but in those early years, the auction format that Overture (and later Yahoo!) used had not yet been optimized to maximize revenue. Enter Michael Schwarz, an auction-theory expert who joined Yahoo!'s newly-formed economics research unit in 2006. Schwarz had previously been an economics professor at Harvard University, but now he was tasked with improving the auction design that Yahoo! relied on for its revenue.

Schwarz and his collaborator Michael Ostrovsky (a Stanford economics professor) identified reserve prices as an aspect of the auction design that could be improved, showing that Yahoo! could substantially raise its revenue by increasing its reserve prices on many keywords. The idea was that, by refusing to place ads at bargain-basement rates, Yahoo! could induce bidders to make higher bids in the first place. The results were felt quickly, as Yahoo! experienced an impressive boost in revenues in the quarters immediately following its adoption of the new reserve pricing system.

How would you figure out the optimal reserve price to set for a particular object? In our earlier numerical example, we assumed that there were just two bidders having values drawn uniformly from the interval [0, 1]. Here we suppose that there may be *any number*, $n \geqslant 1$, of bidders,[26] but we continue to assume that the bidders have uncorrelated values drawn uniformly from the interval [0, 1]. We also suppose that you can set any reserve price, $r \geqslant 0$. According to the revenue equivalence theorem, your expected revenue if you use a first-price auction (FPA) is equal to your expected revenue in a second-price auction (SPA) with the same number of bidders and the same reserve price. So let's focus on the SPA, which is easier to analyze because truthful bidding is a weakly dominant strategy.

As we show in the appendix to this chapter, the expected revenue (denoted by *REV*) that you get from choosing the SPA with n bidders and reserve price r is given by the following formula:

$$REV(n,r)=\frac{n-1}{n+1}+r^n\left(1-\frac{2nr}{n+1}\right).$$

Let $r*(n)$ denote the optimal reserve price when there are n bidders; $r*(n)$ is the level of r that maximizes $REV(n, r)$. Remarkably, $r*(n) = ½$, no matter how many bidders there are. So, your expected revenue when setting an optimal reserve price [call it $REV*(n)$] takes the form

$$REV^*(n)=REV\left(n,\frac{1}{2}\right)=\frac{n-1}{n+1}+\frac{(1/2)^n}{n+1}.$$

Note that if you didn't use a reserve price, your expected revenue would be $$REV(n,0)=\frac{n-1}{n+1};$$ so the extra expected revenue you get from using an optimal reserve price (rather than no reserve price at all) is $$\frac{(1/2)^n}{n+1}.$$ When n = 2, this extra revenue, (, is an increase of 25% relative to what you would have earned without any reserve price [$REV(2, 0)$ = ⅓]. On the other hand, if there are 10 bidders, the extra revenue from using an optimal reserve price is only $^{1}/_{11,264}$, an increase of 0.027% relative to not having any reserve price.

In summary, the key observation regarding reserve prices for anyone designing an auction is that *setting a reserve price can raise your expected revenue by a substantial amount, but only when the number of bidders is small*.

B. Bidder Attitudes Toward Risk

If you choose to sell your object using an English auction, each bidder has an incentive to remain in the bidding until the price reaches his value, no matter what his attitude toward risk is. So your expected revenue in the English auction is unaffected by risk aversion. On the other hand, bidders in the Dutch auction have an incentive to jump in at higher prices when they are risk averse. To understand why, consider again the logic of bidding in the Dutch auction (see Section 3.B). To earn a profit, bidders need to wait until the price has dropped below their value and then continue to wait a little longer, taking the risk that someone else will jump in and grab the object. Risk-averse bidders will be willing to pay a higher price in order to avoid that risk, causing revenues in the Dutch auction to be higher when bidders are risk averse than when they are risk neutral. So, if you use a Dutch auction (or the strategically equivalent first-price auction) when your bidders are risk averse, you will have greater expected revenue than you would have from an English auction.

In summary, the key observation regarding bidder attitudes toward risk for anyone designing an auction is that *when bidders are risk averse, the first-price/Dutch auction generates greater expected revenue than the English auction*.

C. Common Values and Correlated Estimates

Suppose that the object you plan to auction has a common value and that each bidder has an estimate, $\hat{V}_i$, of that common value. For instance, before a drilling-rights auction, oil companies interested in bidding on a given tract will each run their own seismic tests and make an estimate of how much oil is under the ground. These estimates are naturally correlated with one another because if there is actually a large amount of oil under the ground, each oil company is likely to make a high estimate, and vice versa.

Each oil company in this example would love to know what the others know, because this would allow it to improve its own estimate and to bid more successfully in the auction. But consider for a moment what would happen if all of them somehow managed to learn everything that any of them knew about the oil tract up for auction. Having the same information, all of them would then have the same estimate, $\hat{V}$, of the common value. With all bidders having the same valuation, $\hat{V}$, the auction would then be one with *objective values*—like an auction of a $20 bill. In such an auction, the auction designer could extract the bidders' full willingness to pay, $\hat{V}$, and leave them with zero profit.

Notice that this kind of "private-information leakage" is good for you as the seller here; it increases the intensity of bidder competition, leading to higher expected revenue. As we saw in Chapter 9, players in a game who possess private information must be given an incentive to act on (and hence reveal) that information, and such incentives must always leave money on the table for the player in question. In the auction context, the "money on the table" corresponds to bidder surplus (or profit), the difference between the winning price and the bidder's true value. Anything you as the seller can do to reduce bidders' private information will therefore tend to have the effect of reducing their surplus—and hence increasing your revenue.

In a first-price auction, nothing is revealed about bidders' private information during the auction because bidders submit sealed bids. On the other hand, in an English auction, some private information is revealed every time a bidder drops out of the auction, Others can make inferences about what that bidder's value estimate must have been for him to drop out at that particular price.

The overall effect of this private-information revelation on expected revenue in an English auction is not intuitively obvious, because seeing others' bids may cause bidders to bid more or less aggressively. If surprisingly many bidders drop out early at low prices, others will adjust their estimates downward and will also drop out earlier. On the other hand, if surprisingly few bidders drop out at low prices, others will adjust their estimates upward and stay in the bidding even longer. Auction theorists have determined that private-information revelation generally increases expected revenue, so your expected revenue with an English auction will be larger than with a first-price auction—the intuition again being that the English auction induces more of bidders' private information to be revealed.

In summary, the key observation regarding common values and correlated estimates for anyone designing an auction is that *when bidders have common values and positively correlated estimates of the object's value, the English auction tends to generate greater expected revenue than the first-price/Dutch auction.*[27]

D. Incorrect Beliefs about the Game

DealDash.com sells flat-screen TVs and other big-ticket items using a penny auction (also known as a *bid-fee auction*). In a DealDash auction, the price starts at zero and the auction ends as soon as 10 seconds go by without any bidding. During each 10-second window, bidders can jump in to restart the clock by paying a "bidding fee" of 60 cents. The name *penny auction* comes from the fact that when someone jumps in to bid, the price increases by a penny. In a video on DealDash.com that explains how its auction works,[28] a 52″ HDTV Sony flat-screen TV that retails for about $600 ultimately sells at a price of $35.63. This may seem like a great deal, until you realize that—because DealDash collected 60 cents for every penny of the sale price—bidders overall paid 61 × $35.63 = $2,173.43, well over three times the TV's actual value. And this isn't an anomaly. Studies of penny auctions have found that penny-auction sites enjoy remarkable profit margins of 50% or more.

It's possible that bidders enjoy a thrill from participating, and occasionally winning, in penny auctions, so much so that they don't mind losing money on average—much like gamblers in a casino or people playing the lottery. However, it's equally plausible that many people who participate in these auctions have *incorrect beliefs* about the game and hence are prone to make choices that are actually not in their best interest. For instance, some bidders may overestimate their own strategic prowess, thinking that they can outbid others even though, in reality, their skills and ability to win are just average. Many studies have documented this tendency toward overconfidence, known as *illusory superiority*.[29] For instance, 88% of University of Oregon students thought that they were better-than-average drivers (relative to other students who participated in the same study), 87% of Stanford MBA students thought that their academic performance was above average, and 90% of professors at the University of Nebraska–Lincoln believed that they were better-than-average teachers.

People who *think* that they can outwit others to win money consistently in penny auctions will be the ones who wind up participating in these auctions. But, of course, only some people can be better than average, and most will end up consistently losing money. The issue is not that these people are dumb or incompetent, but that they have incorrect beliefs about the game and hence are unable to play their best response—which, with penny auctions, would be not to play at all!

The key observation to take away from this discussion is that *in reality, some bidders fail to play a best response and, hence, cannot be playing a Nash equilibrium. Auctions designed to take advantage of such mistakes can generate greater expected revenue*.

E. Collusion

Each summer, school districts across the country solicit bids on annual supply contracts for milk. Dairies submit bids in these procurement auctions, and the low bidder is selected to supply the milk that will be served to students during the following school year. Auctions could in principle be a great way for school districts to secure a good price for milk, but there's a problem: *pervasive collusion* among the dairies bidding in these auctions.

In the mid-1980s, Florida Attorney General Robert A. Butterworth began uncovering milk-auction collusion, using bid-analysis techniques developed by auction theorists who had been studying collusion in highway-construction procurement auctions. In 1993, the *New York Times* reported that "federal and state investigators have found evidence in at least 20 states that executives at the nation's largest national and regional dairy companies have conspired—sometimes for decades—to rig bids on milk products sold to schools and military bases."[30]

As the auctioneer Mike Brandly explains on *Auctioneer's Blog*,[31] "Bidders collude by agreeing with each other to not bid against each other." One way that dairies did that was to divide the school districts in a given state into *territories*, and then promise not to bid on contracts in another dairy's territory. Of course, each dairy would have had the temptation to break that promise and try to steal some extra business. However, because the winner of any public contract is a matter of public record, any such betrayal would have been easily detectable, ending the collusive arrangement and hence returning the dairies to a less profitable competitive mode.

The longevity of the dairy-auction cartels relied in large part on the fact that the dairies were playing a repeated game (like those we analyzed in Chapter 10), with the ability to observe others' actions and to retaliate against anyone who violated the collusive agreement. But the same principle also applies in auctions that are not repeated. If an auction format allows bidders to observe others' actions and to retaliate, then colluding will be easier than if the auction were designed to obscure bidders' actions until it is too late to retaliate.

For instance, suppose that two bidders in a first-price auction have made a "handshake agreement" to collude. Bidder 1 has value \$100, bidder 2 has value \$80, and both are fairly certain that no one else is willing to pay more than \$50. They agree that bidder 2 will bid \$50 (thereby allowing bidder 1 to win the item at a slightly higher price) and that after the auction, bidder 1 will pay \$10 to bidder 2. If bidder 2 goes along with this plan, she will walk away with \$10. However, if she cheats and bids \$60, she can win the item for a profit of \$20 = \$80 − \$60. She can get away with such cheating in a first-price auction because there is no way for bidder 1 to see what bidder 2 has bid until it's too late. On the other hand, in an English auction, bidder 2 has no hope of winning by cheating. If she stays in the bidding past \$50, bidder 1 will just stay in the bidding himself. Bidder 2 therefore has no temptation to cheat, making collusion easier to sustain in an English auction.

The key observation to take away from this analysis of bidder collusion is that *the English auction is more prone to collusion than the Dutch/first-price auction*.

Endnotes

- We provide an appendix to this chapter for those interested in more mathematical details. For now, please just take our word for it that with two bidders, the lower bidder's value, on average, equals ⅓ and the higher bidder's value, on average, equals ⅔. Return to reference 22
- See Vijay Krishna, *Auction Theory*, 2nd ed. (Burlington, Mass.: Academic Press, 2010), Chapter 3 for details. Return to reference 23
- A *strategy* in an auction specifies the bid $b_i(V_i)$ that bidder i will make for any given value V_i. Bidding strategies are "symmetric" if $b_i(V_i) = b(V_i)$ for all i or, in words, if all bidders with the same value submit the same bid. Return to reference 24
- In 2012, Google hosted over 5.1 billion searches per day; the corresponding sponsored-search auctions generated $43.7 billion for Google that year, 95% of its total revenue. For more on the auction formats used by different search engines, see Hal Varian, "Online Ad Auctions," *American Economic Review*, vol. 99, no. 2 (May 2009), pp. 430–34. Return to reference 25
- The case with one bidder ($n = 1$) corresponds to having a single buyer, also known as "monopoly pricing." So, when you learned about monopoly pricing in your introductory microeconomics class, you were actually learning about auction theory! Return to reference 26
- In a landmark 1982 paper, Paul Milgrom and Robert Weber established the revenue ranking theorem: so long as bidders' value estimates are positively correlated (in a particular technical sense known as "affiliation"), expected revenue in the English auction is greater than in the second-price auction, which is itself greater than in the first-price/Dutch auction. See Paul Milgrom and

Robert Weber, “A Theory of Auctions and Competitive Bidding,” *Econometrica*, vol. 50, no. 5 (September 1982), pp. 1089–1122. Return to reference 27

- See https://www.dealdash.com/help/how-it-works(accessed July 23, 2018). Return to reference 28
- Illusory superiority is also known as “superiority bias” and “the Lake Wobegon effect” (after a fictional Minnesota town where “all the children are above average”). See Ola Svenson, “Are We All Less Risky and More Skillful Than Our Fellow Drivers?” *Acta Psychologica*, vol. 47, no. 2 (February 1981), pp. 143–48; “It’s Academic,” *Stanford GSB Reporter*, April 24, 2000, pp. 14–15; and K. Patricia Cross, “Not Can, But *Will* College Teaching Be Improved?” *New Directions for Higher Education*, vol. 1977, no. 17 (Spring), pp. 1–15. Return to reference 29
- Diana B. Henriques and Dean Baquet, “Evidence Mounts of Rigged Bidding in Milk Industry,” *New York Times*, May 23, 1993. Return to reference 30
- Mike Brandly, “What is Collusion at an Auction?” May 14, 2000, available at https://mikebrandlyauctioneer.wordpress.com/2010/05/14/what-is-collusion-at-an-auction/(accessed May 27, 2019). Return to reference 31

Glossary

revenue equivalence theorem (RET)
: A famous result in auction theory specifying conditions under which two auctions will generate the same expected revenue for the seller.

penny auction
: An auction format in which each bidder may pay a bidding fee (say, 60 cents) to advance the price by one cent. The auction continues until no one is willing to advance the price any longer, at which point the last bidder wins and pays the final price.

5 FURTHER READING

This chapter only scratches the surface of auction theory, auction design, and the broader field known as *market design*. In this section, we provide reading recommendations for those who would like to dig in deeper and learn more about the subject.

Books with compelling stories about auctions and other markets:

John McMillan, *Reinventing the Bazaar: A Natural History of Markets* (New York: W. W. Norton, 2003).

Alvin Roth, *Who Gets What—and Why: The New Economics of Matchmaking and Market Design* (Wilmington, Mass.: Mariner Books, 2016).

Paul Milgrom, *Putting Auction Theory to Work* (New York: Cambridge University Press, 2004).

Scholarly books and articles on auction-theory topics:

Multi-unit auctions (Section 1): Ali Hortaçsu and David McAdams, “Empirical Work on Auctions of Multiple Objects,” *Journal of Economic Literature*, vol. 56, no. 1 (March 2018), pp. 157–84.

Combinatorial auctions (Section 1): Peter Cramton, Yoav Shoham, and Richard Steinberg (eds.), *Combinatorial Auctions* (Cambridge, Mass: MIT Press, 2005).

Internet-enabled auctions (Section 1): David Lucking-Reiley, “Auctions on the Internet: What’s Being Auctioned, and How?” *Journal of Industrial Economics*, vol. 48, no. 3 (September 2000), pp. 227–52; and on the issue of participation and winner’s curse, Patrick Bajari and Ali Hortaçsu, “The Winner’s Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions,” *RAND Journal of Economics*, vol. 34, no. 2 (Summer 2003), pp. 329–55.

Experimental evidence on bidder behavior, including winner’s curse (Section 2): John H. Kagel and Dan Levin, “Auctions: A Survey of Experimental Research,” in *The Handbook of Experimental Economics*, vol. 2, ed. John H. Kagel and Alvin E. Roth (Princeton: Princeton University Press, 2016), pp. 563–637.

Jump bidding (Section 3.A): Christopher Avery, “Strategic Jump Bidding in English Auctions,” *Review of Economic Studies*, vol. 65, no. 2 (April 1998), pp. 185–210; Robert F. Easley and Rafael Tenorio, “Bidding Strategies in Internet Yankee Auctions,” SSRN Working Paper 170028 (June 1999).

Shill bidding (Section 3.A): David McAdams and Michael Schwarz, “Who Pays when Auction Rules are Bent?” *International Journal of Industrial Organization*, vol. 25, no. 5 (October 2007), pp. 1144–57; Georgia Kosmopoulou and Dakshina G. De Silva, “The Effect of Shill Bidding upon Prices: Experimental Evidence,” *International Journal of Industrial Organization*, vol. 25, no. 2 (April 2007), pp. 291–313.

Sniping (Section 3.A): Axel Ockenfels and Alvin Roth, “Late and Multiple Bidding in Second-Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction,” *Games and Economic Behavior*, vol. 55, no. 2 (May 2006), pp. 297–320.

Vickrey auctions (Section 3.D): Vickrey’s original article is “Counterspeculation, Auctions, and Competitive Sealed Tenders,” *Journal of Finance*, vol. 16, no. 1 (March 1961), pp. 8–37. See also David Lucking-Reiley, “Vickrey Auctions in Practice: From Nineteenth-Century Philately to Twenty-First-Century e-Commerce," *Journal of Economic Perspectives*, vol. 14, no. 3 (Summer 2000), pp. 183–92.

All-pay auctions (Section 3.E): Many real-world contests can be interpreted as all-pay auctions.

- **In lobbying and political corruption:** Michael R. Baye, Dan Kovenock, and Casper G. de Vries, "Rigging the Lobbying Process: An Application of the All-Pay Auction," *American Economic Review*, vol. 83, no. 1 (March 1993), pp. 289–94.
- **In crowdsourcing over the Internet:** Dominic DiPalantino and Milan Vojnovic, "Crowdsourcing and All-Pay Auctions," *Proceedings of the 10th ACM Conference on Electronic Commerce* (July 2009), pp. 119–28.
- **In patent races:** Michael R. Baye and Heidrun C. Hoppe, "The Strategic Equivalence of Rent-Seeking, Innovation, and Patent-Race Games," *Games and Economic Behavior*, vol. 44, no. 2 (August 2003), pp. 217–26.

War of attrition (Section 3.F): Many real-world contests can be interpreted as wars of attrition.

- **Animal contests over territory and mates:** Steven N. Austad, "A Game Theoretical Interpretation of Male Combat in the Bowl and Doily Spider (*Frontinella pyramitela*)," *Animal Behaviour*, vol. 31, no. 1 (February 1983), pp. 59–73; James H. Marden and Jonathan K. Waaget, "Escalated Damselfly Territorial Contests Are Energetic Wars of Attrition," *Animal Behaviour*, vol. 39, no. 5 (May 1990), pp. 954–59.
- **Business exit from a declining industry:** Yuya Takahashi, "Estimating a War of Attrition: The Case of the US Movie Theater Industry," *American Economic Review*, vol. 105, no. 7 (July 2015), pp. 2204–41.
- **Political delay before stabilizing out-of-control deficits:** Alberto Alesina and Allan Drazen, "Why Are

Stabilizations Delayed?" *American Economic Review*, Vol. 81, no. 5 (December 1991), pp. 1170–88.

Auctions with risk-averse bidders (Section 4.B): Eric Maskin and John Riley, "Optimal Auctions with Risk Averse Buyers," *Econometrica*, vol. 52, no. 6 (November 1984), pp. 1473–1518; Steven Matthews, "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View," *Econometrica*, vol. 55, no. 3 (May 1987), pp. 633–46.

Auctions with collusion (Section 4.E): Robert C. Marshall and Leslie Marx, *The Economics of Collusion: Cartels and Bidding Rings* (Cambridge, Mass.: MIT Press, 2012).

SUMMARY

Auctions arise whenever players compete over a scarce resource and the game they play involves bidding and direct competition among *bidders*. Auctions take many forms, from art lovers raising their paddles in an art-house auction and firms competing to hire an employee to lobbyists raising campaign cash in an all-pay contest over legislative policy. Auction theory identifies the key features of an auction environment that determine how much you should bid (if you are a bidder) and how you design the auction (if you are the seller).

The most important feature of an auction environment is whether bidders have private values or common values for the object to be auctioned. In an auction with *common values*, how much the object is worth to you depends on what others know about it, and you are more likely to win the object when they do not want it. Bidders who do not account for this effect are likely to suffer the *winner’s curse*, paying more for the object than it is worth. Fortunately, the winner’s curse can be avoided once you understand it.

In an auction with *private values*, how you should bid depends on the details of the auction format. In an *English auction* or a *second-price auction*, *truthful bidding* is your best strategy. In a *first-price auction* or a *Dutch auction*, you should always bid less than your true value. And in an *all-pay auction*, your best strategy may involve submitting a low bid even when your value is high.

Sellers can often increase their expected revenue by committing to keep the object unless a *reserve price* is met. Many other factors also affect how you should bid in or design an auction, including risk aversion, correlation of bidders’ private information, bidders’ incorrect beliefs about the game, and collusion among bidders.

KEY TERMS

all-pay auction (601)

ascending-price auction (595)

auction (587)

auction designer (587)

bidder (587)

combinatorial auction (587)

common value (591)

descending-price auction (597)

drop-out price (595)

Dutch auction (597)

English auction (595)

first-price auction (598)

jump bidding (596)

multi-unit auction (587)

objective value (590)

open-outcry auction (589)

penny auction (609)

private information (590)

Glossary

auction
: A game in which multiple players (called bidders) compete for a scarce resource.

bidder
: A player in an auction game.

valuation
: The benefit that a bidder gets from winning the object in an auction.

auction designer
: A player who sets the rules of an auction game.

procurement auction
: An auction in which multiple bidders compete to supply an item. Bids in a procurement auction are prices that bidders are willing to receive to supply the good. The lowest bidder wins and is paid her bid.

multi-unit auction
: An auction in which multiple identical objects are sold.

combinatorial auction
: An auction of multiple dissimilar objects in which bidders are able to bid on and win combinations of objects.

sealed-bid auction
: An auction mechanism in which bids are submitted privately in advance of a specified deadline, sometimes in sealed envelopes.

open-outcry auction
: An auction mechanism in which bids are made openly for all to hear or see.

private information
: Information known by only one player.

objective value
: An auction is called an objective-value auction when the object up for sale has the same value to all bidders and

each bidder knows that value.

private value

An auction is called a private-value auction when each bidder has private information about their own valuation of the object up for sale, but knowing others’ private information would not change any bidder’s own willingness to pay for the object. An important special case is when each bidder knows their own valuation but others do not.

common value

An auction is called a common-value auction when the object up for sale has the same value to all bidders, but each bidder knows only an imprecise estimate of that value.

winner’s curse

A situation in a common-value auction where the winner fails to take account of the fact that when she wins, she is likely to have made an overly optimistic estimate of the object’s value. Bidders who correctly anticipate this possibility can avoid the winner’s curse by lowering their bids appropriately.

ascending-price auction

An open-outcry auction in which prices are announced in increasing order either by an auctioneer (in the case of an English auction) or by bidders themselves (in the case of jump bidding). The last person to bid or accept the announced price wins the auction and pays that price.

English auction

A type of *ascending-price auction* in which the auctioneer calls out a sequence of increasing prices, bidders decide when to drop out of the bidding, and the last bidder remaining pays the last announced price.

reserve price

The minimum price set by the seller of an item up for auction; if no bids exceed the reserve, the item is not sold.

drop-out price

In an English auction, the price at which a bidder drops out of the bidding.

jump bidding

Submitting a bid that is significantly higher than the previous bid and well beyond whatever minimum bid increment exists.

shill bidder

A fake bidder created by sellers at an auction to place fictitious bids for an object they are selling.

sniping

Waiting until the last moment to make a bid.

descending-price auction

An open-outcry auction in which the auctioneer announces possible prices in descending order. The first person to accept the announced price wins the auction and pays that price. Also called **Dutch auction**.

Dutch auction

Same as a **descending-price auction**.

first-price auction

A sealed-bid auction in which the highest bidder wins and pays the amount of her bid.

shading

A strategy in which bidders bid slightly below their true valuation of an object.

second-price auction

A sealed-bid auction in which the highest bidder wins the auction but pays a price equal to the value of the second-highest bid; a special case of the *Vickrey auction*.

Vickrey auction

An auction design proposed by William Vickrey in which truthful bidding is a weakly dominant strategy for each bidder. When a single object is sold, the Vickrey auction is the same as the *second-price auction*.

truthful bidding

A practice by which bidders in an auction bid their true valuation of an object.

all-pay auction
: An auction in which each person who submits a bid must pay her highest bid amount at the end of the auction, even if she does not win the auction.

war of attrition
: A contest between multiple players in which each player decides when to retreat, the victor is whoever remains the longest, and choosing to remain longer is costly for each player.

revenue equivalence theorem (RET)
: A famous result in auction theory specifying conditions under which two auctions will generate the same expected revenue for the seller.

penny auction
: An auction format in which each bidder may pay a bidding fee (say, 60 cents) to advance the price by one cent. The auction continues until no one is willing to advance the price any longer, at which point the last bidder wins and pays the final price.

single-object auction
: An auction in which a single indivisible object is sold.

SOLVED EXERCISES

1. In each of the following examples, does the object being auctioned have an objective value or not? Why or why not? In each case, imagine that the object in question is being auctioned by your game-theory professor in your game-theory class:
   1. \$20 Amazon gift card
   2. Lunch with the professor
   3. A bottle of water
2. A house painter has a regular contract to work for a builder. On these jobs, her cost estimates are generally right: sometimes a little high, sometimes a little low, but correct on average. When her regular work is slack, she bids competitively for other jobs. "Those are different," she says. "They almost always end up costing more than I estimate." If we assume that her estimating skills do not differ between the two types of jobs, what can explain the difference?
3. Consider an auction where n identical objects are offered, and there are $(n + 1)$ bidders. The actual value of an object is the same for all bidders and equal for all objects, but each bidder has only an independent estimate, subject to error, of this common value. The bidders submit sealed bids. The top n bidders get one object each, and each pays what she has bid. What considerations will affect your bidding strategy? How?
4. You are in the market for a used car and see an ad for the model that you like. The owner has not set a price, but invites potential buyers to make offers. Your prepurchase inspection gives you only a very rough idea of the value of the car; you think it is equally likely to be anywhere in the range of \$1,000 to \$5,000 (so your calculation of the average value is \$3,000). The current owner knows the exact value and will accept your offer if it exceeds that value. If your offer is accepted and you get the car, then you will find out the truth. But you have some special repair skills and know that when you own the car, you will be able to work on it and increase its value by a third (33.3 . . . %) of whatever it is worth.
   1. What is your expected profit if you offer \$3,000? Should you make such an offer?
   2. What is the highest offer that you can make without losing money on the deal?

5. It' s your birthday, and there' s one cupcake left. Your two best friends Bob (B) and Apple (A), each want it. Each friend i = A, B is willing to pay V_i for the entire cupcake or $V_i/3$ for half of the cupcake (they don' t like sharing), where V_i are uncorrelated private values drawn uniformly from \$0 to \$12.
   1. Suppose that you hold a second-price auction with no reserve price. Compute Bob' s expected surplus when his realized value is V_B = \$12, V_B = \$9, V_B = \$6, and V_B = \$3. Hint: Expected surplus = (probability that Bob wins) × (Bob' s surplus when winning). Bob wins whenever he has the highest value, which happens with probability $V_B/2$, and pays Apple' s valuation, which, when Bob wins, is uniformly distributed from \$0 to V_B.
   2. If you instead give Bob half of the cupcake, Bob' s surplus is $v_B/3$. Compare this *gift surplus* with the *auction surplus* that you computed in part (a). For each of the values V_B = \$12, V_B = \$9, V_B = \$6, and V_B = \$3, does Bob prefer to get half of the cupcake for free or to compete in an auction for the whole cupcake?
6. In this problem, we consider a special case of the first-price, sealed-bid auction and show what the equilibrium amount of bid shading should be. Consider a first-price, sealed-bid auction with n risk-neutral bidders. Each bidder has a private value independently drawn from a uniform distribution over the interval [0, 1]. That is, for each bidder, all values between 0 and 1 are equally likely. The complete strategy of each bidder is a "bid function" that will tell us, for any value v, what amount $b(V)$ that bidder will choose to bid. Deriving the equilibrium bid functions requires solving a differential equation, but instead of asking you to derive the equilibrium using a differential equation, this problem proposes a candidate equilibrium and asks you to confirm that it is indeed a Nash equilibrium.

   It is proposed that the equilibrium bid function for n = 2 is $b(V) = V/2$ for each of the two bidders. That is, if we have two bidders, each should bid half her value, which represents considerable shading.

   1. Suppose you' re bidding against just one opponent whose value is uniformly distributed over the interval [0, 1], and who always bids half her value. What is the probability that you

will win if you bid $b = 0.1$? If you bid $b = 0.4$? If you bid $b = 0.6$?
2. Put together the answers to part (a). What is the correct mathematical expression for *Prob*(win), the probability that you win, as a function of your bid b?
3. Find an expression for the expected profit you make when your value is v and your bid is b, given that your opponent is bidding half her value. Remember that there are two cases: either you win the auction, or you lose the auction. You need to average the profit between these two cases.
4. What is the value of b that maximizes your expected profit? The answer should be a function of your value V.
5. Use your results to argue that it is a Nash equilibrium for both bidders to follow the same bid function $b(V) = V/2$.

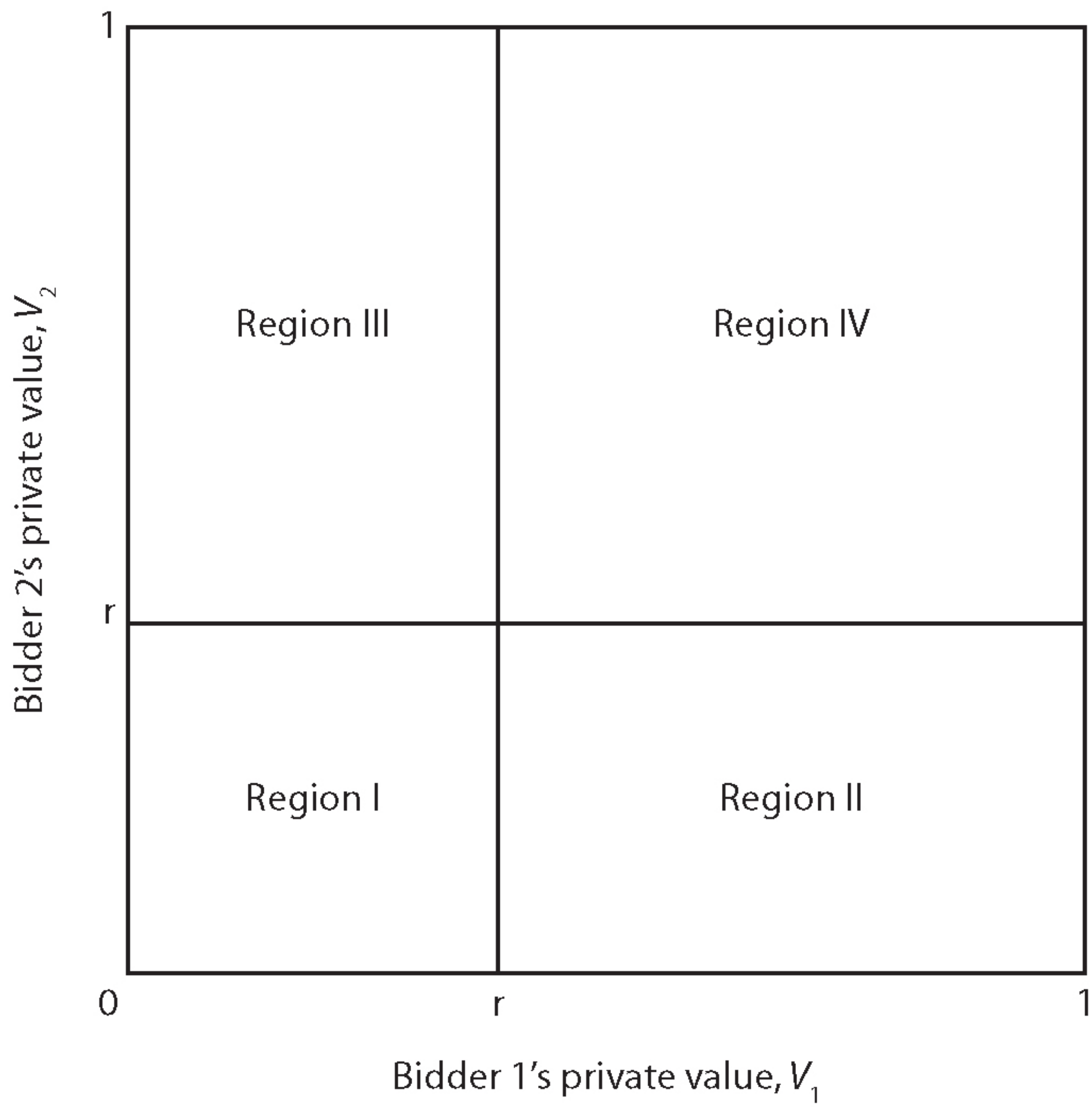
7. **(Optional)** This question looks at equilibrium bidding strategies in all-pay auctions in which bidders have private values for the good, as opposed to the scenario in Section 4, where the all-pay auction is for a good with a publicly known value. Consider an all-pay auction where each player's private value is distributed uniformly between 0 and 1, and those of different players are independent of each other.
   1. Verify that the Nash equilibrium bid function is $b(V) = [(n - 1)/n]V^n$. Use an approach similar to that of Exercise S6. Remember that in an all-pay auction, you pay your bid even when you lose, so your payoff is $V - b$ when you win and $-b$ when you lose.
   2. Now consider the effect of changing the number of bidders n. For clarity, call the bid function $b_n(V)$. Show that $b_3(V) > b_2(V)$ if and only if $V > 3/4$, and that $b_4(V) > b_3(V)$ if and only if $V > 8/9$. Using the results for those cases, what can you say about the general case comparing $b_{n+1}(V) > b_n(V)$?

UNSOLVED EXERCISES

1. In each of the following examples, do bidders have private values or not? Why or why not? In each case, imagine that the object in question is being auctioned to a group of Wall Street investors.
    1. Shares of stock in Mystery Enterprises Agricultural Technologies (MEAT), a (fictional) biotech startup that sells “Mystery Meat” under the tagline “There’s no mystery—it tastes like beef!!”
    2. A lifetime supply of Mystery Meat.
    3. A hamburger (made from traditional beef).
2. “In the presence of very risk-averse bidders, a person selling her house in an auction will have a high expected profit by using a first-price, sealed-bid auction.” True or false? Explain your answer.
3. Suppose that three risk-neutral bidders are interested in purchasing a Princess Beanie Baby. The bidders (numbered 1 through 3) have valuations of $12, $14, and $16, respectively. The bidders will compete in auctions as described in parts (a) through (d); in each case, bids can be made in $1 increments at any value from $5 to $25.
    1. Which bidder wins an open-outcry English auction? What is the final price paid, and what is the profit to the winning bidder?
    2. Which bidder wins a second-price, sealed-bid auction? What is the final price paid, and what is the profit to the winning bidder? Contrast your answer here with that for part (a). What is the cause of the difference in profits in these two cases?
    3. In a sealed-bid, first-price auction, all the bidders will bid a positive amount (at least $1) less than their true valuations. What is the likely outcome in this auction? Contrast your answer with those for parts (a) and (b). Does the seller of the Princess Beanie Baby have any clear reason to choose one of these auction mechanisms over the other?
    4. Risk-averse bidders would reduce the shading of their bids in part (c); assume, for the purposes of this question, that they do not shade at all. If that were the case, what would be the winning price (and profit for the bidder) in part (c)?

Would the seller care about which type of auction she chooses? Why?

4. In this exercise, you will explore how much setting a reserve price can increase the seller's expected revenue. Two bidders $i = 1, 2$ have (uncorrelated) private values that are uniformly distributed over the interval from 0 to 1. The figure below illustrates the four basic possible outcomes. First, if $V_1 < r$ and $V_2 < r$ (Region I), then no sale occurs; this happens with probability $Prob(V_1 < r) \times Prob(V_2 < r) = r \times r = r^2$. Second, if $V_1 > r$ and $V_2 < r$ (Region II), then bidder 1 wins and pays the reserve price r; this happens with probability $Prob(V_1 > r) \times Prob(V_2 < r) = (1 - r) \times r = r - r^2$. Third, if $V_1 < r$ and $V_2 > r$ (Region III), then bidder 2 wins and pays r; this also happens with probability $r - r^2$. Finally, if $V_1 > r$ and $V_2 > r$ (Region IV), then the bidder with the higher value wins and pays the lower value; this happens with probability $(1 - r)^2$, and when it happens, the price on average equals $\frac{2r+1}{3}$.[32]



1. What is the seller's expected revenue without any reserve price?
2. Provide a formula expressing the seller's expected revenue as a function of r. Hint: Determine the probability that bidder values fall in each of the four possible regions shown in the figure and the expected price (with price = 0 if no sale is made) when values fall in each region.
3. Does the seller earn more expected revenue, less expected revenue, or the same amount when setting a reserve price r = ½ than when setting no reserve price? How much expected revenue is gained (or lost) by setting reserve price r = ½? How often is the object left unsold?

4. Does the seller earn more expected revenue, less expected revenue, or the same amount when setting a reserve price r = ¾ than when setting no reserve price? How much expected revenue is gained (or lost) by setting reserve price r = ¾? How often is the object left unsold?

5. You are a turnaround artist, specializing in identifying underperforming companies, buying them, improving their performance and stock price, and then selling them. You have found such a prospect, Sicco. This company’s marketing department is mediocre; you believe that if you take over the company, you will increase its value by 75% of whatever it was before. But its accounting department is very good; it can conceal assets, liabilities, and transactions to a point where the company’s true value is hard for outsiders to identify. (But insiders know the truth perfectly.) You think that the company’s value in the hands of its current management is somewhere between $10 million and $110 million, uniformly distributed over this range. The current management will sell the company to you if, and only if, your bid exceeds the true value known to them.
   1. If you bid $110 million for the company, your bid will surely succeed. Is your expected profit positive?
   2. If you bid $50 million for the company, what is the probability that your bid succeeds? What is your expected profit if you do succeed in buying the company? Therefore, at the time when you make your bid of $50 million, what is your expected profit? (Warning: In calculating this expectation, don’t forget the probability of your getting the company.)
   3. What should you bid if you want to maximize your expected profit? (Hint: Assume your bid is $\$X$ million. Carry out the same analysis as in part (b) above, and find an algebraic expression for your expected profit as seen from the time when you are making your bid. Then choose X to maximize this expression.)
6. The idea of the winner’s curse can be expressed slightly differently from its usage in the chapter: “The only time your bid matters is when you win, which happens when your estimate is higher than the estimates of all the other bidders. Therefore you should focus on this case. That is, you should always act as if all the others have received estimates lower than yours, and use this ‘information’ to revise your own estimate.” Here we ask you to apply this idea to a very different situation.

A jury consists of 12 people who hear and see evidence presented at a trial and collectively reach a verdict of guilt or innocence. To simplify the process somewhat, assume that the jurors hold a single simultaneous vote to determine the verdict. Each juror is asked to vote Guilty or Not Guilty. The accused is convicted if all 12 vote Guilty and is acquitted if one or more vote Not Guilty; this is known as the unanimity rule. Each juror's objective is to arrive at a verdict that is the most accurate verdict in light of the evidence, but each juror interprets the evidence in accord with her own thinking and experience. Thus, she arrives at an estimate of the guilt or the innocence of the accused that is individual and private.

1. If jurors vote truthfully—that is, in accordance with their individual private estimates of the guilt of the accused—will the verdict be Not Guilty more often under a unanimity rule or under a majority rule, where the accused is convicted if 7 jurors vote Guilty? Explain. What might we call the "juror's curse" in this situation?
2. Now consider the case in which each juror votes strategically, taking into account the potential problems of the juror's curse and using all the devices of information inference that we have studied. Are individual jurors more likely to vote Guilty under a unanimity rule when voting truthfully or when voting strategically? Explain.
3. Do you think strategic voting to account for the juror's curse would produce too many Guilty verdicts? Why or why not?

7. (**Optional**) This exercise is a continuation of Exercise S6; it looks at the general case where n is any positive integer. It is proposed that the equilibrium bid function with n bidders is $b(V) = V(n - 1)/n$. For $n = 2$, we have the case explored in Exercise S6: Each of the bidders bids half of her value. If there are nine bidders ($n = 9$), then each should bid $^9/_{10}$ of her value, and so on.
   1. Now there are $n - 1$ other bidders bidding against you, each using the bid function $b(V) = V(n - 1)/n$. For the moment, let's focus on just one of your rival bidders. What is the probability that she will submit a bid less than 0.1? Less than 0.4? Less than 0.6?
   2. Using the above results, find an expression for the probability that the other bidder has a bid less than your

bid amount b.

3. Recall that there are $n - 1$ other bidders, all using the same bid function. What is the probability that your bid b is larger than *all* of the other bids? That is, find an expression for $Prob$(win), the probability that you win, as a function of your bid b.
4. Use this result to find an expression for your expected profit when your value is V and your bid is b.
5. What is the value of b that maximizes your expected profit?
6. Use your results to argue that it is a Nash equilibrium for all n bidders to follow the same bid function $b(V) = V(n - 1)/n$.

Endnotes

- When two bidders have private values uniformly drawn from an interval $[A, B]$, the higher value, on average, equals $(A + 2B)/3$ while the lower value, on average, equals $(2A + B)/3$. When bidder values are in Region IV, they are each drawn from the interval $[r, 1]$. Return to reference 32

■ Appendix: Computing Bidding Equilibria

In this appendix, we derive equilibrium bidding strategies for the second-price auction, first-price auction, and all-pay auction in an extended version of the numerical example in the main text, allowing for any number of bidders and for any reserve price. In particular, there are any number, $n \geqslant 1$, of risk-neutral bidders, each with uncorrelated (independent) private values, V_i, drawn uniformly from the interval [0, 1] (uniformly distributed).

A. Math Facts

Our analysis in this appendix takes advantage of several basic facts about the uniform distribution and leverages mathematical notation to express complex formulas as simply as possible. The following six facts are used throughout our bidding strategy calculations below.

I. MATH FACT ONE (MF1)The first math fact addresses the average highest and average second-highest bidder value. Because bidders’ values are uniformly distributed over the interval [0, 1], each bidder’s value is, on average, equal to ½. But how high do bidders’ values tend to be *when they win* and *when they come in second*? Given bidder values V_1, V_2, . . . , V_n, let $V^{(1)} = \max\{V_1, V_2, . . . , V_n\}$ denote the highest bidder’s value, and let $V^{(2)}$ denote the second-highest bidder’s value. $V^{(1)}$ is the value of the bidder who wins; its average value is denoted by $E[V^{(1)}]$.[33] $V^{(2)}$ is the value of the bidder who comes in second; its average value is denoted by $E[V^{(2)}]$. We will make extensive use of the fact that $E[V^{(1)}]=\frac{n}{n+1}$ and $E[V^{(2)}]=\frac{n-1}{n+1}$.[34] For example, in the special case with two bidders considered in the main text, $E[V^{(1)}]=\frac{2}{3}$ and $E[V^{(2)}]=\frac{1}{3}$.

II. MATH FACT TWO (MF2)The second math fact addresses the average highest bidder value when the reserve price is not met. Suppose that $V^{(1)} < r$, so that all bidders' values are less than the reserve price. Conditional on that being true, $V^{(2)}$, on average, equals $\frac{n-1}{n+1}r.$ This can be written more succinctly, using mathematical notation, as

$$E[V^{(2)}|V^{(1)} < r] = \frac{n-1}{n+1}r.$$

III. MATH FACT THREE (MF3)The third math fact relates to the likelihood that the reserve price is not met. Each bidder, I, is unwilling to pay the reserve price so long as $V_i < r$. Since bidder i's value is uniformly distributed over [0,1], this happens with probability r, or, using mathematical notation, $Prob[V_i < r] = r$. Of course, the highest bidder's value, $V^{(1)}$, can be less than the reserve price, r, only if all n bidders' values are less than r; so, $Prob[V^{(1)} < r] = (Prob[V_i < r])^n = r^n$.

IV. MATH FACT FOUR (MF4)The fourth math fact relates to the average second-highest bidder value when only one bidder meets the reserve price. Suppose that $V^{(2)} < r < V^{(1)}$, so that exactly one bidder's value exceeds the reserve price. Conditional on that being true, the second-highest bidder's value is, on average, equal to $\frac{n-1}{n}r,$ or, in mathematical notation,

$$E[V^{(2)} | V^{(2)} < r < V^{(1)}] = \frac{n-1}{n}.$$

V. MATH FACT FIVE (MF5)The fifth math fact relates to the likelihood that only one bidder meets the reserve price. It happens that exactly one bidder's value exceeds the reserve price with probability $nr^{n-1}(1 - r)$ or, more succinctly, $Prob[V^{(2)} < r < V^{(1)}] = nr^{n-1}(1 - r)$.

VI. MATH FACT SIX (MF6)The sixth math fact determines the average second-highest bidder value, conditional on the highest bidder value. Suppose that bidder i has the highest value. The second-highest value must, on average, be equal to $\frac{n-1}{n}V_i$; that is,

$$E[V^{(2)} | V_i = V^{(1)}] = \frac{n-1}{n}V_i.$$

B. Second-Price Auction Revenue with Any Reserve

Consider a second-price auction (SPA) in which the reserve price is set at some value r, $r \geqslant 0$. Let $REV(n, r)$ denote the expected revenue in this SPA with n bidders and reserve price r. In the SPA, each bidder has a dominant strategy to stay out of the auction if $V_i < r$ and, otherwise, to enter the auction and bid truthfully. If the reserve price were 0, all bidders would participate, and expected revenue would be

$$REV(n,0)=E[V^{(2)}]=\frac{n-1}{n+1}$$ by MF1.

Depending on the highest bidder value, $V^{(1)}$, and the second-highest value, $V^{(2)}$, adding a reserve price may or may not change the auction outcome. There are three possibilities.[35]

Case 1: $V^{(1)} < r$. In this case, setting the reserve price causes the object not to sell, hurting the seller because revenue $V^{(2)}$ is lost. This case occurs with probability Prob($V^{(1)} < r$) = r^n (by MF3), and when it does, revenue decreases from an average of

$$E[V^{(2)} | V^{(1)} < r]=\frac{n-1}{n+1}r$$ to 0 (by MF2).

The overall expected harm to the seller in this case is therefore

$$r^n \frac{n-1}{n+1}r=\frac{n-1}{n+1}r^{n+1}.$$

Case 2: $V^{(2)} < r < V^{(1)}$. In this case, the object sells at price r rather than at price $V^{(2)}$, benefiting the seller. This case occurs with probability Prob($V^{(2)} < r < V^{(1)}$) = $nr^{n-1}(1-r)$ (by MF5), and when it does,

$$E[V^{(2)} \mid V^{(2)} < r < V^{(1)}] = \frac{n-1}{n} r$$

(by MF4). The overall expected benefit to the seller in this case is therefore

$$nr^{n-1}(1-r)\left(r - \frac{n-1}{n} r\right) = r^n (1-r).$$

Case 3: $V^{(2)} > r$. In this case, the reserve price has no effect since the object still sells at price $V^{(2)}$. Adding up all these effects, we conclude that expected revenue for the SPA would be

$$REV(n,r) = REV(n,0) + r^n(1-r) - \frac{n-1}{n+1} r^{n+1},$$

which, after plugging in $REV(n,0) = \frac{n-1}{n+1}$ and collecting terms, yields the formula

$$REV(n,r) = \frac{n-1}{n+1} + r^n - \frac{2n}{n+1} r^{n+1}.$$

C. Optimal Reserve Price

We can show that a seller's optimal reserve price, $r*(n)$, the reserve price that maximizes expected revenue, is equal to ½ for all n. (This derivation uses calculus.) To calculate $r*(c)$, we first take the derivative of $REV(n, r)$, derived in Section B , with respect to r:

$$r: \frac{dREV(n,r)}{dr} = nr^{n-1} - 2nr^{n}.$$

Since $r*(n)$ maximizes $REV(n, r)$ (and $REV(n, r)$ is a smooth function), it must be true that

$$\frac{dREV(n,r^*(n))}{dr} = 0.$$

We conclude that $nr*(n)^{n-1}(1 - 2r*(n)) = 0$, which is possible only if $r*(n) = 0$

$$\text{if } r^*(n) = 0 \text{ or } r^*(n) = \frac{1}{2}.$$

or

Comparing these two possibilities, note that

$$REV(n,0) = \frac{n-1}{n+1}$$

and that

$$REV\left(n, \frac{1}{2}\right) = \frac{n-1}{n+1} + \frac{(1/2)^n}{n+1} > REV(n,0).$$

So, it must be true that the optimal reserve price

$$r^*(n) = \frac{1}{2}$$

for all n.

The argument thus far applies to the second-price auction (SPA), but, according to the revenue equivalence theorem (RET), it also applies to the first-price auction (FPA). To see why, let $REV^{FPA}(n, r)$ and $REV^{SPA}(n, r)$ denote expected revenue in the FPA and SPA, respectively, given n bidders and reserve price r. Because all RET assumptions hold in this example, RET implies that $REV^{FPA}(n, r) = REV^{SPA}(n, r)$ for all (n, r). The fact that $REV^{SPA}(n, r)$ is maximized at

$$r^*(n) = \frac{1}{2}$$

for all n therefore implies that $REV^{FPA}(n, r)$ is maximized at $r^*(n) = ½$. Indeed, the same argument shows that $r^*(n) = ½$ is the optimal reserve price in any of the standard auction formats; it is true not just for the SPA and the FPA/Dutch auction, but also for the English auction, the all-pay auction, and the war of attrition.

D. Equilibrium Bidding Strategies in the First-Price Auction

There's another way that the revenue equivalence theorem (RET) can come in handy when analyzing auctions. It provides a shortcut when computing equilibrium bidding strategies in tough-to-analyze auctions. (In what follows, we focus for simplicity on the case with no reserve price.)

Bidder i wins the object in the SPA when his value is the highest ($V_i = V^{(1)}$) and pays a price equal to the second-highest value, which, on average, equals

$$E[V^{(2)} \mid V_i = V^{(1)}] = \frac{n-1}{n} V_i$$

(by MF6). By RET, we know that the FPA must generate expected revenue equal to that of the SPA, and that for this to happen, bidders in the FPA must pay the same amount (when they win) as they do in the SPA. Since bidders pay their full bid when they win in the FPA, but pay only $\frac{n-1}{n} V_i$ in the SPA, this tells us that the equilibrium bidding strategy in the FPA must be

$$b^{\mathrm{FPA}}(V_i) = \frac{n-1}{n} V_i.$$

That's it!

To check that this solution is indeed an equilibrium, suppose that all other bidders use this bidding strategy and that bidder i has value V_i. Let $S(b_i, V_i)$ denote bidder i's expected surplus when bidding b_i given value V_i, and let $b_i^*(V_i)$ be the bid that maximizes $S(b_i, V_i)$. Bid b_i is high enough to win the auction so long as

$$b_i > \max_{j \neq i} \frac{n-1}{n} V_j;$$

this happens with probability

$$Prob\left[\max_{j \neq i} V_j < b\frac{n}{n-1}\right]$$

$$= \left(b\frac{n}{n-1}\right)^{n-1}$$

(by the same logic behind MF3), and when it happens, generates bidder surplus $V_i - b_i$. So,

$$V_i - b_i. \text{ So, } \; S(b_i, V_i) = \left(b\frac{n}{n-1}\right)^{n-1} \times (V_i - b_i),$$

which, after rearranging, becomes

$$S(b_i, V_i) = \left(\frac{n}{n-1}\right)^{n-1} (b_i^{n-1} V_i - b_i^n).$$

Taking the derivative with respect to b,

$$\frac{dS(b_i, V_i)}{db} = \left(\frac{n}{n-1}\right)^{n-1} ((n-1)(b_i)^{n-2} V_i - n(b_i)^{n-1}),$$

which equals zero only at $b_i = 0$ or $b_i = \frac{n-1}{n} V_i$.

Checking these two bid levels, note that $S(0, V_i) = 0$ while

$$S\left(\frac{n-1}{n} V_i, V_i\right) = \frac{(V_i)^n}{n} > 0.$$

We conclude that the optimal bid $b_i^*(V_i) = \frac{n-1}{n} V_i,$ as desired. In particular, in the special case considered in the text with $n = 2$, $n = 2,\ b_i^*(V_i) = \frac{V_i}{2}.$

E. Equilibrium Bidding Strategies in the All-Pay Auction

By RET, we know that the all-pay auction must generate expected revenue equal to that of the SPA, and that for this to happen, bidders in the all-pay auction must pay the same amount (on average) as they do in the SPA. Since bidders always pay their full bid in the all-pay auction, but pay only $\frac{n-1}{n}V_i$ with probability $Prob\left[\max_{j \neq i} V_j < V_i\right] = (V_i)^{n-1}$ in the SPA, this tells us that the equilibrium bidding strategy in the all-pay auction must be

$$b^{\text{APA}}(V_i) = \frac{n-1}{n}V_i \times (V_i)^{n-1} = \frac{n-1}{n}(V_i)^n.$$

That’s it!

Endnotes

- Given any random variable X, the notation $E[X]$ means "the average (or 'expected') value of X." Similarly, given random variables X and Y, $E[X \mid Y < y]$ means "the average value of X, conditional on Y being less than y." Return to reference 33
- Interested readers can verify these facts for themselves—or you can just trust us. To dive deeper into the fascinating mathematics of these so-called 'order statistics,' consult Barry C. Arnold, N. Balakrishnan, and H. N. Nagaraja, *A First Course in Order Statistics*, Classics in Applied Mathematics (Philadelphia, Pa.: Society for Industrial and Applied Mathematics, 2008). Return to reference 34
- We ignore the zero-probability events in which $V^{(1)} = r$ or $V^{(2)} = r$. Return to reference 35

16■ Strategy and Voting

WHEN MANY OF YOU think about voting, you probably imagine first a national presidential election, then perhaps a local mayoral election, and maybe even an election for class president at your school. But some of you may also be reminded of last year’s Heisman Trophy–winning college football player, the latest Academy Award–winning film, or the most recent Supreme Court decision. *All* of these situations involve voting, although they differ in the number of voters involved, the ballot length or number of choices available to those voters, and the procedures used to tally the votes and determine the final winner. In each case, strategic thinking may play a role in how ballots are marked. And strategic considerations can be critical in choosing the method by which votes are taken and then counted.

Voting procedures vary widely not because some votes elect Oscar winners and others elect presidents, but because certain procedures have attributes that make them better (or worse) for specific voting situations. In the past decade, for example, concerns about how elections based on the plurality rule (the candidate with the most votes wins) encourage the existence of a two-party political system have led to changes in voting rules in more than a dozen U.S. cities.[1] These changes have led in some cases to election outcomes that differed from those that would have arisen under the old plurality-rule system. Jean Quan, the mayor of Oakland, California, for example, won her post in November 2010 under that city’s new ranked-choice voting system despite being the first choice of only 24% of the voters, while the eventual runner-up had 35% of the first-place votes. In the last round of the ranked-choice vote, Quan won 51% of the votes, with 49% going to the runner-up. We investigate such seemingly paradoxical outcomes in Section 2 of this chapter.

Given the fact that different voting procedures can produce different outcomes, you should immediately see the scope for strategic behavior in choosing a procedure that can generate an outcome you prefer. Perhaps, then, you can also imagine a situation in which voters might find it beneficial to vote for someone, or something, that is not their top choice in order to avoid having their absolute last choice option be the winner. This type of strategic behavior is common when the voting procedure allows it. As a voter, you should be aware of the benefits associated with such *strategic misrepresentation of preferences* and of the possibility that others may use such tactics against you.

In this chapter, we first introduce you to the range of voting procedures available and to some of the paradoxical outcomes that can arise when specific procedures are used. We then consider how one might judge the performance of those procedures before addressing the impact of strategic voting. Finally, we present two different versions of a well-known result known as the *median voter theorem*—as a two-person zero-sum game with discrete strategies and with continuous ones.

Endnotes

- This result is known in political science as Duverger’s law. We discuss it in greater detail in Section 4.A.
  Return to reference 1

1 VOTING RULES AND PROCEDURES

Numerous voting procedures are available to help voters choose from a slate of alternatives (that is, candidates or issues). With as few as three available alternatives, election design becomes interestingly complex. We describe in this section a variety of procedures from three broad classes of voting, or vote-aggregation, methods. We have not attempted to provide an exhaustive survey—the number of possible voting procedures is enormous, and the simple taxonomy that we provide here can be broadened extensively by using combinations of these procedures—but rather to give you a flavor of the broader literatures in economics and political science on the subject.[2]

A. Binary Methods

Vote-aggregation methods can be classified according to the number of options or candidates considered by the voters at any given time. Binary methods require voters to choose between only two alternatives at a time. In elections in which there are exactly two candidates, votes can be aggregated by using the well-known principle of majority rule, which simply requires that the alternative with a majority of votes wins. When dealing with a slate of more than two alternatives, pairwise voting—a method consisting of a repetition of binary votes—can be used. Pairwise voting is a multistage procedure; it entails voting on pairs of alternatives in a series of majority votes to determine which alternative is most preferred.

One pairwise voting procedure is called the Condorcet method, after the eighteenth-century French theorist Antoine Nicholas Caritat, marquis de Condorcet. This procedure requires a separate majority vote for each possible pair of alternatives (a “round robin”); if there are n alternatives, there are $n(n-1)/2$ such pairs. If one alternative wins all of its $(n-1)$ binary elections, it wins the entire election and is termed a Condorcet winner. There may be no Condorcet winner, but there cannot be two or more (except in the very unusual circumstances that there are ties in all of the binary elections between the two, or more, alternatives that win all of their other elections). Other pairwise procedures produce “scores” such as the Copeland index, which measures an alternative’s win-loss record in a round robin of contests. The first round of the World Cup soccer tournament uses a type of Copeland index to determine which teams from each group move on to the second round of play.[3]

Another well-known pairwise procedure, used when there are three possible alternatives, is the amendment procedure, required by the parliamentary rules of the U.S. Congress when legislation is brought to a vote. When a bill is brought before Congress, any amended version of the bill must first win a majority vote against the original version of the bill. The winner of that vote is then paired against the status quo, and members vote on whether to adopt the version of the bill that won the first round. The amendment procedure can be used to consider any three alternatives by pairing two in a first-round election and then putting the third up against the winner in a second-round vote.

B. Plurative Methods

Plurative methods allow voters to consider three or more alternatives simultaneously. One group of plurative voting methods applies information on the positions of alternatives on a voter’s ballot to assign points used when tallying ballots; these voting methods are known as positional methods. The familiar plurality rule is a special-case positional method in which each voter casts a single vote for her most preferred alternative. That alternative is assigned a single point when votes are tallied; the alternative with the most votes (or points) wins. Note that a plurality winner need *not* gain a majority, or 51%, of the vote. Thus, for instance, in the 2012 presidential election in Mexico, Enrique Peña Nieto captured the presidency with only 38.2% of the vote; his opponents gained 31.6%, 25.4%, and 2.3% of the vote. Such narrow margins of victory have led to concerns about the legitimacy of past Mexican presidential elections, especially in 2006, when the margin of victory was a mere 0.58 percentage points. Another special-case positional method, the antiplurality method, asks voters to vote against one of the available alternatives or, equivalently, to vote for all but one. For counting purposes, the alternative voted against is allocated −1 point, or else all alternatives except that one receive 1 point while the alternative voted against receives 0.

One of the best-known positional methods is the Borda count, named after Jean-Charles de Borda, a fellow countryman and contemporary of Condorcet. Borda described the new procedure as an improvement on the plurality rule. The Borda count requires voters to rank all of the possible alternatives in an election and to indicate their rankings on their ballot cards. Points are assigned to each alternative on the basis of its position on each voter’s ballot. In a three-person election, for example, the candidate at the top of a ballot gets 3 points, the next candidate 2 points, and the bottom candidate 1 point. After the ballots are collected, each candidate’s points are summed, and the one with the most points wins the election. A Borda count procedure is used in a number of sports-related elections, including professional baseball’s Cy Young Award and college football’s championship elections.

Many variations on the Borda count can be devised simply by altering the rule used for the allocation of points to alternatives based on their positions on a voter’s ballot. One system might allocate points in such a way as to give the top-ranked alternative relatively more than the others—for example, 5 points for the most preferred alternative in a three-way election, but only 2 and 1 for the second- and third-ranked options. In elections with larger numbers of candidates—say, eight—the top two choices on a voter’s ballot might receive preferred treatment, gaining 10 and 9 points, respectively, while the others receive 6 or fewer.

An alternative to these positional plurative methods is the relatively recently invented approval voting method, which allows voters to cast a single vote for each alternative of which they “approve.”[4] Unlike positional methods, approval voting does not distinguish between alternatives on the basis of their positions on the ballot. Rather, all approval votes are treated equally, and the alternative that receives the most approvals wins. In elections in which more than one winner can be selected (in electing a school board, for instance), a threshold level of approvals is set in advance, and alternatives with more than the required minimum of approvals are elected. Proponents of this method argue that it favors relatively moderate alternatives over those at either end of a spectrum; opponents claim that unwary voters could elect an unwanted novice candidate by indicating too many “encouragement” approvals on their ballots. Despite these disagreements, several professional societies and the United Nations have adopted approval voting to elect their officers, and some states have used or are considering using this method for public elections.

C. Mixed Methods

Some multistage voting procedures combine plurative and binary voting in mixed methods. The majority runoff procedure, for instance, is a two-stage method used to reduce a large group of possibilities to a binary decision. In a first-stage election, voters indicate their most preferred alternative, and these votes are tallied. If one candidate receives a majority of votes in the first stage, she wins. However, if there is no majority choice, a second-stage election pits the two most preferred alternatives against each other. The winner is chosen by majority rule in the second stage. French presidential elections use the majority runoff procedure, which can yield unexpected results if three or four strong candidates split the vote in the first round. In the spring of 2002, for example, the far-right candidate Jean-Marie Le Pen came in second, ahead of France's socialist Prime Minister Lionel Jospin, in the first round of the presidential election. This result aroused surprise and consternation among French citizens, 30% of whom hadn't even bothered to vote in the first round, and some of whom had taken it as an opportunity to express their preference for various candidates of the far and fringe left. Le Pen's advance to the runoff election led to considerable political upheaval, although he lost in the end to the incumbent president, Jacques Chirac.

Another mixed procedure consists of voting in successive rounds. Voters consider a number of alternatives in each round of voting. The worst-performing alternative is eliminated after each round, and voters then consider the remaining alternatives in the next round. The elimination continues until only two alternatives remain; at that stage, the method becomes binary, and a final majority runoff determines a winner. A procedure with rounds is used to choose sites for the Olympic Games.

One can eliminate the need for successive rounds of voting by having each voter indicate her preference ordering by ranking all the candidates on a single ballot. Then a single transferable vote method can be used to tally votes. If no alternative receives a majority of all first-place votes, the bottom-ranked alternative is eliminated, and all first-place votes for that candidate are "transferred" to the candidate ranked second on those ballots; similar reallocation occurs in later rounds as additional alternatives are eliminated until a majority winner emerges. This voting method, also referred to as instant runoff voting (IRV) or ranked-choice voting, has been slowly gaining traction. In 2010, IRV was used in a dozen U.S. cities. In 2016, Maine became the first state to use IRV for both primary and general elections for both statewide and national offices (governor, state legislature, U.S. House, and U.S. Senate), and Maine voters reaffirmed IRV in a 2018 referendum. However, some who experimented with IRV have abandoned it. For example, North Carolina used IRV for statewide judicial elections in 2010, but then returned to plurality rule, and Aspen, Colorado used IRV for city council elections in 2008, but then returned to a more traditional runoff system.[5]

In 2015, Democrat Ethan Strimling was elected mayor of Portland, Maine, in a three-way race using instant runoff voting, defeating incumbent mayor Michael Brennan (also a Democrat) and Green Independent Party leader Tom MacMillan. Strimling led the voting from the first to the final round, but some outside observers commented that Strimling might have lost if the election had used the plurality rule—the idea being that Democratic voters might have rallied around the incumbent Brennan to avoid splitting the Democratic vote and allowing MacMillan to win.[6] (You are asked to consider this example in more detail in Exercise S3.)

The single transferable vote method is sometimes combined with other procedures meant to promote proportional representation in an election. Proportional representation means that a state electorate consisting of 55% Republicans, 25% Democrats, and 20% Independents, for example, would yield a body of representatives mirroring the party affiliations of that electorate—in other words, that 55% of the U.S. representatives from such a state would be Republican, and so on. Such methods contrast starkly with the plurality rule, which would elect *all* Republicans (assuming that the voter mix in each district exactly mirrors the overall voter mix in the state). Under proportional representation rules, candidates who attain a certain quota of votes are elected, and others who fall below a certain quota are eliminated, depending on the exact specifications of the voting procedure. Votes for those candidates who are eliminated are again transferred by using the voters' preference orderings. This procedure continues until an appropriate number of candidates from each party is elected. Versions of this type of procedure are used in parliamentary elections in both Australia and New Zealand.

Clearly, there is room for considerable strategic thinking in the choice of a vote-aggregation method, and strategy is important even after the rule has been chosen. We examine some of the issues related to rule making and agenda setting in Section 2. Furthermore, strategic behavior on the part of voters, often called strategic voting or strategic misrepresentation of preferences, can also alter election outcomes under any set of rules, as we will see later in this chapter.

Endnotes

- The classic textbook on this subject, which was instrumental in making game theory popular in political science, is William Riker, *Liberalism against Populism* (San Francisco: W. H. Freeman, 1982). A general survey is "Symposium: Economics of Voting," *Journal of Economic Perspectives*, vol. 9, no. 1 (Winter 1995). An important early research contribution is Michael Dummett, *Voting Procedures* (Oxford: Clarendon Press, 1984). Donald Saari, *Chaotic Elections* (Providence, R.I.: American Mathematical Society, 2000), develops some new ideas that we use later in this chapter. Return to reference 2
- Note that such indexes, or scores, must have precise mechanisms in place to deal with ties; World Cup soccer uses a system that undervalues a tie to encourage more aggressive play. See Barry Nalebuff and Jonathan Levin, "An Introduction to Vote Counting Schemes," *Journal of Economic Perspectives*, vol. 9, no. 1 (Winter 1995), pp. 3–26. Return to reference 3
- Unlike the many voting methods that have histories going back several centuries, the approval voting method was designed and named by then-graduate student Robert Weber in 1971. Weber is now a professor of managerial economics and decision sciences at Northwestern University, specializing in game theory. Return to reference 4
- Some Aspen residents complained that they did not know all of the candidates well enough to rank them, and that more time was needed for them to make informed choices in the runoff. See Carolyn Sackariason, "Aspen Voters to Vote on How They Vote-Again," *Aspen Times*, July 22, 2009. Return to reference 5
- Drew Spencer Penrose, "Seven Ways Ranked Choice Voting is Empowering Voters in 2015," FairVote.org, November 4, 2015, available at https://www.fairvote.org/seven-ways-

ranked-choice-voting-is-empowering-voters-in-2015(accessed May 29, 2019). Return to reference 6

Glossary

binary method
: A class of voting methods in which voters choose between only two alternatives at a time.

majority rule
: A voting method in which the winning alternative is the one that garners a majority (more than 50%) of the votes.

pairwise voting
: A voting method in which only two alternatives are considered at the same time.

multistage procedure
: A voting procedure in which there are multiple rounds of voting.

Condorcet method
: A voting method in which the winning alternative must beat each of the other alternatives in a round-robin of pairwise contests.

Condorcet winner
: The alternative that wins an election run using the *Condorcet method*.

Copeland index
: An index measuring an alternative's record in a round-robin of contests where different numbers of points are allocated for wins, ties, and losses.

amendment procedure
: A procedure in which any amended version of a proposal must win a vote against the original version before the winning version is put to a vote against the status quo.

plurative method
: Any voting method that allows voters to consider a slate of three or more alternatives simultaneously.

positional method
: A voting method that determines the identity of the winning alternative using information on the position of

alternatives on a voter's ballot to assign points used when tallying ballots.

plurality rule
: A voting method in which two or more alternatives are considered simultaneously and the winning alternative is the one that garners the largest number of votes; the winner needs only gain more votes than each of the other alternatives and does not need 50% of the vote as would be true in *majority rule*.

antiplurality method
: A positional voting method in which the electorate is asked to vote against one item on the slate (or to vote for all but one).

Borda count
: A positional voting method in which the electorate indicates its order of preference over a slate of alternatives. The winning alternative is determined by allocating points based on an alternative's position on each ballot.

approval voting
: A voting method in which voters cast votes for all alternatives of which they approve.

mixed method
: A multistage voting method that uses plurative and binary votes in different rounds.

majority runoff
: A two-stage voting method in which a second round of voting ensues if no alternative receives a majority in the first round. The top two vote-getters are paired in the second round of voting to determine a winner.

round
: A single vote within a larger *multistage procedure* that consists of multiple sequentially held votes.

single transferable vote
: A voting method in which each voter indicates her preference ordering over all candidates on a single initial ballot. If no alternative receives a majority of

all first-place votes, the bottom-ranked alternative is eliminated and all first-place votes for that candidate are "transferred" to the candidate listed second on those ballots; this process continues until a majority winner emerges. Also called **instant-runoff voting (IRV)** or **ranked-choice voting**.

instant-runoff voting (IRV)

Same as **single transferable vote**.

ranked-choice voting

Another name for **single transferable vote**.

proportional representation

This voting system requires that the number of seats in a legislature be allocated in proportion to each party's share of the popular vote.

strategic voting

Voting in conformity with your optimal rational strategy found by doing rollback analysis on the game tree of the voting procedure.

strategic misrepresentation of preferences

Refers to strategic behavior of voters when they use rollback to determine that they can achieve a better outcome for themselves by not voting strictly according to their preference orderings.

2 VOTING PARADOXES

Election outcomes can depend critically on the type of procedure used to aggregate votes. Even when people vote according to their true preferences, certain voter preferences or voting procedures can give rise to counterintuitive outcomes. This section describes some of the most famous of those outcomes—the so-called voting paradoxes—as well as some examples of how election results can change under different vote-aggregation methods with no change in voter preferences and no strategic voting.

A. The Condorcet Paradox

The Condorcet paradox is one of the most famous and important of the voting paradoxes.[7] As mentioned earlier, the Condorcet method calls for the winner to be the candidate who gains a majority of votes in each round of a round robin of pairwise comparisons. The paradox arises when no Condorcet winner emerges from this process.

To illustrate the paradox, we construct an example in which three people vote on three alternatives using the Condorcet method. Consider three city councillors (Left, Center, and Right) who are asked to rank their preferences for three alternative welfare policies: one that extends the welfare benefits currently available (call this one Generous, or G), another that decreases available benefits (Decreased, or D), and yet another that maintains the status quo (Average, or A). They are then asked to vote on each pair of policies to establish a *group ranking*, or social ranking. This ranking is meant to describe how the council as a whole judges the merits of the alternative welfare systems.

Suppose Councillor Left prefers to keep benefits as high as possible, whereas Councillor Center is most willing to maintain the status quo, but is concerned about the state of the city budget, and so is least willing to extend welfare benefits. Finally, Councillor Right most prefers reducing benefits, but prefers an increase in benefits to the status quo; she expects that extending benefits will soon cause a serious budget crisis and turn public opinion so much against benefits that a more permanent state of low benefits will result, whereas the status quo could go on indefinitely. We illustrate these preference orderings in Figure 16.1, where the “curly” greater-than symbol, $\succ$, is used to indicate that one alternative is preferred to another. (Technically, $\succ$ is referred to as a *binary ordering relation*.)

With these preferences, if Generous is paired against Average, Generous wins. In the next pairing, of Average against Decreased, Average wins. And in the final pairing, of Generous against Decreased, the vote is again 2 to 1, this time in favor of Decreased. Therefore, if the council votes on alternative pairs of policies, a majority prefer Generous over Average, Average over Decreased, *and* Decreased over Generous. No one policy has a majority over both of the others. The group's preferences are cyclical: G $\succ$ A $\succ$ D $\succ$ G.

Left	Center	Right
G > A > D	A > D > G	D > G > A

FIGURE 16.1 Councillor Preference Orderings for Welfare Policies

---

This cycle of preferences is an example of an intransitive ordering of preferences. The concept of rationality is usually taken to mean that individual preference orderings are transitive (the opposite of intransitive). If someone is given choices A, B, and C, and you know that she prefers A to B and B to C, then transitivity implies that she also prefers A to C. (The terminology comes from the transitivity of numbers in mathematics; for instance, if $3 > 2$ and $2 > 1$, then we know that $3 > 1$.) A transitive preference ordering does not cycle as the social ranking derived in our city council example does; hence, we say that such an ordering is intransitive.

Notice that all of the *councillors* have transitive preference orderings of the three welfare policy alternatives, but the *council* does not. This is the Condorcet paradox: Even if all individual preference orderings are transitive, there is no guarantee that the social preference ordering induced by Condorcet's voting procedure will also be transitive. This result has far-reaching implications for public servants as well as for the general public. It calls into question the basic notion of the "public interest" because such interests may not be easily defined, or may not even exist. Our city council does not have any well-defined set of group preferences among the welfare policies. The lesson is that societies, institutions, or other large groups of people should not always be analyzed as if they acted like individuals.

The Condorcet paradox can even arise more generally. There is no guarantee that the social ranking induced by *any* formal group voting process will be transitive just because individual preferences are. However, some estimates have shown that the paradox is most likely to arise when large groups of people are considering large numbers of alternatives. Smaller groups considering smaller numbers of alternatives are more likely to have similar preferences among those alternatives; in such situations, the paradox is much less likely to appear.[8] In fact, the paradox arose in our example because the council completely disagreed not only about which alternative was best, but also about which was worst. The smaller the group, the less likely such outcomes are to occur.

B. The Agenda Paradox

The second paradox that we consider also entails a binary voting procedure, but this example considers the ordering of alternatives in that procedure. In a parliamentary setting with a committee chair who determines the specific order of voting for a three-alternative election, the chair has substantial power over the final outcome. In fact, the chair can take advantage of the intransitive social preference ordering that arises from some sets of individual preferences and, by selecting an appropriate agenda, manipulate the outcome of the election in any manner she desires.

Consider again the city councillors Left, Center, and Right, who must decide among Generous, Average, and Decreased welfare policies. The councillors' preference rankings of the alternatives were shown in Figure 16.1. Let us now suppose that one of the councillors has been appointed chair of the council by the mayor, and the chair is given the right to decide which two welfare policies get voted on first and which goes up against the winner of that initial vote. With the given set of councillor preferences, as long as the chair knows all of the preference orderings, she can get any outcome that she wants. If Left were chosen as chair, for example, she could orchestrate a win for Generous by setting Average against Decreased in the first round, with the winner to go up against Generous in round two. The fact that any final ordering can be obtained by choosing an appropriate procedure under these circumstances is known as the agenda paradox.

The only determinant of the outcome in our city council example is the ordering of the agenda. Thus, setting the agenda is the real game here, and because the chair sets the agenda, the appointment or election of the chair is the true outlet for strategic behavior. Here, as in many other strategic situations, what appears to be the game (in this case, choice of a welfare policy) is not the true game at all; rather, those participating in the game engage in strategic play at an earlier point (deciding the identity of the chair) and vote according to set preferences in the eventual election.

There is a subtlety here worth noting: The preceding demonstration of the agenda setter's power assumes that in the first round, voters choose between the two alternatives (Average and Decreased) only on the basis of their rankings of these two alternatives, with no regard for the eventual outcome of the procedure. Such behavior is called sincere voting or truthful voting; actually, myopic or nonstrategic voting would be a better name. If Center is a strategic game player, she should realize that if she votes for Decreased in the first round (even though she prefers Average between the pair presented at that stage), then Decreased will win the first round and will also win against Generous in the second round with support from Right. Center prefers Decreased over Generous as the eventual outcome. Therefore, she should do this rollback analysis and vote strategically in the first round. But should she do so if everyone else is also voting strategically? We examine the game of strategic voting and find its equilibrium in Section 4.

C. The Reversal Paradox

Positional voting methods can also lead to paradoxical results. The Borda count, for example, can yield the reversal paradox when the slate of candidates presented to voters changes. This paradox arises in an election with at least four alternatives when one of them is removed from consideration after votes have been submitted, making recalculation necessary.

Suppose there are four candidates for a (hypothetical) special commemorative Cy Young Award to be given to a retired major-league baseball pitcher. The candidates are Steve Carlton, Sandy Koufax, Robin Roberts, and Tom Seaver. Seven prominent sportswriters are asked to rank these pitchers on their ballot cards. The top-ranked candidate on each card will get 4 points; decreasing numbers of points will be allotted to candidates ranked second, third, and fourth.

Across the seven voting sportswriters, there are three different preference orderings of the candidate pitchers; these preference orderings, with the number of writers having each ordering, are shown in Figure 16.2. When the votes are tallied, Seaver gets $(2 \times 3) + (3 \times 2) + (2 \times 4) = 20$ points; Koufax gets $(2 \times 4) + (3 \times 3) + (2 \times 1) = 19$ points; Carlton gets $(2 \times 1) + (3 \times 4) + (2 \times 2) = 18$ points; and Roberts gets $(2 \times 2) + (3 \times 1) + (2 \times 3) = 13$ points. Seaver wins the election, followed by Koufax, Carlton, and Roberts in last place.

Now suppose it is discovered that Roberts is not really eligible for the commemorative award, because he never actually won a Cy Young Award, having reached the pinnacle of his career in the years just before the institution of the award in 1956. This discovery requires points to be recalculated, ignoring Roberts on the ballots. The top spot on each card now gets 3 points, while the second and third spots receive 2 and 1, respectively. Ballots from sportswriters with preference ordering 1, for example, now give Koufax and Seaver 3 and 2 points, respectively, rather than the 4 and 3 from the first calculation; those ballots also give Carlton a single point for last place.

Adding votes with the revised point system shows that Carlton receives 15 points, Koufax receives 14 points, and Seaver receives 13 points. Winner has turned loser as the new results reverse the standings in the original election. No change in preference orderings accompanies this result. The only difference between the two elections is the number of candidates being considered. In Section 3, we identify the key vote-aggregation principle violated by the Borda count that leads to the reversal paradox.

Ordering 1 (2 voters)	Ordering 2 (3 voters)	Ordering 3 (2 voters)
Koufax > Seaver > Roberts > Carlton	Carlton > Koufax > Seaver > Roberts	Seaver > Roberts > Carlton > Koufax

FIGURE 16.2 Sportswriter Preference Orderings for Pitchers

D. Change the Voting Method, Change the Outcome

As should be clear from the preceding discussion, election outcomes are likely to differ under different sets of voting rules. As an example, consider 100 voters who can be broken down into three groups on the basis of their preference rankings of three candidates (A, B, and C), as shown in Figure 16.3. With these preferences as shown, and depending on the vote-aggregation method used, any of the three candidates could win the election.

With a simple plurality rule, candidate A wins with 40% of the vote, even though 60% of the voters rank her lowest of the three candidates. Supporters of candidate A would obviously prefer this type of election. If they had the power to choose the voting method, then the plurality rule, a seemingly fair procedure, would win the election for A in spite of the majority's strong dislike for that candidate.

The Borda count would produce a different outcome. In a Borda system with 3 points going to the most preferred candidate, 2 points to the middle candidate, and 1 to the least preferred candidate, A would get 40 first-place votes and 60 third-place votes, for a total of (40 × 3) + (60 × 1) = 180 points. Candidate B would get 25 first-place votes and 75 second-place votes, for a total of (25 × 3) + (75 × 2) = 225 points; and C would get 35 first-place votes, 25 second-place votes, and 40 third-place votes, for a total of (35 × 3) + (25 × 2) + (40 × 1) = 195 points. With this procedure, B wins, with C in second place and A last. Candidate B would also win with the antiplurality method, in which electors cast votes for all but their least preferred candidate.

And what about candidate C? She could win the election if a majority or an instant runoff system were used. In either method, A and C, with 40 and 35 votes in the first round, would survive to face each other in the runoff. The majority runoff system would call voters back to the polls to consider A and C; the instant runoff system would eliminate B and reallocate B’s votes from Group 2 voters to their next preferred alternative, candidate C. Then, because A is the least preferred alternative for 60 of the 100 voters, candidate C would win the runoff 60 to 40.

Another example of how different procedures can lead to different outcomes can be seen in the case of the 2010 Oakland mayoral election described in the introduction to this chapter. Olympics site selection voting is now done using instant runoff instead of several rounds of plurality-rule voting with elimination. The change was made after similar unusual results in voting for the 1996 and 2000 host cities. In both cases, the plurality winner in all but the final round lost to the one remaining rival city in the last round: Athens lost out to Atlanta for the 1996 Games, and Beijing lost out to Sydney for the 2000 Games.

Group 1 (40 voters)	Group 2 (25 voters)	Group 3 (35 voters)
A > B > C	B > C > A	C > B > A

FIGURE 16.3 Group Preference Orderings for Candidates

Endnotes

- It is so famous that economists have been known to refer to it as *the* voting paradox. Political scientists appear to know better, in that they are far more likely to use its formal name. As we will see, there are any number of possible voting paradoxes, not just the one named for Condorcet. Return to reference 7
- See Peter Ordeshook, *Game Theory and Political Theory* (Cambridge: Cambridge University Press, 1986), p. 58. Return to reference 8

Glossary

Condorcet paradox

Even if all individual voter preference orderings are transitive, there is no guarantee that the social preference ordering generated by Condorcet's voting method will also be transitive.

social ranking

The preference ordering of a group of voters that arises from aggregating the preferences of each member of the group.

intransitive ordering

A preference ordering that cycles and is not *transitive*. For example, a preference ordering over three alternatives A, B, and C is intransitive if A is preferred to B and B is preferred to C but it is not true that A is preferred to C.

transitive ordering

A preference ordering for which it is true that if option A is preferred to B and B is preferred to C, then A is also preferred to C.

agenda paradox

A voting situation where the order in which alternatives are paired when voting in multiple rounds determines the final outcome.

sincere voting

Voting at each point for the alternative that you like best among the ones available at that point, regardless of the eventual outcome. Also called **truthful voting**.

truthful voting

Same as **sincere voting**.

reversal paradox

This paradox arises in an election with at least four alternatives when one of these is removed from consideration after votes have been submitted and the removal changes the identity of the winning alternative.

3 EVALUATING VOTING SYSTEMS

The discussion of the various voting paradoxes in Section 2 suggests that voting methods can suffer from a number of faults that lead to unusual, unexpected, or even unfair outcomes. This suggestion leads us to ask, Is there one voting system with properties that most people would regard as most desirable, including being “fair”—that is, most accurately capturing the preferences of the electorate? Kenneth Arrow listed some such properties and proved that no vote-aggregation system can satisfy them all; the result is his celebrated impossibility theorem.[9]

The technical content of Arrow’s theorem makes it beyond our scope to prove completely. But the sense of the theorem is straightforward. Arrow argued that no vote-aggregation method could conform to all six of the critical principles that he identified:

1. The social or group ranking must rank all alternatives (be complete).
2. It must be transitive.
3. It must satisfy a condition known as *positive responsiveness*, or the Pareto property: Given two alternatives, A and B, if the electorate unanimously prefers A to B, then the aggregate ranking should place A above B.
4. The ranking must not be imposed by external considerations (such as customs) independent of the preferences of individual members of the society.
5. It must not be dictatorial—no single voter should determine the group ranking.
6. And it must be independent of irrelevant alternatives; that is, no change in (i.e., addition to or subtraction from) the set of alternatives should change the rankings of the unaffected alternatives.

Often, the theorem is abbreviated by imposing the first four conditions and focusing on the difficulty of simultaneously obtaining the last two; this simplified form states that we cannot have independence of irrelevant alternatives (IIA) without dictatorship.[10]

You should be able to see immediately that some of the voting methods considered earlier do not conform to all of Arrow’s principles. The requirement of IIA, for example, is violated by the single transferable vote procedure as well as by the Borda count, as we saw in Section 2.C. The Borda count is, however, nondictatorial and satisfies the Pareto property. All the other systems that we have considered satisfy IIA but break down on one of the other principles.

Arrow’s theorem has provoked extensive research into the robustness of his conclusion to changes in the underlying assumptions. Political scientists, economists, and mathematicians have searched for a way to reduce the number of criteria or relax Arrow’s principles minimally to find a procedure that satisfies the criteria without sacrificing the core principles; their efforts have been largely unsuccessful. Most economic and political theorists now accept the idea that some form of compromise is necessary when choosing a vote- or preference-aggregation method. Here are a few prominent examples of such compromises, each representing the approach of a particular field—political science, economics, or mathematics.

A. Black’s Condition

As the discussion in Section 2.A showed, the pairwise voting procedure does not satisfy Arrow’s requirement for transitivity of the social ranking, even when all individual rankings are transitive. One way to surmount this obstacle to meeting Arrow’s conditions, as well as to prevent the Condorcet paradox, is to place restrictions on the preference orderings held by individual voters. Such a restriction, known as the requirement of single-peaked preferences, was put forth by the political scientist Duncan Black in the late 1940s.[11] Black’s seminal paper on group decision making actually pre-dates Arrow’s impossibility theorem and was formulated with the Condorcet paradox in mind, but voting theorists have since shown its relevance to Arrow’s work; in fact, the requirement of single-peaked preferences is sometimes referred to as Black’s condition.

For a preference ordering to be single peaked, it must be possible to order the alternatives being considered along some specific dimension (for example, by the expenditure level associated with each of several alternative policies). To illustrate this requirement, we draw a graph in Figure 16.4, with the specified dimension on the horizontal axis and a voter’s preference ranking (or payoffs) on the vertical axis. For the single-peaked requirement to be met, each voter must have a single ideal or most preferred alternative, and alternatives farther away from the most preferred point must provide steadily lower payoffs. The two voters in Figure 16.4, Mr. Left and Ms. Right, have different ideal points along the policy dimension, but for each, the payoff falls steadily as the policy moves away from his or her ideal point.

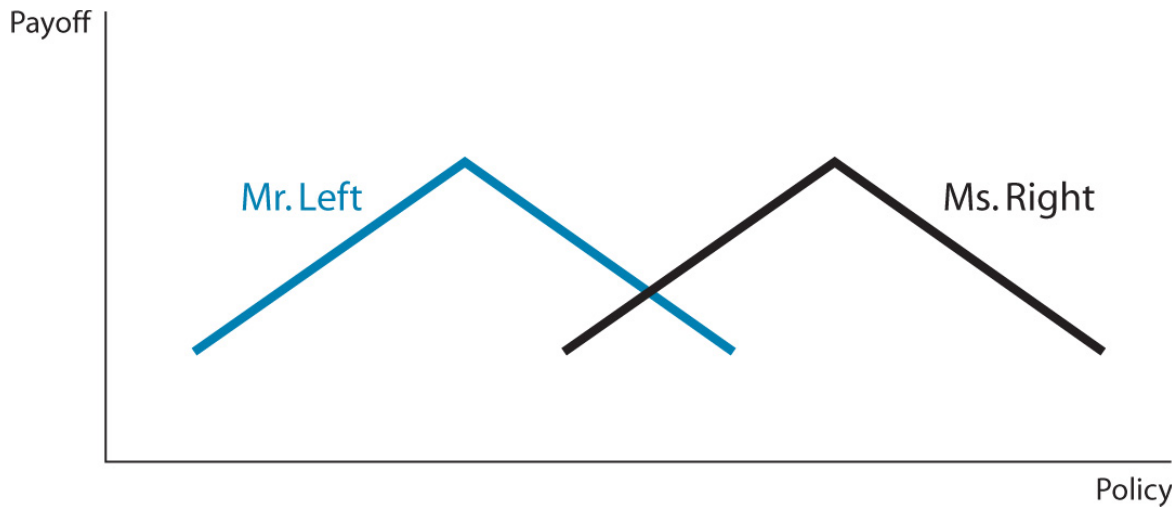


FIGURE 16.4 Single-Peaked Preferences

Black shows that if the preferences of each voter are single peaked, then the pairwise (majority) voting procedure produces a transitive social ranking. The Condorcet paradox is prevented, and pairwise voting satisfies Arrow’s transitivity condition.

B. Robustness

An alternative, more recent method of compromise with Arrow's principles comes from the economic theorists Partha Dasgupta and Eric Maskin.[12] They suggest a new criterion, called robustness, by which to judge voting methods. Robustness is measured by considering how often a voting procedure that is nondictatorial and that satisfies IIA as well as the Pareto property also satisfies the requirement of transitivity: For how many sets of voter preference orderings does such a procedure satisfy transitivity?

With the use of the robustness criterion, simple majority rule can be shown to be *maximally robust*—that is, it is nondictatorial, satisfies IIA and the Pareto property, and provides transitive social rankings for the largest possible set of voter preference orderings. Behind majority rule on the robustness scale lie other voting procedures, including the Borda count and the plurality rule. The robustness criterion is appealing in its ability to establish one of the most commonly used voting procedures—the one most often associated with the democratic process—as a candidate for the best vote-aggregation procedure.

C. Intensity Ranking

Another class of attempts to escape from Arrow's impossibility theorem focuses on the difficulty of satisfying Arrow's IIA requirement. A recent theory of this kind comes from the mathematician Donald Saari.[13] He suggests that a vote-aggregation method might use more information about voters' preferences than is contained in their mere ordering of any pair of alternatives, X and Y; rather, it could take into account each individual voter's *intensity* of preference between that pair of alternatives. This intensity can be measured by counting the number of other alternatives, Z, W, V, . . . , that a voter places between X and Y. Saari therefore replaces the IIA condition, number 6 of Arrow's principles, with a different condition, which he labels IBI (intensity of binary independence) and which we will number 6′:

6′. Society's relative ranking of any two alternatives should be determined only by (1) each voter's relative ranking of the pair and (2) the intensity of this ranking.

This condition is weaker than IIA because it effectively applies IIA only to such additions or deletions of "irrelevant" alternatives that do not change the intensity of people's preferences between the "relevant" ones. With this revision, the Borda count satisfies the Arrow theorem. It is the only positional voting method that does so.

Saari also hails the Borda count as the only procedure that appropriately observes ties within collections of ballots, a criterion that he argues is essential for a good vote-aggregation system to satisfy. Ties can occur two ways: through Condorcet terms or through reversal terms within voter preference orderings. In a three-candidate election among alternatives A, B, and C, the Condorcet terms are the preference orderings $A \succ B \succ C$, $B \succ C \succ A$, and $C \succ A \succ B$. In a set of three ballots with these preferences appearing on one ballot apiece, the ballots should logically offset one another, or constitute a tie.

Reversal terms are preference orderings that contain a reversal in the location of a *pair* of alternatives. In the same election, two ballots with preference orderings of A ≻ B ≻ C and B ≻ A ≻ C should logically lead to a tie in a pairwise contest between A and B. Only the Borda procedure treats collections of ballots with Condorcet terms or reversal terms as tied. Although the Borda count can lead to the reversal paradox, as shown in Section 2.C, it retains many proponents. The *only* time that the Borda procedure produces paradoxical results is when alternatives are dropped from consideration after ballots have been collected. Because such results can be prevented by using only ballots for the complete set of final candidates, the Borda procedure has gained favor in some circles as one of the best vote-aggregation methods.

Other researchers have made different suggestions regarding criteria that a good aggregation system should satisfy. Some of them include the *Condorcet criterion* (that a Condorcet winner should be selected by a voting system, if such a winner exists), the *consistency criterion* (that an election including all voters should elect the same alternative as would two elections held for two arbitrary divisions of the entire set of voters), and lack of manipulability (that a voting system should not encourage strategic voting on the part of voters). We cannot consider each of these suggestions at length, but we do address strategic voting in the following section.

Endnotes

- A full description of this theorem, often called "Arrow's general possibility theorem," can be found in Kenneth Arrow, *Social Choice and Individual Values*, 2nd ed. (New York: Wiley, 1963). Return to reference 9
- See Nicholson and Snyder's treatment of Arrow's impossibility theorem in their *Microeconomic Theory*, 11th ed. (New York: Cengage Learning, 2012), Chapter 19 , for more detail at a level appropriate for intermediate-level economics students. Return to reference 10
- Duncan Black, "On the Rationale of Group Decision-Making," *Journal of Political Economy*, vol. 56, no. 1 (February 1948), pp. 23–34. Return to reference 11
- See Partha Dasgupta and Eric Maskin, "On the Robustness of Majority Rule," *Journal of the European Economic Association*, vol. 6 (2008), pp. 949–73. Return to reference 12
- For more precise information about Saari's work on Arrow's theorem, see D. Saari, "Mathematical Structure of Voting Paradoxes I: Pairwise Vote," *Economic Theory*, vol. 15 (2000), pp. 1–53. Additional information on this result and on the robustness of the Borda count can be found in D. Saari, *Chaotic Elections* (Providence, R.I.: American Mathematical Society, 2000). Return to reference 13

Glossary

impossibility theorem

A theorem that indicates that no preference aggregation method can satisfy the six critical principles identified by Kenneth Arrow.

single-peaked preferences

A preference ordering in which alternatives under consideration can be ordered along some specific dimension and each voter has a single ideal or most-preferred alternative with alternatives farther away from the most-preferred point providing steadily lower payoffs. Also called **Black's condition**.

Black's condition

Same as the condition of **single-peaked preferences**.

robustness

A measure of the number of sets of voter preference orderings that are nondictatorial, satisfy independence of irrelevant alternatives and the Pareto property, and also produce a transitive *social ranking*.

Condorcet terms

A set of ballots that would generate the Condorcet paradox and that should together logically produce a tied vote among three possible alternatives. In a three-candidate election among A, B, and C, the Condorcet terms are three ballots that show A preferred to B preferred to C; B preferred to C preferred to A; C preferred to A preferred to B.

reversal terms

A set of ballots that would generate the *reversal paradox* and that should together logically produce a tied vote between a pair of alternatives. In a three-candidate election among A, B, and C, the reversal terms are two ballots that show a reversal in the location of a pair of alternatives. For example, one ballot with A preferred to B preferred to C and another with B preferred to A preferred to C should produce a tie between A and B.

4 STRATEGIC VOTING

Several of the voting systems that we have considered yield considerable scope for strategic misrepresentation of preferences by voters. In Section 2.B, we showed how the power of an agenda-setting chair could be countered by a councillor voting in the first round against her true preference, so as to knock out her least preferred alternative and send a more preferred one into the second round. More generally, voters can choose to vote for candidates, issues, or policies that are not actually their most preferred among the alternatives presented in an early round of voting if such behavior can alter the final election results in their favor. In this section, we consider a number of ways in which strategic voting behavior can affect elections.

A. Plurality Rule

Plurality-rule elections, often perceived as the fairest by many voters, still provide opportunities for strategic behavior. In U.S. presidential elections, for instance, there are generally two candidates, representing the two major political parties, in contention. When such a race is relatively close, there is potential for a third candidate to enter the race and divert votes away from the leading candidate; if the entry of this third player truly threatens the chances of the leader winning the election, the third player is called a spoiler.

In the 2016 U.S. presidential election, Libertarian candidate Gary Johnson and Green Party candidate Jill Stein got, respectively, 3.3% and 1.1% of the popular vote, earning zero electoral votes but potentially still swinging the election. Johnson had a centrist appeal and so probably drew voters from both Hillary Clinton, the Democratic candidate, and Donald Trump, the Republican. Jill Stein, on the other hand, got her support from the far left and hence probably drew voters mainly from Hillary Clinton. If those Stein voters had instead cast their ballots for Clinton, they would have flipped Michigan, Pennsylvania, and Wisconsin in Clinton's favor, which would have been enough to put Clinton in the White House.[14] Stein also harshly criticized Clinton during the campaign, potentially causing some voters to stay at home who would have otherwise voted for Clinton, even as she praised Trump (saying at one point that "maybe he's the peace president"[15]) and denied that Russia was in any way interfering in the election. However, as later revealed by the special counsel investigation led by Robert Mueller, Russian agents aggressively promoted Stein's candidacy (and anti-Clinton views) on social media as part of their broader effort to elect Trump.[16]

In plurality-rule elections with a spoiler candidate, those who prefer the spoiler to the leading major candidate but least prefer the trailing major candidate may do best to strategically misrepresent their preferences to prevent the election of their least favorite candidate. That is, you should vote for the leader in such a case, even though you would prefer the spoiler, because the spoiler is unlikely to garner a plurality; voting for the leader then prevents the trailing candidate, your least favorite, from winning.[17]

Of course, sometimes a third-party candidate who appears to be a spoiler might actually have a shot at winning an election. If voters who prefer the third-party candidate strategically misrepresent their preferences, voting instead for a major-party candidate, the third-party candidate could lose even if she has more support than anyone else. For instance, in the 1992 U.S. presidential election, businessman Ross Perot ran as a third-party candidate against Republican George H. W. Bush and Democrat Bill Clinton. In a survey conducted after the election, *Newsweek* found that a plurality of 40% of voters surveyed said they would have voted for Perot if they had thought he could have won. However, because they thought he was likely to lose, they cast their ballots for their second-favorite candidate (Bush or Clinton), causing Perot to lose the election with only 18.9% of the popular vote and zero electoral votes.[18]

Because of strategic voting, third parties in the American political system face a “chicken-and-egg problem”: They cannot successfully challenge the two major parties until people believe that they can make a successful challenge. One way to crack this problem would be to switch from plurality-rule elections to another system, such as instant runoff voting, in which voters have an incentive to vote according to their true preferences. For example, consider again the 1992 election between Ross Perot (P), Bill Clinton (C), and George H. W. Bush (B), and suppose that voters’ preferences among the three candidates were as follows: $P \succ B \succ C$ (20%); $P \succ C$

$\succ$ B (20%); C $\succ$ P $\succ$ B (30%); C $\succ$ B $\succ$ P (5%); B $\succ$ P $\succ$ C (20%); and B $\succ$ C $\succ$ P (5%). In the actual plurality-rule election, some voters who preferred Perot voted instead for their second-favorite candidate, fearing that Perot was a spoiler with no chance of winning. Using our (imagined) numbers, if half of Perot supporters voted strategically in this way, then Perot would get 20%, Clinton would get 45%, and Bush would get 35%; this is roughly what transpired in the actual election.

But what if, instead, the election had been held using instant runoff voting? Voters who preferred Perot would have had an incentive to vote according to their true preferences, listing him as their favorite candidate and then either Bush or Clinton as their second favorite. (If Perot had turned out to have the least support, their second-favorite votes would then have been tallied in the instant runoff between Bush and Clinton.) With all votes reflecting voters' true preferences, the first-round votes would have gone 40% to Perot, 35% to Clinton, and 25% to Bush, at which point Bush would have lost, and his supporters' votes would have been redistributed to their favorite remaining candidate. The instant runoff would have added 20% to Perot's support and 5% to Clinton's, leaving Perot the winner with 60% in the final tally.

In elections for legislatures, where many candidates are chosen, the performance of third parties is very different under a system of proportional representation of the whole population in the whole legislature than under a system of plurality in separate constituencies. Britain uses the constituency and plurality system. In the past 50 years, the Labor and Conservative Parties have shared power. The Liberal Party, despite sizable third-place support in the electorate, has suffered from strategic voting and therefore has had disproportionately few seats in Parliament. Italy uses the nationwide candidate list and proportional representation system; there is no need to vote strategically in such a system, and even small parties can have a significant presence in the legislature. Often, no party has a clear majority of seats, and small parties can affect policy through bargaining for alliances.

A party cannot flourish if it is largely ineffective in influencing a country’s political choices. Therefore, we tend to see just two major parties in countries with the plurality system and several parties in those with the proportional representation system. Political scientists call this observation *Duverger’s law*. The plurality system tends to produce only two major parties—often one of them with a clear majority of seats in the legislature—and therefore more decisive government. But it runs the risk that the minority’s interests will be overlooked—that is, of producing a “tyranny of the majority.” A proportional representation system gives more of a voice to minority views. But it can produce inconclusive bargaining for power and legislative gridlock. Interestingly, each country seems to believe that its system performs worse than the other and considers switching; in Britain, there are strong voices calling for proportional representation, and Italy has been seriously considering a constituency system.

B. Pairwise Voting

When you know that you are bound by a pairwise voting method such as the amendment procedure, you can use your prediction of the second-round outcome to determine your optimal voting strategy in the first round. It may be in your interest to appear committed to a particular candidate or policy in the first round, even if it is not your most preferred alternative, so that your least favorite alternative cannot win the entire election in the second round.

We return here to our example of the city council with an agenda-setting chair from Section 2.B. Here, we add the assumption that, because they have worked together closely for so long, all three councillors' preference orderings (shown in Figure 16.1) are known to the entire council. Suppose Councillor Left, who most prefers the Generous (G) welfare package, is appointed chair and sets the Average (A) and Decreased (D) policies against each other in a first vote, with the winner facing off against G in the second round. If the three councillors vote strictly according to their preferences, A will beat D in the first vote and G will then beat A in the second vote; the chair's preferred outcome will be chosen. The city councillors are likely to be well-trained strategists, however, who can look ahead to the final round of voting and use rollback to determine which way to vote in the opening round.

In the scenario just described, Councillor Center's least preferred policy will be chosen in the election. Therefore, rollback analysis says that she should vote strategically in the first round to alter the election's outcome. If Center votes for her most preferred policy in the first round, she will vote for A, which will then beat D in that round and lose to G in round two. However, she could instead vote strategically for D in the first round, which would lift D over A on the first vote. Then, when D is set up against G in the second round, G will lose to D. Councillor Center's misrepresentation of her preference ordering with respect to A and D helps her to change the winner of the election from G to D. Although D is not her most preferred outcome, it is better than G from her perspective.

This strategy works well for Center if she can be sure that no other strategic votes will be cast in the election. Thus, we need to analyze both rounds of voting fully to verify the Nash equilibrium strategies for the three councillors. We do so by using rollback on the two simultaneous-vote rounds of the election, starting with the two possible second-round contests, A versus G and D versus G.

Figure 16.5 illustrates the outcomes that arise in each of the possible second-round elections. The two tables in Figure 16.5a show the winning policy (not payoffs to the players) when A has won the first round and is pitted against G; the tables in Figure 16.5b show the winning policy when D has won the first round. In both cases, Councillor Left chooses the row of the final outcome, Center chooses the column, and Right chooses the actual table (left or right).

You should be able to establish that each councillor has a dominant strategy in each second-round election. In the A versus G election, Left's dominant strategy is to vote for G, Center's dominant strategy is to vote for A, and Right's dominant strategy is to vote for G; G will win this election. If the councillors consider D versus G, Left's dominant strategy is still to vote for G, and Right and Center both have a dominant strategy to vote for D; in this vote, D wins. A quick check shows that all the councillors vote according to their true preferences in this round. Thus, these dominant strategies are all the same: "Vote for the alternative that I prefer." Because there is no future to consider in the second-round vote, the councillors simply vote for whichever policy ranks higher in their preference ordering.[19]

(a) A versus G election

RIGHT votes:

A

		CENTER	
		A	**G**
LEFT	A	A	A
	G	A	G

G

		CENTER	
		A	**G**
LEFT	A	A	G
	G	G	G

(b) D versus G election

RIGHT votes:

D

		CENTER	
		D	**G**
LEFT	D	D	D
	G	D	G

G

		CENTER	
		D	**G**
LEFT	D	D	G

	CENTER		
		D	G
	G	G	G

FIGURE 16.5 Election Outcomes in Two Possible Second-Round Votes

We can now use the results from our analysis of Figure 16.5 to consider optimal strategies in the first round of voting, in which the councillors choose between policies A and D. Because we know how the councillors will vote in the next round given each winner here, we can show the outcome of the entire election in each case (Figure 16.6).

As an example of how we arrived at these outcomes, consider the G in the upper-left cell of the right-hand table in Figure 16.6. The outcome in that cell is obtained when Left and Center both vote for A in the first round while Right votes for D. Thus, A and G are paired in the second round, and as we saw in Figure 16.5, G wins. The other outcomes are derived in similar fashion.

Given the outcomes in Figure 16.6, Councillor Left (who is the chair and has set the agenda) has a dominant strategy to vote for A in this first round. Similarly, Councillor Right has a dominant strategy to vote for D. Neither of these councillors misrepresents her preferences or votes strategically in either round. Councillor Center, however, has a dominant strategy to vote for D here even though she strictly prefers A to D. As the preceding discussion suggested, she has a strong incentive to misrepresent her preferences in the first round of voting, and she is the only one who votes strategically. Center’s behavior changes the winner of the election from G (the winner without strategic voting) to D.

RIGHT votes:

A

		CENTER	
		A	D
LEFT	A	G	G
	D	G	D

D

		CENTER	
		A	D
LEFT	A	G	D
	D	D	D

FIGURE 16.6 Election Outcomes Based on First-Round Votes

Remember that the chair, Councillor Left, set the agenda in the hope of having her most preferred alternative chosen. Instead, her *least* preferred alternative has prevailed. So it might appear that the power to set the agenda may not be so beneficial after all. But Councillor Left should anticipate the other councillors' strategic behavior. Then she can choose the agenda so as to take advantage of her understanding of games of strategy. In fact, if she sets D against G in the first round and then the winner against A, the Nash equilibrium outcome is G, the chair's most preferred outcome. With that agenda, Right misrepresents her preferences in the first round to vote for G over D to prevent A, her least preferred outcome, from winning. You should verify that setting this agenda is Councillor Left's optimal strategy. In the full voting game, where setting the agenda is considered an initial, pre-voting round, we should expect to see the Generous welfare policy adopted when Councillor Left is chair.

We can also see an interesting pattern emerge when we look more closely at voting behavior in the strategic version of the election. There are pairs of councillors who vote "together" (i.e., for the same policy) in both rounds. Under the original agenda (with A versus D in the first round), Right and Center vote together in both rounds, and in the suggested alternative agenda (with D versus G in the first round), Right and Left vote together in both rounds. In other words, a sort of long-lasting coalition has formed between two councillors in each case.

Strategic voting of this type appears to have taken place in the U.S. Congress on more than one occasion. One example was a federal school construction funding bill considered in 1956.[20] Before being brought to a vote against the status quo of no funding, the bill was amended in the House of Representatives to require that the funding be offered only to states with no racially segregated schools. Under the parliamentary voting rules of Congress (described in Section 1.A), a vote on whether to accept this so-called Powell Amendment was taken first, with the winning version of the bill considered afterward. Political scientists who have studied the history of this bill argue that opponents of school funding strategically misrepresented their preferences regarding the amendment in an effort to defeat the full bill. A key group of representatives voted for the amendment, but then joined opponents of racial integration in voting against the full bill in the final vote; the bill was defeated. The voting records of this group indicate that many of them had voted against racial integration in other circumstances, implying that their vote for integration in the case of this amendment was merely an instance of strategic voting and not an indication of their true feelings regarding school integration.

C. Strategic Voting with Incomplete Information

The analysis in Section 4.B showed that sometimes group members have incentives to vote strategically to prevent their least preferred alternative from winning an election. Our example assumed that the council members knew the preference orderings that were possible and how many other councillors had those preference orderings. Now suppose their information is incomplete: Each council member knows the possible preference orderings, her own actual ordering, and the probability that each of the others has each particular ordering, but not the actual distribution of the different preference orderings among the other councillors. In this situation, each councillor's strategy needs to be conditioned on her beliefs about that distribution and on her beliefs about how sincere the other councillors' votes will be.[21]

For example, suppose we still have a three-member council considering the three alternative welfare policies described earlier, following the (original) agenda set by Councillor Left; that is, the council considers policies A and D in the first round of voting, with the winner facing G in the second round. We assume that there are still three different possible preference orderings, as illustrated in Figure 16.1, and that the councillors know that these orderings are the only possibilities. The difference is that no councillor knows for sure exactly how many other councillors have each preference ordering. Rather, each councillor knows her own type, and she knows that there is some positive probability of observing each type of voter (Left, Center, or Right), with the probabilities p_L, p_C, and p_R summing to 1.

We saw earlier that all three councillors vote truthfully in the last round of balloting. We also saw that Left-type and

Right-type councillors vote truthfully in the first round as well. This result remains true in the incomplete information case. Right-type voters prefer to see D win the first-round election; given this preference, Right always does at least as well by voting for D over A (if both other councillors have voted the same way) as she would by voting otherwise, and she sometimes does better by voting this way (if the other two votes split between D and A). Similarly, Left-type voters prefer to see A survive to vie against G in round two; these voters always do at least as well as otherwise—and sometimes do better—by voting for A over D.

At issue, then, is only the behavior of the Center-type voters. Because they do not know the types of the other councillors, and because they have an incentive to vote strategically with some preference distributions—specifically, the case in which it is known for certain that there is one voter of each type—their behavior will depend on the probabilities that the various voter types occur within the council. We consider here one of two polar cases, in which a Center-type voter believes that other Center types will vote truthfully, and we look for a symmetric, pure-strategy Nash equilibrium. The other case, in which she believes that other Center types will vote strategically, is taken up in Exercise U9.

To make outcome comparisons possible, we specify the payoffs for the Center-type voters associated with the possible winning policies. Center-type preferences are A $\succ$ D $\succ$ G. Suppose that if A wins, Center types receive a payoff of 1, and if G wins, Center types receive a payoff of 0. If D wins, Center types receive some intermediate payoff, call it u, where $0 < u < 1$.

Now suppose our Center-type councillor must decide how to vote in the first round (A versus D) in an election in which she believes that both other councillors will vote truthfully, regardless of their type. If both other councillors choose either A or D, then Center's vote is immaterial to the final outcome; she is indifferent between A and D. If the other two councillors split their votes, however, then Center can influence the election outcome. Her problem is that she needs to decide whether to vote truthfully herself.

If the other two councillors split between A and D, and if both are voting truthfully, then the vote for D must have come from a Right-type voter. But the vote for A could have come from *either* a Left type *or* a (truthful) Center type. If the A vote came from a Left-type voter, then Center knows that there is one voter of each type. If she votes truthfully for A in this situation, A will win the first round but lose to G in the end; Center's payoff will be 0. If Center votes strategically for D, D beats A and G, and Center's payoff is u. In contrast, if the A vote came from a Center-type voter, then Center knows there are two Center types and a Right type, but no Left type, on the council. In this case, a truthful vote for A helps A win the first round, and then A also beats G by a vote of 2 to 1 in round two; Center gets her highest payoff of 1. If Center were to vote strategically for D, D would win both rounds again, and Center would get u.

To determine Center's optimal strategy, we need to compare her expected payoff from truthful voting with her expected payoff from strategic voting. When she votes truthfully for A in the first round, Center's payoff depends on how likely it is that the other A vote comes from a Left type versus a Center type. Those probabilities are straightforward to calculate. The probability that the other A vote comes from a Left type is just the probability of a Left type being one of the remaining voters, or $p_L/(p_L + p_C)$; similarly, the probability that the other A vote comes from a Center type is $p_C/(p_L + p_C)$. Then Center's payoffs from truthful voting are 0 with probability $p_L/(p_L + p_C)$ and 1 with probability $p_C/(p_L + p_C)$, so her expected payoff is $p_C/(p_L + p_C)$. With a strategic vote for D, D wins regardless of the identity of the third voter—D wins with certainty—so Center's expected payoff is just u. Center's final decision is to vote truthfully as long as $p_C/(p_L + p_C) > u$.

Note that Center's decision-making condition is an intuitively reasonable one. If the probability of there being other Center-type voters is large, or relatively larger than the probability of there being a Left-type voter, then the Center types vote truthfully. Voting strategically is useful to Center only when she is the only voter of her type on the council.

We add two additional comments on the existence of imperfect information and its implications for strategic behavior. First, if the number of councillors, n, is larger than three but odd, then the expected payoff to a Center type from voting strategically remains equal to u, and $[p_C/(p_L + p_C)]^{(n - 1)/2}$ is her expected payoff from voting truthfully.[22] Thus, a Center type should vote truthfully only when $[p_C/(p_L + p_C)]^{(n - 1)/2} > u$. Because $p_C/(p_L + p_C) < 1$ and $u > 0$, this inequality will *never* hold for large enough values of n. This result tells us that a truthful-voting equilibrium can never persist in a large enough council! Second, imperfect information about the preferences of other voters yields additional scope for strategic behavior. With agendas that include more than two rounds of voting, voters can use their early-round votes to signal their types. The extra rounds give other voters the opportunity to update their prior beliefs about the probabilities p_C, p_L, and p_R and a chance to act on that information. With only two rounds of pairwise votes, there is no time to use any information gained during round one, because truthful voting is a dominant strategy for all voters in the final round.

D. Scope for Strategic Voting

The extent to which a voting procedure is susceptible to strategic misrepresentation of preferences—that is, to strategic voting of the types illustrated in this section—is another topic that has generated considerable interest among voting theorists. Arrow's principles do not require invulnerability to strategic behavior, but the literature has considered how such a requirement would relate to his conditions. Similarly, theorists have considered the scope for strategic voting in various procedures, producing rankings of voting methods.

In an election with a large number of voters, any individual's vote, and likewise any individual's strategic voting, is unlikely to have any effect; strategic voting requires collective action by a group of like-minded voters, which is quite difficult. Strategic voting is more likely to have an effect in smaller electorates, for example, in committees rather than in nationwide elections.

The economist William Vickrey, perhaps better known for his Nobel Prize–winning work on auctions (see Chapter 15), did some of the earliest work considering strategic behavior of voters. He pointed out that procedures satisfying Arrow's IIA condition were most immune to strategic behavior. He also set out several conditions under which strategic voting is more likely to be attempted and to be successful. In particular, he noted that situations with smaller numbers of informed voters and smaller sets of available alternatives may be most susceptible to strategic voting, given a voting method that is itself susceptible to such behavior. This result means, however, that weakening the IIA requirement to help voting procedures satisfy Arrow's conditions makes way for procedures that are more prone to strategic voting. In particular, Saari's intensity ranking version of IIA (called

IBI), mentioned in Section 3.C, may allow more procedures to satisfy this modified version of Arrow's theorem but may simultaneously allow more of those procedures that are susceptible to strategic voting to do so.

Like Arrow's impossibility theorem on the limits of vote aggregation, the most general result on invulnerability to strategic voting is a negative one. Specifically, the Gibbard–Satterthwaite theorem shows that if there are three or more alternatives to consider, the only voting procedure that prevents strategic voting is dictatorship: One voter is assigned the role of dictator, and her preferences determine the election outcome.[23] Combining the Gibbard–Satterthwaite outcome with Vickrey's discussion of IIA may help the reader understand why Arrow's theorem is often reduced to a consideration of which procedures can simultaneously satisfy nondictatorship and IIA.

Finally, some theorists have argued that voting procedures should be evaluated not on their ability to satisfy Arrow's conditions, but on their vulnerability to strategic voting. The relative vulnerability of a voting system can be determined by the amount of information about the preferences of other voters that is required by voters to vote strategically and alter an election successfully. Some research based on this criterion suggests that of the procedures so far discussed, the plurality rule is the most vulnerable to strategic voting (that is, requires the least information). In decreasing order of vulnerability are approval voting, the Borda count, the amendment procedure, majority rule, and the single transferable vote method (IRV).[24]

It is important to note that this ranking of voting procedures by level of vulnerability to strategic voting depends only on the amount of information necessary to alter the outcome; it is not based on the ease of putting such information to good use or on whether strategic voting is most easily achieved by individual voters or by groups. In practice, strategic voting by *individual* voters to alter the outcome in a plurality-rule election is quite difficult.

Endnotes

- "Jill Stein: Democratic Spoiler Or Scapegoat?" FiveThirtyEight Chat, December 7, 2016, available at https://fivethirtyeight.com/features/jill-stein-democratic-spoiler-or-scapegoat/(accessed May 29, 2019). Return to reference 14
- Mike Pesca, "Jill Stein Thinks Nuclear War Is Less Likely under Trump," *Slate*, October 19, 2016, available at https://slate.com/news-and-politics/2016/10/jill-stein-thinks-nuclear-war-is-less-likely-under-trump.xhtml(accessed May 29, 2019). Return to reference 15
- According to Andrew Weiss, a Russian expert at the Carnegie Endowment for International Peace, "The Russian embrace of fringe voices like Stein goes back more than a decade to the earlier days of RT [Russia Today]," an English-language television network funded by the Russian government on which Stein appeared as a frequent guest. See Robert Windrem, "Russians Launched Pro–Jill Stein Social Media Blitz to Help Trump Win Election, Reports Say," *NBC News*, Dec. 22, 2018. Return to reference 16
- Note that an approval voting method would not suffer from this same problem. Return to reference 17
- "Ross Reruns," *Newsweek*, Special Election Recap Issue, November 18, 1996, p. 104. Return to reference 18
- It is a general result in the voting literature that voters faced with pairs of alternatives will always vote truthfully at the last round of voting. Return to reference 19
- A more complete analysis of the case can be found in Riker, *Liberalism against Populism*, pp. 152–57. Return to reference 20
- This result can be found in P. Ordeshook and T. Palfrey, "Agendas, Strategic Voting, and Signaling with Incomplete Information," *American Journal of Political Science*, vol. 32, no. 2 (May 1988), pp. 441–66. The structure of the

example to follow is based on Ordeshook and Palfrey's analysis. Return to reference 21

- A Center type can affect the election outcome only if all other votes are split evenly between A and D. Thus, there must be exactly $(n - 1)/2$ Right-type voters choosing D in the first round and $(n - 1)/2$ other voters choosing A. If those A voters are Left types, then A won't win the second-round election, and Center will get 0 payoff. For Center to get a payoff of 1, it must be true that all of the other A voters are Center types. The probability of this occurring is $[p_C/(p_L+p_C)]^{(n-1)/2}$; then Center's expected payoff from voting truthfully is as stated. See Ordeshook and Palfrey, p. 455. Return to reference 22
- For the theoretical details on this result, see A. Gibbard, "Manipulation of Voting Schemes: A General Result," *Econometrica*, vol. 41, no. 4 (July 1973), pp. 587–601, and M. A. Satterthwaite, "Strategy-Proofness and Arrow's Conditions," *Journal of Economic Theory*, vol. 10 (1975), pp. 187–217. The theorem carries both their names because each proved the result independently of the other. Return to reference 23
- H. Nurmi's classification can be found in his *Comparing Voting Systems* (Norwell, Mass.: D. Reidel, 1987). Return to reference 24

Glossary

spoiler
: Refers to a third candidate who enters a two-candidate race and reduces the chances that the leading candidate actually wins the election.

Gibbard–Satterthwaite theorem
: With three or more alternatives to consider, the only voting method that prevents strategic voting is dictatorship; one person is identified as the dictator and her preferences determine the outcome.

5 THE MEDIAN VOTER THEOREM

All of the preceding sections have focused on the behavior, strategic and otherwise, of voters in elections. However, strategic analysis can also be applied to *candidate* behavior in such elections. Given a particular distribution of voters and voter preferences, candidates will, for instance, need to determine their optimal strategies for building their political platforms. The median voter theorem tells us that when there are just two candidates in an election, when voters are distributed in a “reasonable” way along the political spectrum, and when each voter has “reasonably” consistent (meaning singled-peaked) preferences, both candidates will position themselves on the political spectrum at the same place as the median voter. The median voter is the “middle” voter in that distribution—more precisely, the one at the 50th percentile.

The full game here has two stages. In the first stage, candidates choose their locations on the political spectrum. In the second stage, voters elect one of the candidates. The second-stage game is open to all the varieties of strategic misrepresentation of preferences discussed earlier; hence, we have reduced the choice of candidates to two for our analysis to prevent such behavior from arising in equilibrium. With only two candidates, second-stage votes directly correspond to voter preferences, and the first-stage location decisions of the candidates remain the interesting part of the larger game. It is in that first stage that the median voter theorem defines Nash equilibrium behavior.

A. Discrete Political Spectrum

Let us first consider a population of 90 million voters, each of whom has a preferred position on a five-point political spectrum: Far Left (FL), Left (L), Center (C), Right (R), or Far Right (FR). We suppose that these voters are spread across the political spectrum. The discrete distribution of their locations is shown by a histogram, or bar chart, in Figure 16.7. The height of each bar indicates the number of voters located at that position. In this example, we have supposed that, of the 90 million voters, 40 million are Left, 20 million are Far Right, and 10 million each are Far Left, Center, and Right.

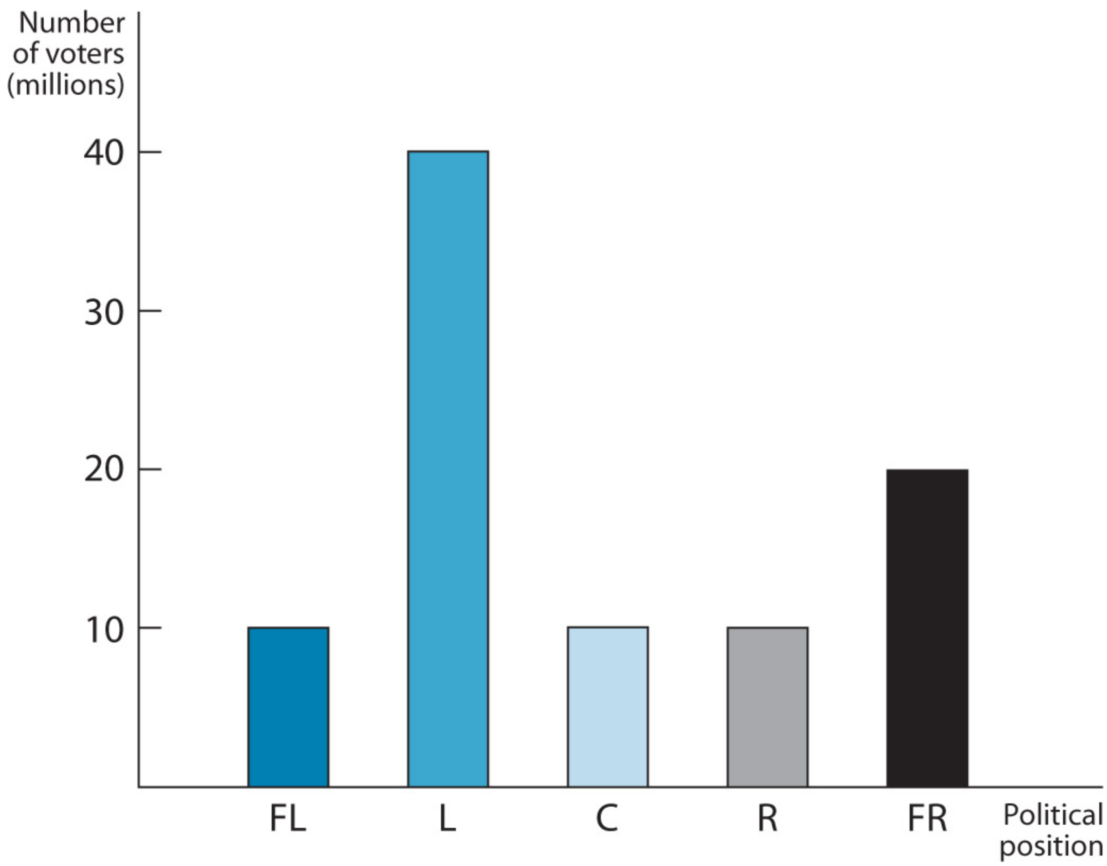


FIGURE 16.7 Discrete Distribution of Voters

In an election, each voter will vote for the candidate who publicly identifies herself as being closer to her own position on the spectrum. If both candidates are politically equidistant from a group of like-minded voters, each voter will flip a coin to decide which candidate to choose; this process will give each candidate one-half of the voters in that group.

Now suppose there is an upcoming presidential election between a former first lady (Claudia) and a former first lady hopeful (Dolores), each now running for office on her own.[25] Given the distribution of voters illustrated in Figure 16.7, we can construct a payoff table for the two candidates showing the number of votes that each can expect to receive under all of the possible combinations of political platform choices. This five-by-five table is shown in Figure 16.8, with totals denoted in millions of votes. The candidates will choose their location strategies to maximize the number of votes that they receive (and thus to increase their chances of winning).[26]

Here is how the votes are allocated. When both candidates choose the *same* position (the five cells along the top-left to bottom-right diagonal of the table), each candidate gets exactly one-half of the votes; because all voters are equidistant from each candidate, all of them flip coins to decide their choices, and each candidate garners 45 million votes. When the two candidates choose *different* positions, the more left-leaning candidate gets all the votes at or to the left of her position, while the more right-leaning candidate gets all the votes at or to the right of her position. In addition, each candidate gets the votes in central positions closer to her than to her rival, and the two of them split the votes from any voters in a position equidistant between them. Thus, if Claudia locates herself at L while Dolores locates herself at FR, Claudia gets the 40 million votes at L, the 10 million at FL, *and* the 10 million at C (because C is closer to L than to FR). Dolores gets the 20 million votes at FR and the 10 million at R (because R is closer to FR than to L). The payoffs are (60, 30). Similar calculations determine the outcomes in the rest of the table.

The table in Figure 16.8 is large, but the game can be solved very quickly. We begin with the now familiar search for dominant, or dominated, strategies for the two players. Immediately we see that for Claudia, FL is dominated by L and FR is dominated by R. For Dolores, too, FL is dominated by L and FR by R. With these extreme strategies eliminated, R is dominated by C for each candidate. With the two R strategies gone, C is dominated by L for each candidate.

The only remaining cell in the table is (L, L); this is the Nash equilibrium.

		DOLORES				
		FL	L	C	R	FR
CLAUDIA	FL	45, 45	10, 80	30, 60	50, 40	55, 35
	L	80, 10	45, 45	50, 40	55, 35	60, 30
	C	60, 30	40, 50	45, 45	60, 30	65, 25
	R	40, 50	35, 55	30, 60	45, 45	70, 20
	FR	35, 55	30, 60	25, 65	20, 70	45, 45
You may need to scroll left and right to see the full figure.						

FIGURE 16.8 Payoff Table for Candidates' Positioning Game

We now note three important characteristics of the equilibrium in the candidate-location game. First, both candidates locate at the *same* position in equilibrium. This outcome illustrates the principle of minimum differentiation, a general result in all two-player games of locational competition, whether it be political platform choices by presidential candidates, hot-dog-cart location choices by street vendors, or product feature choices by electronics manufacturing firms.[27] When the persons who vote for or buy from you can be arranged on a well-defined spectrum of preferences, you do best by looking as much like your rival as possible. This principle explains a diverse collection of behaviors on the part of political candidates and businesses. It may help you understand, for example, why there is never just one gas station at a heavily traveled intersection or why all brands of four-door sedans (or minivans, or sport utility vehicles) seem to look the same even though every brand claims to be coming out continually with a "new" look.

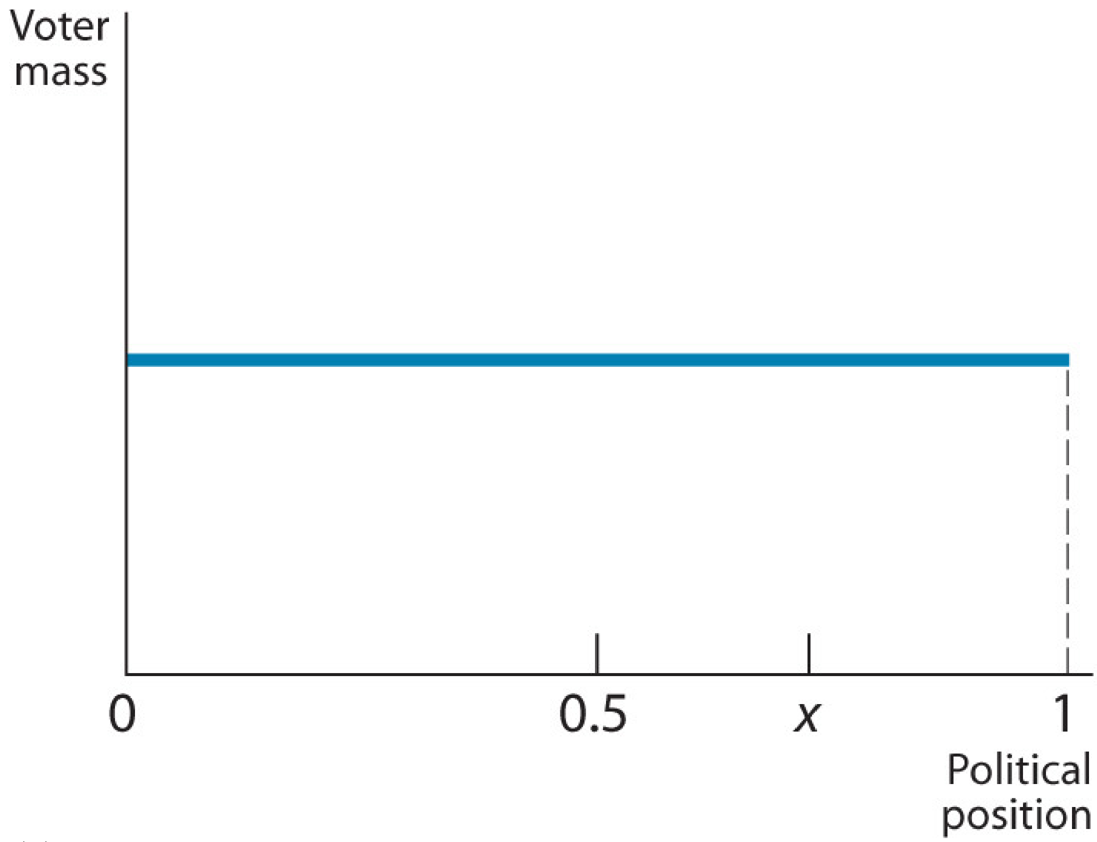
Second, and perhaps most crucially, both candidates locate at the position of the median voter in the population. In our example, with a total of 90 million voters, the median voter is number 45 million from each end. The numbers within one location can be assigned arbitrarily, but the location of the median voter is clear; here, the median voter is located at the L position on the political spectrum. So that is where both candidates locate themselves, which is the result predicted by the median voter theorem.

Third, observe that the location of the median voter need not coincide with the geometric center of the spectrum. The two will coincide if the distribution of voters is symmetric, but the median voter can be to the left of the geometric center if the distribution is skewed to the left (as is true in Figure 16.7) and to the right if the distribution is skewed to the right. This observation helps explain why state political candidates in Massachusetts, for example, *all* tend to be more liberal than candidates for similar positions in Texas or South Carolina.

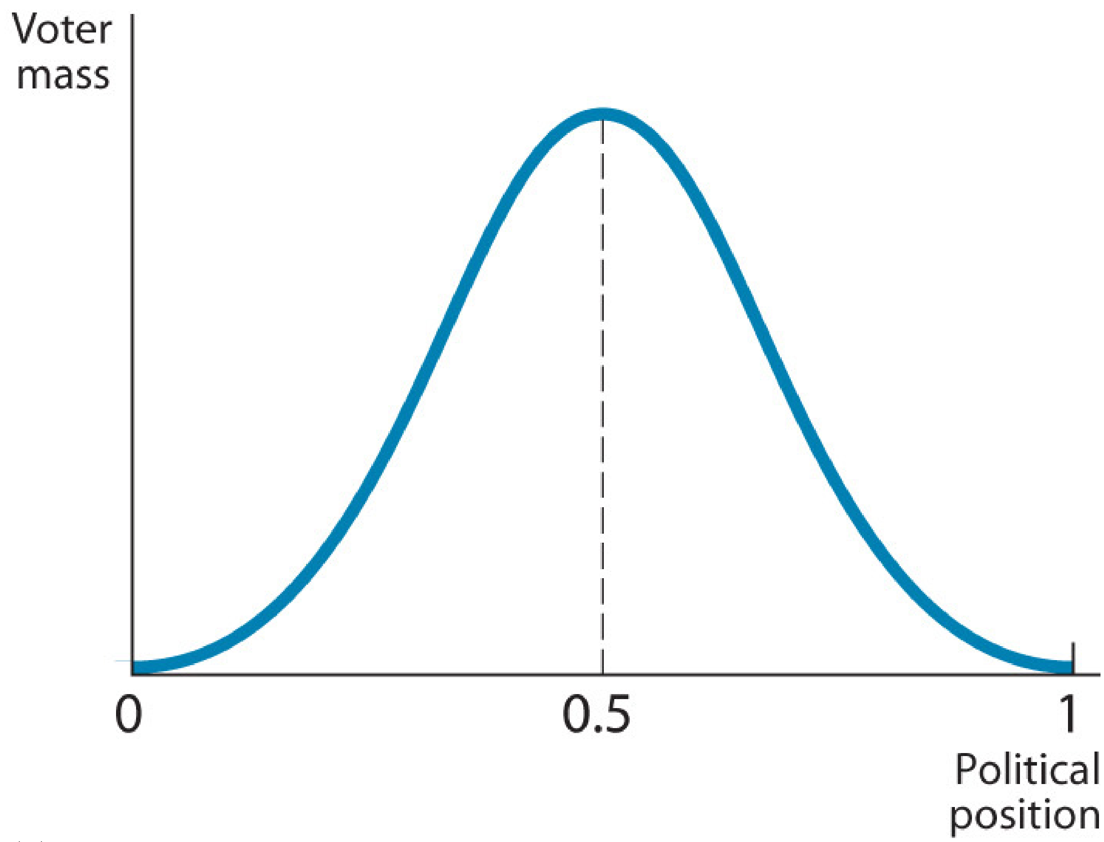
The median voter theorem can be expressed in different ways. One version states simply that the position of the median voter is the equilibrium position of the candidates in a two-candidate election. Another version says that the position that the median voter most prefers will be the Condorcet winner; this position will defeat every other position in a pairwise contest. For example, if M is this median position and L is any position to the left of M, then a candidate positioned at M will get all the votes of people who most prefer a position at or to the right of M, plus some to the left of M but closer to M than to L. Thus, M will get more than 50% of the votes. The two versions amount to the same thing because, in a two-candidate election, both candidates seeking to win a majority of votes will adopt the Condorcet-winner position. In addition, to guarantee that the result holds for a particular population of voters, the theorem (in either form) requires that each voter’s preferences be “reasonable.” *Reasonable* here means single peaked, as in Black’s condition, described in Section 3.A and Figure 16.4. Each voter has a unique, most preferred position on the political spectrum, and her payoff decreases away from that position in either direction.[28] In actual U.S. presidential elections, the theorem is borne out by the tendency for the candidates from the two major parties to make very similar promises to the electorate.

B. Continuous Political Spectrum

The median voter theorem can also be proved for a continuous distribution of political positions. Rather than having five, three, or any finite number of positions from which to choose, a continuous distribution assumes there are effectively an infinite number of political positions. These political positions are then associated with locations along the real number line between 0 and 1.[29] Voters are still distributed along the political spectrum as before, but because the distribution is now continuous rather than discrete, we use a voter distribution function rather than a histogram to illustrate voter locations. Two common distribution functions—the uniform distribution and the (symmetric) normal distribution—are illustrated in Figure 16.9.[30] The area under each curve represents the total number of votes available; for any given point along the interval from 0 to 1, such as x in Figure 16.9a, the number of votes up to that point is determined by finding the area under the distribution function from 0 to x. It should be clear that the median voter in each of these distributions is located at the center of the spectrum, at position 0.5.



(a) Uniform distribution



(b) Normal distribution

FIGURE 16.9Continuous Distribution of Voters

It is not feasible to construct a payoff table for our two candidates in the continuous-spectrum case; such tables must necessarily be finitely dimensioned and thus cannot accommodate an infinite number of possible strategies for players. We can, however, solve the game by applying the same strategic logic that we used for the discrete-spectrum case discussed in Section 5.A.

Consider the options of Claudia and Dolores as they contemplate the possible political positions open to them. Each knows that she must find her Nash equilibrium strategy—her best response to the equilibrium strategy of her rival. We can define a set of strategies that are best responses quite easily in this game, even though the complete set of possible strategies is impossible to delineate.

Suppose Dolores locates at a random position on the political spectrum, such as *x* in Figure 16.9a. Claudia can then calculate how the votes will be split for all possible positions that she might choose. If she chooses a position to the left of x, she gets all the votes to her left and half of the votes lying between her position and Dolores's. If she locates to the right of x, she gets all the votes to her right and half of the votes lying between her position and x. Finally, if she, too, locates at x, she and Dolores split the votes 50–50. These three possibilities effectively summarize all of Claudia's location choices, given that Dolores has chosen to locate at x.

But which of the response strategies just outlined is Claudia's *best* response? The answer depends on the location of x relative to the median voter. If x is to the right of the median, then Claudia knows that her best response will be to maximize the number of votes that she gains, which she can do by locating an infinitely small bit to the left of x.[31] In that case, she effectively gets all the votes from 0 to x, and Dolores gets those from x to 1. When x is to the right of the median, as in Figure 16.9a, then the number of voters represented by the area under the distribution curve from 0 to x is by definition larger than the number of voters from x to 1, so Claudia would win the election. Similarly, if x is to the left of the median, Claudia's best response will be to locate an infinitely small bit to the right of x and thus gain all the votes from x to 1. When x is exactly at the median, Claudia does best by also choosing to locate at x. The best-response strategies for Dolores are constructed exactly the same way and, given the location of her rival, are exactly the same as those described for Claudia. Our analysis shows that Claudia's best response to Dolores's location is to locate slightly closer to the median voter and, when Dolores locates at the position of the median voter, to locate in the same place. The same is true in reverse for Dolores's best response to Claudia's location. Therefore, the only equilibrium is when both locate at the median voter's position.

More complex mathematics is needed to prove the continuous version of the median voter theorem to the satisfaction of a true mathematician. For our purposes, however, the discussion here should convince you of the validity of the theorem in both its discrete and continuous forms. The most important limitation of the median voter theorem is that it applies when there is just one issue, or a one-dimensional spectrum of political differences. If there are two or more dimensions—for example, if being conservative versus liberal on social issues does not coincide with being conservative versus liberal on economic issues—then the voter population is spread out in a two-dimensional “issue space,” and the median voter theorem no longer holds. The preferences of every individual voter can still be single peaked, in the sense that the individual voter has a most preferred point and her payoff value drops away from this point in all directions, like the height going away from the peak of a hill. But we cannot identify a median voter in two dimensions, such that exactly the same number of voters have their most preferred point to one side of the median voter position as to the other side. In two dimensions, there is no unique sense of “side,” and the numbers of voters to the two sides can vary, depending on just how we define *side*.

Endnotes

- Any resemblance between our hypothetical candidates and actual past or possible future candidates in the United States is not meant to imply an analysis or prediction of their performances relative to the Nash equilibrium. Nor is our distribution of voters meant to typify U.S. voter preferences. Return to reference 25
- To keep the analysis simple, we ignore the complications created by the Electoral College and suppose that only the popular vote matters. Return to reference 26
- Economists learn this result within the context of Hotelling’s model of spatial location. See Harold Hotelling, “Stability in Competition,” *Economic Journal*, vol. 39, no. 1 (March 1929), pp. 41–57. Return to reference 27
- However, the distribution of voters’ positions along the political spectrum does not have to be single peaked, as indeed the histogram in Figure 15.7 is not—there are two peaks, at L and FR. Return to reference 28
- This construction is the same one used in Chapters 11 and 12 for analyzing large populations of individual members. Return to reference 29
- We do not delve deeply into the mechanics underlying distribution theory or the integral calculus required to calculate the exact proportion of the voting population lying to the left or right of any particular position on the continuous political spectrum. Here, we present only enough information to convince you that the median voter theorem continues to hold in the continuous case. Return to reference 30
- Such a location, infinitesimally removed from x to the left, is feasible in the continuous case. In our discrete example, candidates had to locate at exactly the same position. Return to reference 31

Glossary

median voter theorem
: If the political spectrum is one-dimensional and every voter has single-peaked preferences, then [1] the policy most preferred by the median voter will be the Condorcet winner, and [2] power-seeking politicians in a two-candidate election will choose platforms that converge to the position most preferred by the median voter. (This is also known as the **principle of minimum differentiation.**)

median voter
: The voter in the middle—at the 50th percentile—of a distribution.

discrete distribution
: A probability distribution in which the random variables may take on only a discrete set of values such as integers.

histogram
: A bar chart; data are illustrated by way of bars of a given height (or length).

principle of minimum differentiation
: Same as part [2] of the *median voter theorem*.

continuous distribution
: A probability distribution in which the random variables may take on a continuous range of values.

distribution function
: A function that indicates the probability that a variable takes on a value less than or equal to some number.

uniform distribution
: A common statistical distribution in which the *distribution function* is horizontal; data are distributed uniformly at each location along the range of possible values.

normal distribution
: A commonly used statistical distribution for which the *distribution function* looks like a bell-shaped curve.

SUMMARY

Elections can be held using a variety of different voting procedures that differ in the order in which issues are considered or the manner in which votes are tallied. Voting procedures are classified as *binary*, *plurative*, or *mixed methods*. Binary methods include *majority rule* as well as *pairwise* procedures such as the *Condorcet method* and the *amendment procedure*. *Positional methods* such as the *plurality rule* and the *Borda count*, as well as *approval voting*, are plurative methods. And *majority runoffs*, *instant runoffs*, and *proportional representation* are mixed methods.

Voting paradoxes (such as the *Condorcet*, the *agenda*, and the *reversal paradox*) show that counterintuitive election results can arise owing to difficulties associated with aggregating voter preferences or to small changes in the list of candidates or issues being considered. Outcomes in any given election under a given set of voter preferences can differ depending on the voting procedure used. Certain principles for evaluating voting methods can be described, although Arrow's *impossibility theorem* shows that no one system satisfies all of these criteria at the same time. Researchers in a broad range of fields have considered alternatives to the principles that Arrow identified.

Voters have scope for strategic behavior in the game that chooses the voting procedure or in an election itself through the *misrepresentation of their own preferences*. Voters may strategically misrepresent their preferences to achieve their most preferred or to avoid their least preferred outcome. In the presence of imperfect information, voters may decide whether to vote strategically on the basis of their beliefs about others' behavior and their knowledge of the distribution of preferences.

Candidates may also behave strategically in building a political platform. A general result known as the *median voter theorem* shows that in elections with only two candidates, both locate at the preference position of the *median voter*. This result holds when voters are distributed along the preference spectrum either *discretely* or *continuously*.

KEY TERMS

agenda paradox (635)

amendment procedure (629)

antiplurality method (630)

approval voting (630)

binary method (628)

Black's condition (639)

Borda count (630)

Condorcet method (629)

Condorcet paradox (633)

Condorcet terms (641)

Condorcet winner (629)

continuous distribution (655)

Copeland index (629)

discrete distribution (652)

distribution function (656)

Gibbard‑Satterthwaite theorem (651)

histogram (652)

impossibility theorem (638)

Glossary

binary method
: A class of voting methods in which voters choose between only two alternatives at a time.

majority rule
: A voting method in which the winning alternative is the one that garners a majority (more than 50%) of the votes.

pairwise voting
: A voting method in which only two alternatives are considered at the same time.

multistage procedure
: A voting procedure in which there are multiple rounds of voting.

Condorcet method
: A voting method in which the winning alternative must beat each of the other alternatives in a round-robin of pairwise contests.

Condorcet winner
: The alternative that wins an election run using the *Condorcet method*.

Copeland index
: An index measuring an alternative's record in a round-robin of contests where different numbers of points are allocated for wins, ties, and losses.

amendment procedure
: A procedure in which any amended version of a proposal must win a vote against the original version before the winning version is put to a vote against the status quo.

plurative method
: Any voting method that allows voters to consider a slate of three or more alternatives simultaneously.

positional method
: A voting method that determines the identity of the winning alternative using information on the position of

alternatives on a voter's ballot to assign points used when tallying ballots.

plurality rule
: A voting method in which two or more alternatives are considered simultaneously and the winning alternative is the one that garners the largest number of votes; the winner needs only gain more votes than each of the other alternatives and does not need 50% of the vote as would be true in *majority rule*.

antiplurality method
: A positional voting method in which the electorate is asked to vote against one item on the slate (or to vote for all but one).

Borda count
: A positional voting method in which the electorate indicates its order of preference over a slate of alternatives. The winning alternative is determined by allocating points based on an alternative's position on each ballot.

approval voting
: A voting method in which voters cast votes for all alternatives of which they approve.

mixed method
: A multistage voting method that uses plurative and binary votes in different rounds.

majority runoff
: A two-stage voting method in which a second round of voting ensues if no alternative receives a majority in the first round. The top two vote-getters are paired in the second round of voting to determine a winner.

round
: A single vote within a larger *multistage procedure* that consists of multiple sequentially held votes.

single transferable vote
: A voting method in which each voter indicates her preference ordering over all candidates on a single initial ballot. If no alternative receives a majority of

all first-place votes, the bottom-ranked alternative is eliminated and all first-place votes for that candidate are "transferred" to the candidate listed second on those ballots; this process continues until a majority winner emerges. Also called **instant-runoff voting (IRV)** or **ranked-choice voting**.

instant-runoff voting (IRV)

Same as **single transferable vote**.

ranked-choice voting

Another name for **single transferable vote**.

proportional representation

This voting system requires that the number of seats in a legislature be allocated in proportion to each party's share of the popular vote.

strategic voting

Voting in conformity with your optimal rational strategy found by doing rollback analysis on the game tree of the voting procedure.

Condorcet paradox

Even if all individual voter preference orderings are transitive, there is no guarantee that the social preference ordering generated by Condorcet's voting method will also be transitive.

social ranking

The preference ordering of a group of voters that arises from aggregating the preferences of each member of the group.

intransitive ordering

A preference ordering that cycles and is not *transitive*. For example, a preference ordering over three alternatives A, B, and C is intransitive if A is preferred to B and B is preferred to C but it is not true that A is preferred to C.

transitive ordering

A preference ordering for which it is true that if option A is preferred to B and B is preferred to C, then A is also preferred to C.

agenda paradox
: A voting situation where the order in which alternatives are paired when voting in multiple rounds determines the final outcome.

sincere voting
: Voting at each point for the alternative that you like best among the ones available at that point, regardless of the eventual outcome. Also called **truthful voting**.

truthful voting
: Same as **sincere voting**.

reversal paradox
: This paradox arises in an election with at least four alternatives when one of these is removed from consideration after votes have been submitted and the removal changes the identity of the winning alternative.

impossibility theorem
: A theorem that indicates that no preference aggregation method can satisfy the six critical principles identified by Kenneth Arrow.

single-peaked preferences
: A preference ordering in which alternatives under consideration can be ordered along some specific dimension and each voter has a single ideal or most-preferred alternative with alternatives farther away from the most-preferred point providing steadily lower payoffs. Also called **Black's condition**.

Black's condition
: Same as the condition of **single-peaked preferences**.

robustness
: A measure of the number of sets of voter preference orderings that are nondictatorial, satisfy independence of irrelevant alternatives and the Pareto property, and also produce a transitive *social ranking*.

Condorcet terms
: A set of ballots that would generate the Condorcet paradox and that should together logically produce a tied vote among three possible alternatives. In a three-

candidate election among A, B, and C, the Condorcet terms are three ballots that show A preferred to B preferred to C; B preferred to C preferred to A; C preferred to A preferred to B.

reversal terms
: A set of ballots that would generate the *reversal paradox* and that should together logically produce a tied vote between a pair of alternatives. In a three-candidate election among A, B, and C, the reversal terms are two ballots that show a reversal in the location of a pair of alternatives. For example, one ballot with A preferred to B preferred to C and another with B preferred to A preferred to C should produce a tie between A and B.

spoiler
: Refers to a third candidate who enters a two-candidate race and reduces the chances that the leading candidate actually wins the election.

Gibbard–Satterthwaite theorem
: With three or more alternatives to consider, the only voting method that prevents strategic voting is dictatorship; one person is identified as the dictator and her preferences determine the outcome.

median voter theorem
: If the political spectrum is one-dimensional and every voter has single-peaked preferences, then [1] the policy most preferred by the median voter will be the Condorcet winner, and [2] power-seeking politicians in a two-candidate election will choose platforms that converge to the position most preferred by the median voter. (This is also known as the **principle of minimum differentiation.**)

median voter
: The voter in the middle—at the 50th percentile—of a distribution.

discrete distribution
: A probability distribution in which the random variables may take on only a discrete set of values such as

integers.

histogram

A bar chart; data are illustrated by way of bars of a given height (or length).

principle of minimum differentiation

Same as part [2] of the *median voter theorem*.

continuous distribution

A probability distribution in which the random variables may take on a continuous range of values.

distribution function

A function that indicates the probability that a variable takes on a value less than or equal to some number.

uniform distribution

A common statistical distribution in which the *distribution function* is horizontal; data are distributed uniformly at each location along the range of possible values.

normal distribution

A commonly used statistical distribution for which the *distribution function* looks like a bell-shaped curve.

strategic misrepresentation of preferences

Refers to strategic behavior of voters when they use rollback to determine that they can achieve a better outcome for themselves by not voting strictly according to their preference orderings.

SOLVED EXERCISES

1. Consider a vote being taken by three roommates, A, B, and C, who share a triple dorm room. They are trying to decide which of three elective courses to take together this term. (Each roommate has a different major and is taking required courses in her major for the rest of her courses.) Their choices are Philosophy, Geology, and Sociology, and their preferences for the three courses are as shown here:

A	B	C
Philosophy	Sociology	Geology
Geology	Philosophy	Sociology
Sociology	Geology	Philosophy

The roommates have decided to have a two-round vote and will draw straws to determine who sets the agenda. Suppose A wins the right to set the agenda and wants the Philosophy course to be chosen. How should she set the agenda to achieve this outcome if she knows that everyone will vote truthfully in all rounds? What agenda should she set if she knows that they will all vote strategically?

2. Suppose that voters 1 through 4 are being asked to consider three different candidates—A, B, and C—in a positional (Borda-count) election. Their preference orderings are as shown here:

1	2	3	4
A	A	B	C
B	B	C	B
C	C	A	A

Assume that the voters will cast their votes truthfully (no strategic voting). Find a variation on the Borda system—a number of points to be allotted to a voter’s first, second, and third preferences—in which candidate A wins.

3. In 2015, Democrat Ethan Strimling (S) was elected mayor of Portland, Maine, in a three-way race using instant runoff voting, over incumbent mayor Michael Brennan (B, also a Democrat) and Green Independent Party leader Tom MacMillan (M). Suppose instead that this election is run

under the plurality rule and that there are three types of voters: Green Party supporters, who are 45% of all voters and whose preference ranking is M ≻ S ≻ B; Brennan Democrats, who are 15% of all voters and whose preference ranking is B ≻ S ≻ M; and Strimling Democrats, who are 40% of all voters and whose preference ranking is S ≻ B ≻ M. For simplicity, suppose that the voters in each of these three groups can coordinate their actions and vote together as a bloc.

1. Suppose that M voters are nonstrategic, always casting their votes for MacMillan, but that B and S voters are strategic. Show that the game B and S voters play has ordinal payoffs identical to those in the battle-of-the-sexes game shown in Figure 4.15. Explain why, in the mixed-strategy equilibrium of this game, all three candidates have a chance of winning the election.
2. Suppose now that B voters are nonstrategic, always casting their votes for Brennan, but that M and S voters are strategic. Construct the ordinal payoff matrix for the game played by M and S voters. Show that, in any Nash equilibrium of this game, Brennan always wins the election.

4. Nashville has been the capital of Tennessee since 1826, but before that, the capital had moved from Knoxville to Kingston and then to Murfreesboro. Might the capital move again someday? To explore the likelihood of that happening, consider a voting game played by residents of Tennessee's four biggest cites: Nashville (N), population 660,000; Memphis (M), population 650,000; Knoxville (K), population 190,000; and Chattanooga (C), population 180,000. The figure below shows the location of each city within the state. Residents of these four cities vote to choose the state's new capital.

Residents prefer the capital to be as close to them as possible. Nashville residents have the preference ordering N ≻ C ≻ K ≻ M, Memphis residents have the preference ordering M ≻ N ≻ C ≻ K, Chattanooga residents have the preference ordering C ≻ K ≻ N ≻ M, and Knoxville residents have the preference ordering K ≻ C ≻ N ≻ M.

1. Would there a Condorcet winner in this voting game?
2. Which city would win under instant runoff voting (IRV), assuming nonstrategic voting?

3. What if Nashville were smaller and Chattanooga were bigger? In particular, suppose that Nashville shrunk to Chattanooga's current size (population 180,000) and Chattanooga grew to Nashville's current size (population 660,000). Would there be a Condorcet winner under these circumstances? Which city would win under IRV, again assuming nonstrategic voting?

5. Consider a group of 50 residents attending a town meeting in Massachusetts. They must choose one of three proposals for dealing with town garbage. Proposal 1 asks the town to provide garbage collection as one of its services; Proposal 2 calls for the town to hire a private garbage collector to provide collection services; and Proposal 3 calls for residents to be responsible for their own garbage. There are three types of voters. The first type prefers Proposal 1 to Proposal 2 and Proposal 2 to Proposal 3; there are 20 of these voters. The second type prefers Proposal 2 to Proposal 3 and Proposal 3 to Proposal 1; there are 15 of these voters. The third type prefers Proposal 3 to Proposal 1 and Proposal 1 to Proposal 2; there are 15 of them.
   1. Under the plurality rule, which proposal wins?
   2. Suppose voting proceeds with the use of a Borda count in which voters list the proposals, in order of preference, on their ballots. The proposal listed first (or at the top) on a ballot gets three points; the proposal listed second gets two points; and the proposal listed last gets one point. In this situation, with no strategic voting, how many points does each proposal gain? Which proposal wins?
   3. What strategy can the second and third types of voters use to alter the outcome of the Borda count vote in part (b) to one that both types prefer? If they use this strategy, how many points does each proposal get, and which one wins?
6. During the Cuban missile crisis (described more fully in Chapter 13), serious differences of opinion arose within ExComm, the group advising President John F. Kennedy, which we summarize here. There were three policy options: Soft (a blockade of Cuba), Medium (a limited air strike), and Hard (a massive air strike or invasion of Cuba). There were also three groups in ExComm. The civilian doves ranked the alternative Soft best, Medium next, and Hard last. The civilian hawks ranked Medium best, Hard next, and Soft last. The military ranked Hard best, but they felt "so strongly about the dangers inherent in the limited strike that they would prefer taking no military action rather than to take that limited strike."[32] In other words, they ranked Soft second and Medium last. Each group constituted about one-third of ExComm, and so any two of the groups would form a majority.
   1. If the policy were to be decided by a majority vote in ExComm and the members voted sincerely, which alternative, if any, would win?

2. What outcome would arise if members voted strategically? What outcome would arise if one group had agenda-setting power? (Model your discussion in these two cases on the analyses found in Sections 2.B and 4.B.)

7. In his book *A Mathematician Reads the Newspaper*, John Allen Paulos gives the following caricature based on the 1992 Democratic presidential primary caucuses.[33] There are five candidates: Jerry Brown, Bill Clinton, Tom Harkin, Bob Kerrey, and Paul Tsongas. There are 55 voters, with different preference orderings for these candidates. There are six different orderings, which we label I through VI. The preference orderings (1 for best to 5 for worst), along with the numbers of voters with each ordering, are shown in the following table (the candidates are identified by the first letters of their last names):

		GROUPS AND THEIR SIZES					
		I 18	II 12	III 10	IV 9	V 4	VI 2
RANKING	1	T	C	B	K	H	H
	2	K	H	C	B	C	B
	3	H	K	H	H	K	K
	4	B	B	K	C	B	C
	5	C	T	T	T	T	T

You may need to scroll left and right to see the full figure.

---

1. First, suppose that all voters vote sincerely. Consider the outcomes of each of several different election rules. Show each of the following outcomes:
    1. Under the plurality method (the one with the most first preferences), Tsongas wins.
    2. Under the runoff method (the top two first preferences go into a second round), Clinton wins.
    3. Under the elimination method (at each round, the one with the fewest first preferences in that round is eliminated, and the rest go into the next round), Brown wins.
    4. Under the Borda count method (5 points for first preference, 4 for second, and so on; the candidate with the most points wins), Kerrey wins.
    5. Under the Condorcet method (pairwise comparisons), Harkin wins.
2. Suppose that you are a Brown, Kerrey, or Harkin supporter. Under the plurality method, you would get your worst outcome. Can you benefit by voting strategically? If so, how?

3. Are there opportunities for strategic voting under each of the other methods as well? If so, explain who benefits from voting strategically and how they can do so.

8. As mentioned in the chapter, some localities have replaced runoff elections and even primaries with instant runoff voting to save time and money. Most jurisdictions, however, still use a two-stage system in which, if a candidate fails to receive a majority of votes in the first round, a second runoff election is held weeks later between the two candidates who earned the most votes.

   For instance, France employs a two-stage system for its presidential elections. No primaries are held. Instead, all candidates from all parties are on the ballot in the first round, which usually guarantees a second round, since it is very difficult for a single candidate to earn a majority of votes among such a large field. Although a runoff in the French presidential election is always expected, it doesn't mean that French elections are without the occasional surprise, as we saw in Section 1.C.

   Instant runoff voting can be explained in five steps:

   1. Voters rank all candidates according to their preferences.
   2. The votes are counted.
   3. If a candidate has earned a majority of the votes, that candidate is the winner. If not, go to step 4.
   4. Eliminate candidate(s) with the fewest votes. (Eliminate more than one candidate at the same time only if they tie for the fewest votes.)
   5. Redistribute the votes for eliminated candidates to the next-ranked choices on those ballots. Once this is done, return to step 2.

   1. Instant runoff voting is slowly gaining traction. Given its potential for saving money and time, it might be surprising that the institution isn't more widely adopted. Why might some oppose instant runoff voting? (Hint: Which candidates, parties, and interests benefit from the two-stage systems that are currently in place?)
   2. What other concerns or criticisms might be raised about instant runoff voting?

9. An election has a slate of three candidates and takes place under the plurality rule. There are numerous voters, spread along a political spectrum from left to right. Represent this spread by a horizontal line whose extreme points are 0 (left) and 1 (right). Voters are uniformly distributed along this spectrum, so the number of voters in any segment of the line is proportional to the length of that segment. Thus, a third of the voters are in the segment from 0 to $\frac{1}{3}$, a quarter in the

segment from ½ to ¾, and so on. Each voter votes for the candidate whose declared political position is closest to the voter's own position. The candidates have no ideological attachment and take up any position along the line, each seeking only to maximize her share of votes.

1. Suppose you are one of the three candidates. The leftmost of the other two positions herself at point x, and the rightmost at point $(1 - y)$, where $x + y < 1$ (so that the rightmost candidate is a distance y from 1). Show that your best response is to take up the following positions under the given conditions:
    1. Just slightly to the left of x if $x > y$ and $3x + y > 1$
    2. Just slightly to the right of $(1 - y)$ if $y > x$ and $x + 3y > 1$
    3. Exactly halfway between the other candidates if $3x + y < 1$ and $x + 3y < 1$
2. In a graph with x and y along the axes, show the areas (the combinations of x and y values) where each of the response rules (i) - (iii) in part (a) is best for you.
3. From your analysis, what can you conclude about the Nash equilibrium of the game where the three candidates each choose positions?

UNSOLVED EXERCISES

1. Repeat Exercise S1 for the situation in which B sets the agenda and wants to ensure that Sociology wins.
2. Repeat Exercise S2 to find a variation on the Borda weighting system in which candidate B wins.
3. In the small town of Embarrass, Minnesota, famous for its funny name, the Miss Embarrass contest is no laughing matter. Each August, several young ladies from around the Embarrass River Valley vie for the honor of being crowned Queen at the Embarrass Region Fair. (Everything so far is true; the rest is made up.) The Miss Embarrass contest has traditionally used the plurality method to determine the winner. However, one of the organizers recently read about a FairVote.org study that found that instant runoff voting (IRV) promotes more civil campaigning, and that format is now being considered.

   All 1,000 Embarrass area residents who attend the fair are eligible to vote. This year, Ashley (A) and Becky (B) are the front-runners, trailed by Olivia (O). At the start of fair season, 420 people are "Ashley fans," with the preference ordering $A \succ B \succ O$; 400 people are "Becky fans," with the preference ordering $B \succ A \succ O$; and 180 are "Olivia fans," with the preference ordering $O \succ A = B$. However, fairgoers' preferences may change before the vote, depending on whether Ashley and/or Becky engages in "negative campaigning" by spreading a nasty rumor during the fair about her opponent. (For simplicity, we assume that Olivia is sure to run a positive campaign.)

   As you consider the questions below, assume (i) that Ashley and Becky each prefer to remain civil and each prefer not to have a nasty rumor spread about them, but most of all want to win the election; (ii) that negative campaigning suppresses turnout, causing 40 fans of the targeted candidate to forgo voting in the election; and (iii) that if only one candidate stays civil, all Olivia fans will prefer the civil candidate as their second favorite. (If Ashley and Becky are both civil or both negative, Olivia fans will flip a coin if forced to decide between them.)

   1. Suppose that the plurality rule continues to be used for the election. Draw the ordinal payoff matrix for the "civility game" played by Ashley and Becky when they choose between Civil and Negative. Show that this game does not have a pure-strategy Nash equilibrium. Explain in words why, in the unique mixed-strategy Nash equilibrium, Ashley and Becky are each sometimes crowned Miss Embarrass.

2. Suppose that the organizers decide to switch to instant runoff voting for the election. Show that being civil is now a superdominant strategy for both Ashley and Becky. Hint: Draw the ordinal payoff matrix for the civility game with Ashley as the row player. Civil is superdominant for Ashley if the best two payoffs (4 and 3) appear in the Civil row. Similarly, Civil is superdominant for Becky if the two best payoffs appear in the Civil column.

4. Consider a version of recent political developments in Britain related to the issue of leaving the European Union (Brexit), only slightly caricatured here. There are three alternatives: Remain in the EU (labeled R), a negotiated Soft Brexit (S), and crashing out of the EU without a negotiated agreement in a Hard Brexit (H). There are three types of voters: 45% of voters are Remainers, with the preference ranking $R \succ S \succ H$; 25% of voters are Moderate leavers, with the preference ranking $S \succ H \succ R$; and 30% of voters are Extremist leavers, with the preference ranking $H \succ R \succ S$ (because they regard a negotiated Soft Brexit as the worst compromise and irreversible, whereas with R they hope to continue their fight and get H in the future).

   For each of the possible voting methods described below, calculate the outcome, first assuming that all types of voters vote sincerely, and then assuming that they all vote strategically (which requires collective action for each group).

   1. Plurality rule: All alternatives are on the ballot at the same time, and the one with the most votes wins.
   2. Two-round ballot: All alternatives are on the first ballot. Each voter votes for just one. If one gets over 50% of the votes, it wins; if not, a second round of voting pits the top two vote getters against each other. The option with the most votes wins.
   3. Exhaustive ballot: All alternatives are on the first ballot. Each voter votes for just one option. If one gets over 50% of the votes, it wins; if not, the one with the fewest votes is eliminated, and another vote is run among all the remaining options. This process continues until an option gains 50% of the votes in a round.
   4. Ranked choice (instant runoff): All alternatives are on the ballot, and each voter ranks them all from best (1) to worst (3). If one option gets more than 50% first-choice votes, it wins. If not, the one with the fewest first-choice votes is eliminated, and all first-choice votes for that alternative are transferred to the alternative ranked second on those ballots. The one that has the majority in this recalculation is the winner.
   5. Borda count: All alternatives are on the ballot, and each voter ranks them all from best (1) to worst (3). All these numbers are

added over all the voters, and the alternative that gets the smallest total is the winner.

6. Pairwise voting: A ballot for each pair of alternatives is prepared and voters vote on each pairing (R versus S, S versus H, and H versus R). The alternative that wins all of its contests is the winner.
7. Sequential decisions: The first round of voting is Remain versus Leave. If Remain wins, that ends the process. If Leave wins, there a second round of voting between Soft and Hard, and the winning alternative is implemented.

5. Every year, college football' s Heisman Trophy is awarded by means of a Borda count procedure. Each voter submits first-, second-, and third-place votes, worth 3 points, 2 points, and 1 point, respectively. Thus, the Borda count point scheme used may be called (3-2-1), where the first digit is the point value of a first-place vote, the second digit the point value of a second-place vote, and the third digit the point value of a third-place vote. In 2004, the vote totals for the top five candidates under the Borda system were as follows:

Player	1st Place	2st Place	3st Place
Leinhart (USC)	267	211	102
Peterson (Oklahoma)	154	180	175
White (Oklahoma)	171	149	146
Smith (Utah)	98	112	117
Bush (USC)	118	80	83

---

1. Compare the Borda count scores of Leinhart and Peterson. By what margin of Borda points did Leinhart win?
2. It seems only fair that a point scheme should give a first-place vote at least as much weight as a second-place vote and a second-place vote at least as much weight as a third-place vote. That is, for a point scheme (x-y-z), we should have $x \geqslant y \geqslant z$. Given this "fairness" restriction, is there any point scheme under which Leinhart would have lost? If so, provide such a scheme. If not, explain why not.
3. Even though White had more first-place votes than Peterson, Peterson had a higher Borda count total. If first-place votes were weighted enough, White' s edge in first-place votes could have given him a higher Borda count. Assume that second-place votes are worth 2 points and third-place votes are worth 1 point, so that the point scheme is (x−2-1). What is the lowest integer value of x such that White would get a higher Borda count than Peterson?

    4. Suppose that the data in the table represent truthful voting. For simplicity, let’s suppose that the election was a simple plurality vote instead of a Borda count. Note that Leinhart and Bush are both from USC, whereas Peterson and White are both from Oklahoma. Suppose that, due to Oklahoma loyalty, those voters who prefer White all have Peterson as their second choice. If those voters were to vote strategically in a plurality-rule election, could they change the outcome of the election? Explain.
    5. Similarly, suppose that due to USC loyalty, those voters who prefer Bush all have Leinhart as their second choice. If all four voting groups (those who most prefer Leinhart, then Bush; those who most prefer Peterson, then White; those who most prefer White, then Peterson; and those who most prefer Bush, then Leinhart) were to vote strategically in a plurality-rule election, who would be the winner of the Heisman Trophy?
    6. In 2004, there were 923 Heisman voters. Under the actual (3-2-1) system, what is the minimum integer number of first-place votes that it would have taken to guarantee victory (that is, without the help of any second-or third-place votes)? Note that a player’s name may appear on a ballot only once.

6. Olympic figure skaters complete two programs in their competition, one short and one long. In each program, the skaters are scored and then ranked by a panel of nine judges. The skaters’ positions in the rankings are used to determine their final scores. A skater’s ranking depends on the number of judges placing her first (or second or third); the skater judged to be best by the most judges is ranked 1, and so on. In the calculation of a skater’s final score, the short program gets half the weight of the long program. That is, Final score = 0.5(Rank in short program) + Rank in long program. The skater with the lowest final score wins the gold medal. In the event of a tie, the skater judged best in the long program by the most judges takes the gold.

   In the women’s individual figure-skating competition at the 2002 Olympics in Salt Lake City, Michelle Kwan was in first place after the short program. She was followed by Irina Slutskaya, Sasha Cohen, and Sarah Hughes, who were in second, third, and fourth places, respectively. In the long program, the judges’ cards for these four skaters were as follows:

		JUDGE NUMBER								
		1	2	3	4	5	6	7	8	9
KWAN	Points	11.3	11.5	11.7	11.5	11.4	11.5	11.4	11.5	11.4

You may need to scroll left and right to see the full figure.

		JUDGE NUMBER								
		1	2	3	4	5	6	7	8	9
	Rank	2	3	2	2	2	3	3	2	3
SLUTSKAYA	Points	11.3	11.7	11.8	11.6	11.4	11.7	11.5	11.4	11.5
	Rank	3	1	1	1	4	1	2	3	2
COHEN	Points	11.0	11.6	11.5	11.4	11.4	11.4	11.3	11.3	11.3
	Rank	4	2	4	3	3	4	4	4	4
HUGHES	Points	11.4	11.5	11.6	11.4	11.6	11.6	11.3	11.6	11.6
	Rank	1	4	3	4	1	2	1	1	1
You may need to scroll left and right to see the full figure.										

---

1. In the long program, Slutskaya skated last of the top skaters. Use the information from the judges' cards to determine the judges' long-program ranks for Kwan, Cohen, and Hughes *before* Slutskaya skated. Then, using the standings already given for the short program in conjunction with your calculated ranks for the long program, determine the final scores and standings among these three skaters *before* Slutskaya skated. (Note that Kwan's rank in the short program was 1, so her partial score after the short program is 0.5.)
2. Given your answer to part (a), what would have been the final outcome of the competition if the judges had ranked Slutskaya's long program above all three of the others?
3. Use the judges' cards to determine the actual final scores for all four skaters *after* Slutskaya skated. Who won each medal?
4. Which of Arrow's principles does the Olympic figure-skating scoring system violate? Explain.

7. The 2008 presidential nomination season saw 21 Republican primaries and caucuses on Super Tuesday—February 5, 2008. By that day—only a month after the Iowa caucuses that began the process—more than half of the Republican contenders had dropped out of the race, leaving only four: John McCain, Mitt Romney, Mike Huckabee, and Ron Paul. McCain, Romney, and Huckabee had each previously won at least one state. McCain had beaten Romney in Florida the week before Super Tuesday, and at that point it looked like only the two of them stood a realistic chance of wining the nomination. In this primary season, as is typical for the Republican party, nearly every GOP contest (whether primary or caucus) was winner-take-all, so winning a given state would earn a candidate all of the convention delegates allotted to that state by the Republican National Committee.

The West Virginia caucus was the first contest to reach a conclusion on Super Tuesday because that caucus took place in the afternoon, it was brief, and the state is in the eastern time zone. News of the result was available hours before the close of polls in many of the other states voting that day.

The following problem is based on the results of that West Virginia caucus. As we might expect, the caucusers did not all share the same preference orderings of the candidates. Some favored McCain, whereas others liked Romney or Huckabee. The caucusers also had varied preferences about whom they wanted to win if their favorite candidate did not. Simplifying substantially from reality (but based on the actual voting), assume that there were seven types of West Virginia caucusers that day, whose prevalence and preferences were as follows:

	I (16%)	II (28%)	III (13%)	IV (21%)	V (12%)	VI (6%)	VII (4%)
1st	McCain	Romney	Romney	Huckabee	Huckabee	Paul	Paul
2nd	Romney	McCain	Huckabee	Romney	McCain	Romney	Huckabee
3rd	Huckabee	Huckabee	McCain	McCain	Romney	Huckabee	Romney
4th	Paul	Paul	Paul	Paul	Paul	McCain	McCain
You may need to scroll left and right to see the full figure.							

At first, no one knew the distribution of preferences of those in attendance at the caucus, so everyone voted truthfully. Thus, Romney won a plurality of the votes in the first round, with 41%.

After each round of this caucus, if no candidate wins a majority, the candidate with the smallest number of votes is dropped from consideration, and his or her supporters vote for one of the remaining candidates in the following rounds.

1. What would the results of the second round have been under truthful (nonstrategic) voting for the remaining three candidates?
2. If West Virginia had held pairwise votes among the four candidates, which one would have been the Condorcet winner with truthful voting?
3. In reality, the results of the second round of the caucus were as follows:

   Huckabee 52%

   Romney 47%

McCain 1%

Given the preferences of the McCain voters, why might this have happened? (Hint: How would the outcome have been different if West Virginia had voted last on Super Tuesday?)

4. After the fact, Romney's campaign cried foul and accused the McCain and Huckabee supporters of making a backroom deal.[34] Should Romney's campaign have suspected collusion between the McCain and Huckabee camps in this case? Explain why or why not.

8. Return to the discussion of instant runoff voting (IRV) in Exercise S8.
   1. Consider the following IRV ballots of five voters:

	Ana	Bernard	Cindy	Desmond	Elizabeth
1st	Jack	Jack	Kate	Locke	Locke
2nd	Kate	Kate	Locke	Kate	Jack
3rd	Locke	Locke	Jack	Jack	Kate
You may need to scroll left and right to see the full figure.					

Which—if any—of the five voters have an incentive to vote strategically? If so, who and why? If not, explain why not.

   2. Consider the following table, which gives the IRV ballots of a small town of seven citizens voting on five policy proposals put forward by the mayor:

	Anderson	Brown	Clark	Davis	Evans	Foster	García
1st	V	V	W	W	X	Y	Z
2nd	W	X	V	X	Y	X	Y
3rd	X	W	Y	V	Z	Z	X
4th	Y	Y	X	Y	V	W	W
5th	Z	Z	Z	Z	W	V	V
You may need to scroll left and right to see the full figure.							

Assuming that all candidates (or policies) that tie for the fewest votes are eliminated at the same time, under what conditions is an eventual majority winner guaranteed? Put another way, under what conditions might there not be an unambiguous majority winner? (Hint: How important is it for Evans, Foster, and García to fill out their ballots completely?) How will these conditions change if a new citizen, Harris, moves into town and votes?

9. Recall the three-member council considering three alternative welfare policies in Section 4.C. There, three councillors (Left, Center, and Right) considered policies A and D in a first-round vote, with the winner facing policy G in a second-round election. But no one knows for sure exactly how many councillors have each set of possible preferences. The possible preference orderings are shown in Figure 16.1. Each councillor knows her own type, and she knows the probabilities of observing each type of voter, p_L, p_C, and p_R (with $p_L + p_C + p_R = 1$). The behavior of the Center-type voters in the first-round election is the only unknown in this situation and will depend on the probabilities that the various preference types occur. Suppose here that a Center-type voter believes (in contrast to the case considered in the chapter) that other Center types will vote strategically; suppose further that the Center type's payoffs are as in Section 4.C: 1 if A wins, 0 if G wins, and $0 < u < 1$ if D wins.
    1. Under what configuration of the other two votes does the Center-type voter's first-round vote matter to the outcome of the election? Given her assumption about the behavior of other Center-type voters, how would she identify the source of the first-round votes?
    2. Following the analysis in Section 4.C, determine the expected payoff to the Center type when she votes truthfully. Compare this with her expected payoff when she votes strategically. What is the condition under which the Center type votes strategically?

Endnotes

- Ernest R. May and Philip D. Zelikow, eds., *The Kennedy Tapes: Inside the White House during the Cuban Missile Crisis* (Cambridge, Mass.: Harvard University Press, 1997), p. 97. Return to reference 32
- John Allen Paulos, *A Mathematician Reads the Newspaper* (New York: Basic Books, 1995), pp. 104–6. Return to reference 33
- See Susan Davis, "Romney Cries Foul in W. Va. Loss," *Wall Street Journal*, February 5, 2008, available at http://blogs.wsj.com/washwire/2008/02/05/huckabee-wins-first-super-tuesday-contest/?mod=WSJBlog. Return to reference 34

17 ■ Bargaining

PEOPLE ENGAGE IN BARGAINING throughout their lives. Children start by negotiating to share toys and to play games with other children. Couples bargain about matters of housing, child rearing, and the adjustments that each must make for the other's career. Buyers and sellers bargain over price, workers and bosses over wages. Countries bargain over policies of mutual trade liberalization; superpowers negotiate mutual arms reduction. And the authors of this book had to bargain among themselves—generally very amicably—about what to include or exclude, how to structure the exposition, and so forth. To get a good result from such bargaining, the participants must devise good strategies. In this chapter, we explain some of the basic ideas of bargaining and some strategies derived from them.

All bargaining situations have two things in common. First, the total payoff that the parties to the negotiation are capable of creating and enjoying as a result of reaching an agreement should be greater than the sum of the individual payoffs that they could achieve separately—the whole must be greater than the sum of the parts. Without the possibility of this excess value, or *surplus*, the negotiation would be pointless. If two children considering whether to play together cannot see a net gain from having access to a larger total stock of toys or to one another's company, then it is better for each to "take his toys and play by himself." The world is full of uncertainty, and the expected benefits of an agreement may not materialize. But when engaged in bargaining, the parties must at least perceive some gain therefrom: Even Faust, when he agreed to sell his soul to the Devil, thought the benefits of knowledge and power that he would gain were worth the price that he would eventually have to pay.

The second important general point about bargaining follows from the first: Bargaining is not a zero-sum game. When a surplus exists, the negotiation is about how to divide it up. Each bargainer tries to get more for himself and leave less for the others. This may appear to be a zero-sum game, but behind it lies the danger that if an agreement is not reached, no one will get any surplus at all. This mutually harmful alternative, as well as *both* parties' desire to avoid it, is what creates the potential for the threats—explicit and implicit—that make bargaining such a strategic matter.

Before the advent of game theory, one-on-one bargaining was generally thought not to have any determinate equilibrium solution. Observations of widely different outcomes in otherwise similar-looking situations lent support to this view. Theorists were not able to achieve any systematic understanding of why one party got more than another and attributed these results to vague and inexplicable differences in "bargaining power."

Even the simple theory of Nash equilibrium does not take us any further. Suppose two people are to split $1. Let us construct a game in which each is asked to announce how much of the money he wants. The moves are simultaneous. If the players' announced amounts x and y add up to 1 or less, each gets the amount he announced. If they add up to more than 1, neither gets anything. Then *any* pair (x, y) adding up to 1 constitutes a Nash equilibrium in this game; *given* the announcement of the other, each player cannot do better than to stick to his own announcement.[1]

Further advances in game theory have brought progress along two quite different lines, each using a distinct mode of game-theoretic reasoning. In Chapter 2, we distinguished between *cooperative games*, in which the players decide and implement their actions jointly, and *noncooperative games*, in which the players decide and take their actions separately. Each of the two lines of advance in bargaining theory uses one of these two concepts. One approach views bargaining as a cooperative game, in which the parties find and implement a solution jointly, perhaps using a neutral third party such as an arbitrator for enforcement. The other approach views bargaining as a noncooperative game, in which the parties choose strategies separately, and looks for an equilibrium. In contrast to our earlier simple game of simultaneous announcements, whose equilibrium was indeterminate, this approach imposes more structure and specifies a sequential-move game of offers and counteroffers, which leads to a determinate equilibrium. As in Chapter 2, we emphasize that the labels *cooperative* and *noncooperative* refer to joint versus separate actions, not to nice versus nasty behavior or to compromise versus breakdown. The equilibria of noncooperative bargaining games can entail a lot of compromise.

Endnotes

- As we saw in Chapter 5, Section 2.B, this type of game can be used as an example to bolster the critique that the Nash equilibrium concept is too imprecise. In the bargaining context, we might say that the multiplicity of equilibria is just a formal way of showing the indeterminacy that previous analysts had claimed. Return to reference 1

1 THE NASH COOPERATIVE SOLUTION

In this section, we present John Nash’s cooperative-game approach to bargaining.[2] First we present the idea in a simple numerical example; then we develop the more general algebra.

A. Numerical Example

Imagine two Silicon Valley entrepreneurs, Andy and Bill. Andy produces a microchip set that he can sell to any computer manufacturer for $900. Bill has developed a software package that can retail for $100. The two meet and realize that their products are ideally suited to each other and that, with a bit of trivial tinkering, they can produce a combined system of hardware and software worth $3,000 in each computer. Thus, together they can produce an extra value of $2,000 per unit, and they can expect to sell millions of these units each year. The only obstacle that remains on this path to fortune is to agree to a division of the spoils. Of the $3,000 revenue from each unit, how much should go to Andy and how much to Bill?

Bill’s starting position is that without his software, Andy’s chip set is just so much metal and sand, so Andy should get only the $900 he could charge for the software and Bill himself should get $2,100. Andy counters that without his hardware, Bill’s programs are just symbols on paper or magnetic signals on a diskette, so Bill should get only $100, and $2,900 should go to Andy.

Watching them argue, you might suggest they “split the difference.” But that is not an unambiguous recipe for agreement. Bill might offer to split the profit on each unit equally with Andy. Under this scheme, each would get a profit of $1,000, meaning that $1,100 of the revenue would go to Bill and $1,900 to Andy. Andy’s response might be that each of them should get a percentage of the profit reflecting his contribution to the joint enterprise. Thus, Andy should get $2,700 and Bill $300.

The final agreement depends on their stubbornness or patience if they negotiate directly with each other. If they try to have the dispute arbitrated by a third party, the arbitrator’s decision depends on her sense of the relative value of hardware and software and on the rhetorical skills of the two principals as they present their arguments to her. For the sake of definiteness, suppose the arbitrator decides that the division of the profit should be 4:1 in favor of Andy; that is, Andy should get four-fifths of the surplus while Bill gets one-fifth, or Andy should get four times as much as Bill. What is the actual division of revenue under this scheme? Suppose Andy gets a total of x and Bill gets a total of y; thus Andy's profit is $(x - 900)$ and Bill's is $(y - 100)$. The arbitrator's decision implies that Andy's profit should be four times as large as Bill's; so $x - 900 = 4(y - 100)$, or $x = 4y + 500$. The total revenue available to both is \$3,000; so it must also be true that $x + y = 3{,}000$, or $x = 3{,}000 - y$. Then $x = 4y + 500 = 3{,}000 - y$, or $5y = 2{,}500$, or $y = 500$, and thus $x = 2{,}500$. This division mechanism leaves Andy with a profit of $2{,}500 - 900 = \$1{,}600$ and Bill with $500 - 100 = \$400$, which is the 4:1 split in favor of Andy that the arbitrator wants.

Next we develop these simple data into a general algebraic formula that you will find useful in many practical applications. Then we go on to examine more specifics of what determines the ratio between the players' shares of the surplus in a bargaining game.

B. General Theory

Suppose two bargainers, A and B, seek to split a total value v, which they can achieve if and only if they agree on a specific division. If no agreement is reached, A will get a and B will get b, each by acting alone or in some other way acting outside of their relationship. Call these their *backstop* payoffs or, in the jargon of the Harvard Negotiation Project, their BATNAs (best alternatives to a negotiated agreement).[3] Often a and b are both zero, but more generally, we need only assume that $a + b < v$, so that the agreement yields a positive surplus $(v - a - b)$; if this were not the case, the whole bargaining process would be moot because each side would just take up its outside opportunity and get its BATNA.

Consider the following rule: Each player is to be given his BATNA plus a share of the surplus, a fraction h of the surplus for A and a fraction k for B, such that $h + k = 1$. Writing x for the amount that A finally ends up with, and similarly y for B, we translate these statements as

$$x = a + h(v - a - b) = a(1 - h) + h(v - b)$$

$$x - a = h(v - a - b)$$

and

$$y = b + k(v - a - b) = b(1 - k) + k(v - a)$$

$$y - b = k(v - a - b).$$

We call this set of expressions the Nash formula. Another way of looking at them is to say that the surplus $(v - a - b)$ gets divided between the two bargainers in the proportion $h{:}k$, or

$$\frac{y-b}{x-a} = \frac{k}{h}$$

or, in slope-intercept form,

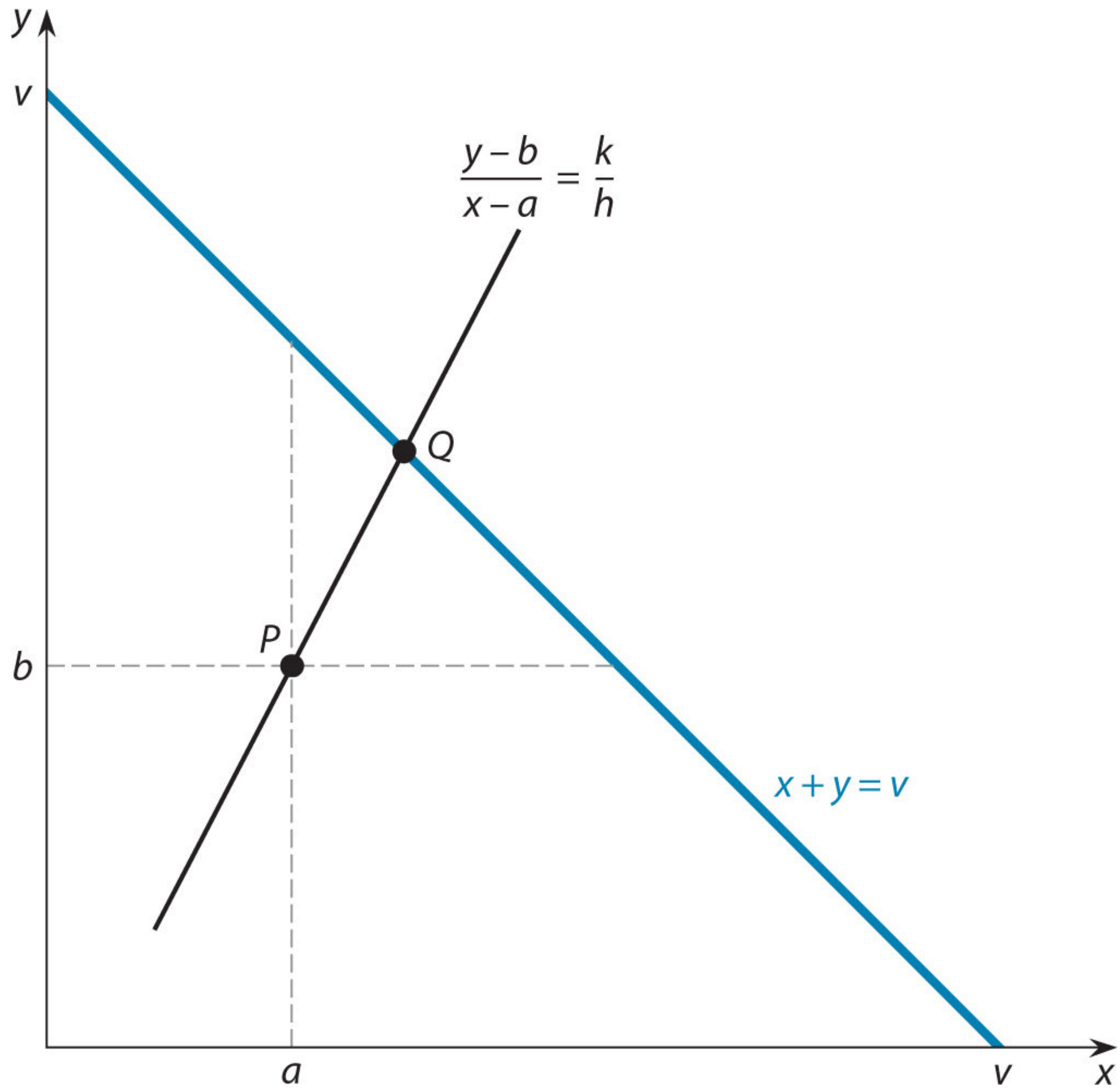


FIGURE 17.1 The Nash Cooperative Solution in the Simplest Case

$$y = b + \frac{k}{h}(x - a) = \left(b - \frac{ak}{h}\right) + \frac{k}{h}x.$$

To use up the whole surplus, x and y must also satisfy $x + y = v$. The Nash formula for x and y actually gives the solutions to these last two simultaneous equations.

The outcome of this formula is the Nash cooperative solution to the bargaining problem.[4] Its geometric representation is shown in Figure 17.1. The backstop, or BATNA, is the point P, with coordinates (a, b). All points (x, y) that divide the gains in proportions h:k between the two players lie along the straight line passing through P and having slope k/h; this slope is just the line $y = b + (k/h)(x - a)$ that we derived earlier. All points (x, y) that use up the whole surplus lie along the straight line joining $(v, 0)$ and $(0, v)$; this line is the second equation that we derived—namely, $x + y = v$. Nash's cooperative solution is at the intersection of the lines, at the point Q. The coordinates of this point are the parties' payoffs after the agreement.

The Nash formula says nothing about how or why such a solution might come about. And this vagueness is its merit—it can be used to encapsulate the results of many different theories taking many different perspectives.

At the simplest, you might think of the Nash formula as a shorthand description of the outcome of a bargaining process that we have not specified in detail. Then h and k can stand for the two parties' relative bargaining strengths. This shorthand description is a cop-out; a more complete theory should explain where these bargaining strengths come from and why one party might have more bargaining strength than the other. We will offer such explanations in a particular context later in this chapter. In the meantime, by summarizing any and all of the sources of bargaining strength in these numbers h and k, the formula has given us a good tool.

Nash's own interpretation of his cooperative solution differed from the view that it was a "shorthand description of some underlying game" as described in the previous paragraph. Indeed, Nash's interpretation differed from the whole approach to game theory that we have taken thus far in this book and deserves more careful explanation. In all the games that we have studied so far, the players chose and played their strategies separately from one another. We have looked for equilibria in which each player's strategy was in his own best interest, given the strategies of the others. Some such outcomes were very bad for some or even all of the players, the prisoners' dilemma being the most prominent example. In such situations, there was scope for the players to get together and agree that all would follow some particular strategy. But in our framework, there was no way in which they could be sure that the agreement would hold. After reaching an agreement, the players would disperse, and when it was each player's turn to act, he would actually take the action that served his own best interest. The agreement on joint action would unravel in the face of such separate temptations. In considering repeated games in Chapter 10, we found that the implicit threat of the collapse of an ongoing relationship might sustain an agreement, and in Chapter 9, we allowed for communication by signals. But individual action was of the essence, and any mutual benefit could be achieved only if it did not fall prey to the selfishness of separate individual actions. In Chapter 2, we called this approach to game theory *noncooperative*, emphasizing that the term signifies how actions are taken, not whether outcomes are jointly good. The important point, again, is that any joint good outcome has to be an equilibrium of separate actions in such games.

But what if joint action *is* possible? For example, the players might take all their actions immediately after the agreement is reached, in one another's presence. Or they might delegate the implementation of their joint agreement to a neutral third party, or to an arbitrator. In other words, the game might be *cooperative* (again, in the sense of joint action). Nash modeled bargaining as a cooperative game.

The thinking of a group that is going to implement a joint agreement by joint action can be quite different from that of a set of individual people who know that they are *interacting* strategically but are *acting* noncooperatively. Whereas the latter set will think in terms of an equilibrium and then delight or grieve depending on whether they like the results, the former can think first of what is a good outcome and then see how to implement it. In other words, the theory defines the outcome of a cooperative game in terms of some general principles or properties that seem reasonable to the theorist.

Nash formulated a set of such principles for bargaining and proved that they imply a unique outcome. His principles are roughly as follows: (1) The outcome should not change if the scale by which the payoffs are measured changes linearly; (2) the outcome should be efficient; and (3) if the set of possibilities is reduced by removing some that are irrelevant, in the sense that they would not be chosen anyway, then the outcome should not be affected.

The first of these principles conforms to the intuition that a change of scale merely changes the units of measurement for the entities being bargained over—for example, inches to centimeters or dollars to euros—and should not affect the result. A nonlinear rescaling would be different; as we mentioned in Chapter 9, that would correspond to a change in the players' attitudes toward risk, but we will leave that topic to more advanced treatments.

The second principle, that the outcome should be efficient, simply means that no available mutual gain should go unexploited. In our simple example of A and B splitting a total value of v, it would mean that x and y must sum to v, and not to any smaller amount; in other words, the solution has to lie on the $x + y = v$ line in Figure 17.1. More generally, the complete set of logically conceivable agreements in a bargaining game, when plotted on a graph as in Figure 17.1, will be bounded above and to the right by the subset of agreements that leave no mutual gain unexploited. This subset need not lie along a straight line such as $x + y = v$ (or $y = v - x$); it could lie along any curve of the form $y = f(x)$.

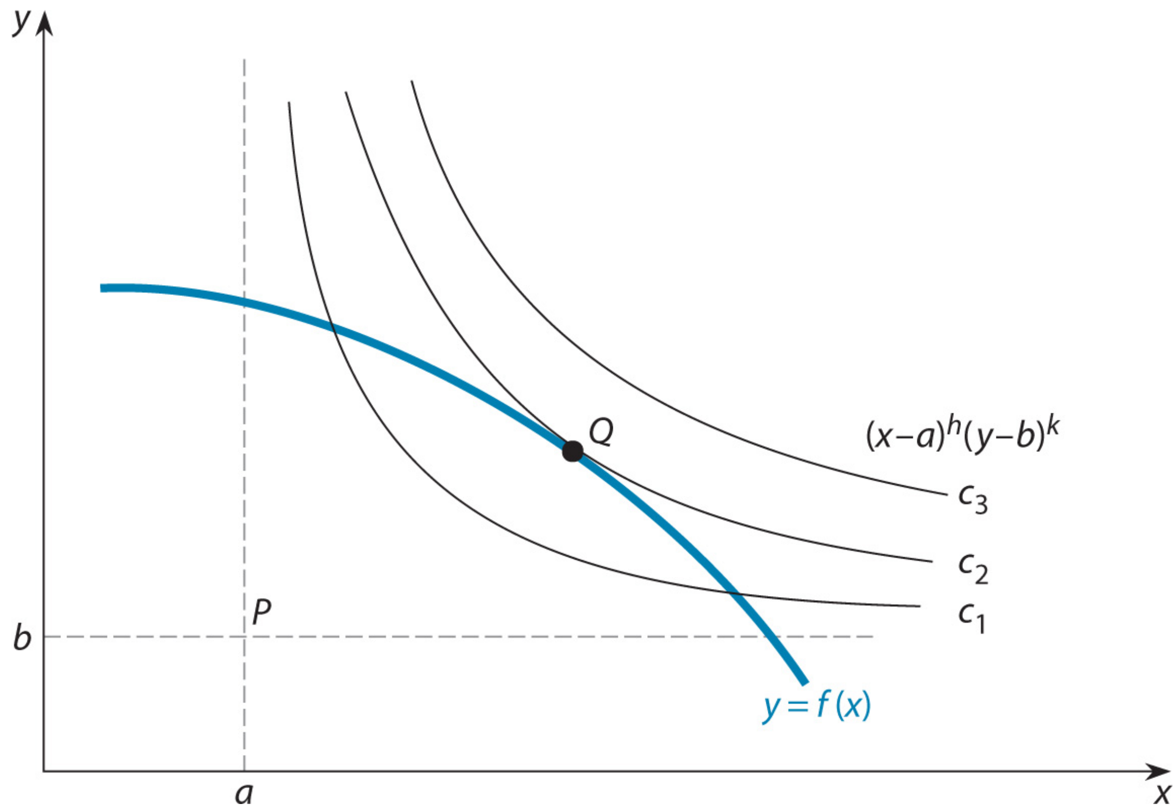


FIGURE 17.2 The General Form of the Nash Cooperative Solution

---

In Figure 17.2, all the points on and below (that is, "southwest" of) the thick blue curve labeled $y = f(x)$ constitute the complete set of conceivable outcomes. The curve itself consists of the efficient outcomes; there are no conceivable outcomes that include more of both x and y than the outcomes on $y = f(x)$, so there are no unexploited mutual gains left. Therefore, we call the curve $y = f(x)$ the efficient frontier of the bargaining problem.

To illustrate a curved efficient frontier, we consider two friends who jointly own a lottery ticket that has won a jackpot of $200,000 and are bargaining over its division. Each wants more money for himself, but the added pleasure each gets from an extra dollar diminishes as he gets more and more money. To be precise, we suppose that the pleasure payoff increases as the square root of the amount of money, so increasing the dollar amount fourfold increases the pleasure only by a factor of 2. If the first friend gets $\$z$ of the money and the other gets the remaining $\$(200,000 - z)$, their respective payoffs x and y will be

$$x = \sqrt{z} \text{ and } y = \sqrt{200,000 - z}.$$

Therefore, we can describe the set of possible payoff outcomes for the two friends by the equation

$x^2 + y^2 = z + (200,000 - z) = 200,000.$

This equation defines a quarter-circle in the positive quadrant and represents the efficient frontier of the friends' bargaining problem.

Now suppose that, if the friends fail to agree on a division of the jackpot, organizers of the lottery will submit their dispute to an arbitrator. This arbitrator would charge a 10% fee and divide the remaining \$180,000 equally between the two friends. So each would get \$90,000 and a payoff of

$\sqrt{90,000} = 300.$ That then is the BATNA of each, and the BATNA point (300, 300) lies inside the quarter-circle efficient frontier.

The third principle also seems appealing. If an outcome that a bargainer wouldn't have chosen anyway drops out of the picture, what should it matter? This assumption is closely connected to the "independence of irrelevant alternatives" assumption of Arrow's impossibility theorem, which we met in Chapter 16, Section 3, but we must leave the development of this connection to more advanced treatments of the subject.

Nash proved that the cooperative outcome that satisfied all three of these assumptions could be characterized by a mathematical maximization problem: Choose x and y to maximize $(x - a)^h(y - b)^k$ subject to $y = f(x)$. Here, x and y are the outcomes, a and b the

BATNAs, and h and k two positive numbers summing to 1, which are like the bargaining strengths of the Nash formula. The values for h and k cannot be determined by Nash's three assumptions alone; thus they leave a degree of freedom in the theory and in the outcome. Nash actually imposed a fourth assumption on the problem—that of symmetry between the two players; this additional assumption led to the outcome $h = k = ½$ and fixed a unique solution. We have given the more general formulation that subsequently became common in game theory and economics.

Figure 17.2 also gives a geometric representation of the objective of the maximization. The black curves labeled c_1, c_2, and c_3 are the level curves, or contours, of the function being maximized; along each such curve, $(x - a)^h(y - b)^k$ is constant and equals c_1, c_2, or c_3 (with $c_1 < c_2 < c_3$) as indicated. The whole space could be filled with such curves, each with its own value of the constant, and curves farther to the northeast would have higher values of the constant.

It is immediately apparent that the highest possible value of the function is at that point of tangency, Q, between the efficient frontier and one of the level curves.[5] The location of Q is defined by the property that the contour passing through Q is tangent to the efficient frontier. This tangency is the usual way to illustrate Nash's cooperative solution geometrically.[6]

In our example of Figure 17.1, we can also derive the Nash cooperative solution mathematically; to do so requires calculus, but the ends here are more important—at least to the study of games of strategy—than the means. For the solution, it helps to write $X = x - a$ and $Y = y - b$. Thus, X is the amount of the surplus that goes to A, and Y is the amount of the surplus that goes to B. The efficiency of the outcome guarantees that $X + Y = x + y - a - b = v - a - b$, which is just the total surplus and which we will write as S. Then $Y = S - X$, and

$$(x - a)^h(y - b)^k = X^hY^k = X^h(S - X)^k.$$

In the Nash cooperative solution, X takes on the value that maximizes this function. Elementary calculus tells us that the way to find X is to take the derivative of this expression with respect to X and set it equal to zero. Using the rules of calculus for taking the derivatives of powers of X and of the product of two functions of X, we get

$$hX^{h-1}(S - X)^{k} - X^{h}k(S - X)^{k-1} = 0.$$

When we cancel the common factor $X^{h-1}(S - X)^{k-1}$, this equation becomes

$$h(S-X)-kX=0$$

$$hY-kX=0$$

$$kX=hY$$

$$\frac{X}{h}=\frac{Y}{k}.$$

Finally, expressing the equation in terms of the original variables x and y, we have $(x - a)/h = (y - b)/k$, which is just a rearranged version of the Nash formula we presented above. The punch line: Nash's three conditions lead to the formula we originally stated as a simple way of splitting the bargaining surplus.

The three principles, or desired properties, that determine the Nash cooperative solution are simple and even appealing. But in the absence of a good mechanism to make sure that the parties take the actions stipulated by the agreement, these principles may come to nothing. A player who can do better by strategizing on his own than by using the Nash cooperative solution may simply reject the principles. If an arbitrator can enforce a solution, the player may simply refuse to go to arbitration. Therefore the

Nash cooperative solution will seem more compelling if it can be given an alternative interpretation—namely, as the Nash equilibrium of a noncooperative game played by the bargainers. This can indeed be done, and we will develop an important special case of it in Section 4.

Endnotes

- John F. Nash Jr., “The Bargaining Problem,” *Econometrica*, vol. 18, no. 2 (1950), pp. 155–62. Return to reference 2
- See Roger Fisher and William Ury, *Getting to Yes*, 2nd ed. (New York: Houghton Mifflin, 1991). Return to reference 3
- This outcome is also commonly known as the *Nash bargaining solution*. Our terminology more clearly distinguishes the cooperative solution from the alternative noncooperative solutions to bargaining games (based on Nash equilibrium) that are analyzed later in this chapter (Sections 3 through 7). Return to reference 4
- One and only one of the (convex) level curves can be tangential to the (concave) efficient frontier; in Figure 17.2, this level curve is labeled c_2. All lower-level curves (such as c_1) cut the frontier at two points; all higher-level curves (such as c_3) do not meet the frontier at all. Return to reference 5
- If you have taken an elementary microeconomics course, you will have encountered the concept of social optimality, illustrated graphically by the tangent point between the production possibility frontier of an economy and a social indifference curve. Our Figure 17.2 is similar in spirit; the efficient frontier in bargaining is like the production possibility frontier, and the level curves of the objective in cooperative bargaining are like social indifference curves. Return to reference 6

Glossary

best alternative to a negotiated agreement (BATNA)
: In a bargaining game, this is the payoff a player would get from his other opportunities if the bargaining in question failed to reach an agreement.

surplus
: A player’s surplus in a bargaining game is the excess of his payoff over his BATNA.

Nash cooperative solution
: This outcome splits the bargainers’ surpluses in proportion to their bargaining powers.

efficient
: An outcome of a bargaining game is called efficient if there is no feasible alternative that would leave one bargainer with a higher payoff without reducing the payoff of the other.

efficient frontier
: This is the northeast boundary of the set of feasible payoffs of the players, such that in a bargaining game it is not possible to increase the payoff of one person without lowering that of another.

2 VARIABLE-THREAT BARGAINING

The outcome of bargaining depends on the players' backstops (BATNAs). The case of equal bargaining strengths suffices to illustrate this result: With $h = k = ½$, the outcomes for players A and B become $x = (v + a - b)/2$ and $y = (v + b - a)/2$. So far, we have taken the BATNAs to be fixed by considerations outside of this bargaining game. But what if one or both players can change the BATNAs? If such manipulation is possible, the bargaining game acquires a pregame stage not unlike the pregame we described in our discussion of strategic moves in Chapter 8. In the pregame to the bargaining game, players can manipulate BATNAs, within some limitations set by the rules of the pregame, to get themselves a better outcome in the upcoming game. This two-stage game (manipulation of BATNAs followed by bargaining) is called variable-threat bargaining. The difference between this situation and the ones described in Chapter 8 is that here, the upcoming game is played cooperatively while the pregame is played noncooperatively, whereas both stages are noncooperative in our strategic move analysis. The principle of rollback applies here just the same, however. When choosing their noncooperative actions to change BATNAs, the players look ahead to the effect of those actions on the outcome of the upcoming cooperative game.

We can see from the formula for outcomes in the case of equal bargaining strengths that Player A stands to gain by increasing $(a - b)$, while Player B stands to gain by decreasing the same difference. It should be intuitive that each player can gain by increasing his own BATNA and by decreasing the other's. But, perhaps surprisingly, it is also true that A can gain even if both a and b *decrease*, so long as $(a - b)$ *increases*; that is, if B's BATNA decreases *more* than A's. On reflection, this should make sense: If a breakdown of negotiations would hurt B more than it hurts A, that should lead to a bargaining outcome that is more favorable to A.

This type of result is nicely illustrated in a scene from the movie *Ransom*. Sean, the son of multimillionaire Tom Mullen (played by Mel Gibson), has been kidnapped. The men holding him are demanding a ransom of $2 million. Mullen goes on live TV and, in a sequence of shots, we see the people involved, including the lead kidnapper, Jimmy Shaker (played by Gary Sinise), watching from various locations. With the full $2 million spread out on a table before him, Mullen announces that he is instead offering it as a reward on the kidnapper's head. But if the kidnapper returns the boy alive and unharmed, Mullen will withdraw the bounty.

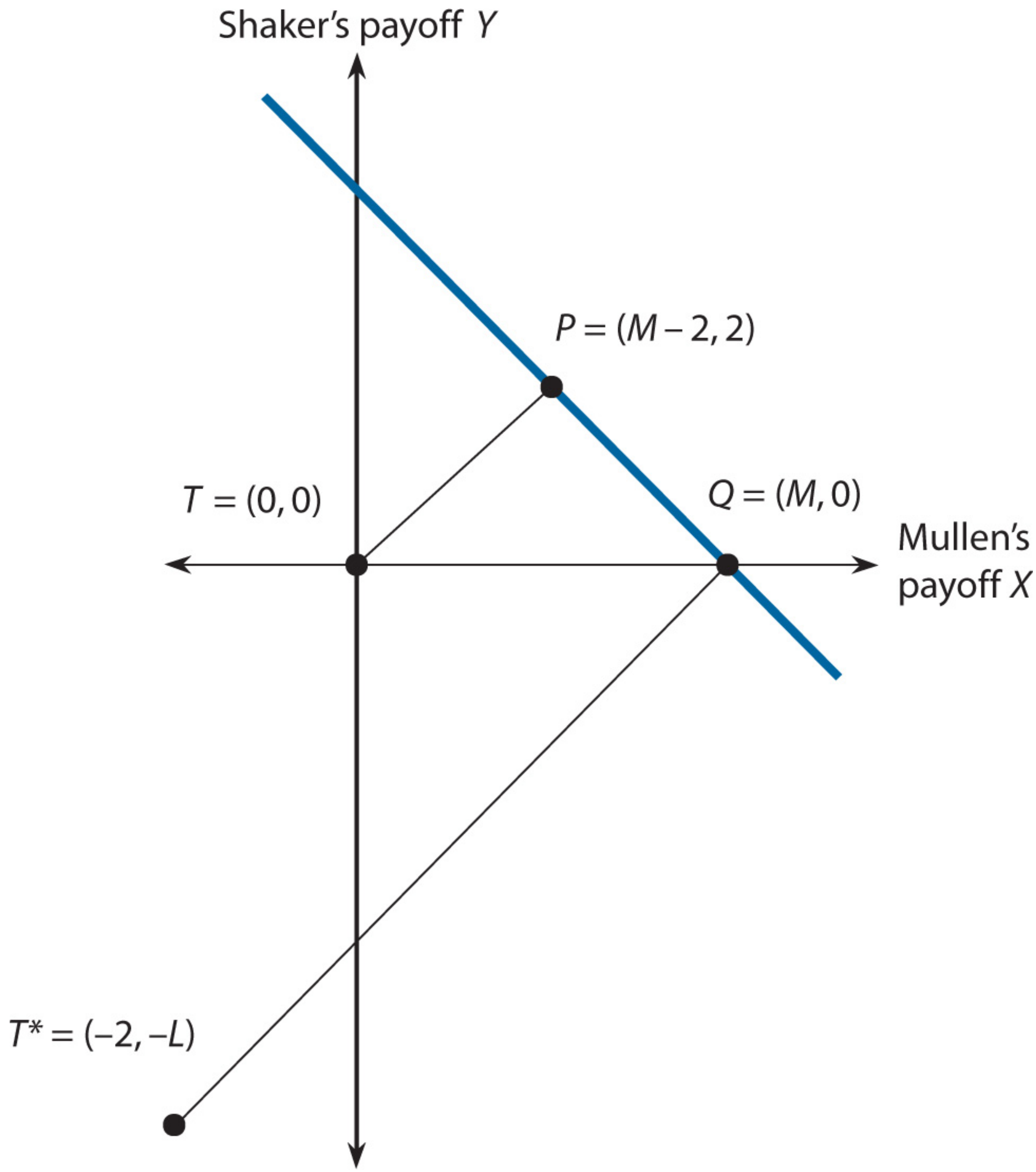


FIGURE 17.3 The Variable-Threat Game of Ransom

Let us represent this interaction in game-theoretic language. Figure 17.3 shows Mullen's payoff on the horizontal axis and Shaker's on the vertical axis. The origin is at the point of their initial wealths. Before the kidnapping, Mullen also has his son; we let M denote his son's value to him in money-equivalent terms (millions of dollars). So the payoff point in the status quo is (M, 0), or the point Q in the figure. The line through Q with slope − 1 shows all payoff points that can be achieved by transferring money from one player to another with Q representing the initial, no-transfers, point; this line is therefore the bargaining frontier.

In the original scenario, after the kidnapping but prior to any additional actions, if negotiation fails, then Mullen loses his son, but Shaker gets no ransom, so the BATNA is the origin, labeled T in Figure 17.3. Shaker has asked for a $2 million ransom. If the ransom is paid and Mullen gets his son back, the payoffs (in millions of dollars) will be $M - 2$ to Mullen and 2 to Shaker. That outcome is shown as point P in the figure.

Mullen's strategy moves the BATNA point. Now, if the negotiation fails, Mullen will lose his son and end up paying the $2 million bounty to the person who kills Shaker (probably one of Shaker's confederates), while Shaker will lose his life. Let Shaker's valuation of his own life be denoted by L (in millions of dollars). The payoffs at the new BATNA are −2 to Mullen and $-L$ to Shaker, and the new BATNA is point T^* in the figure.

With this new BATNA in place, Mullen offers Shaker a solution to the bargaining problem—namely, going back to the status quo point Q. That point is attained if Shaker brings Sean back unharmed and Mullen withdraws the bounty on Shaker's head. The figure is drawn in a way that suggests that this outcome conforms to the Nash cooperative solution with equal bargaining strengths. It is easy to verify that this solution implicitly sets $M = 4$ and $L = 6$. For other values of M and L, the Nash cooperative solution for Mullen's counterproposal need not be exactly at Q.

Under what circumstances will Mullen's strategy give him an outcome better than the P he would get by acceding to Shaker's original demand? The Nash cooperative solution for the threat point T^* would have to be located to the southeast of that for T along the bargaining frontier. This situation arises if the line T^*Q lies below the line TP (as is true in Figure 17.3). That configuration of the two lines occurs when $L > 2$, or when Shaker values his own life more than the ransom money.

Because T^* is to the southwest of T, Mullen's change in the threat point (from T to T^*) worsens both players' BATNAs. His strategy aims to encourage Shaker's acceptance of his alternative proposal, because that proposal carries the threat of an even bigger loss for Shaker than for Mullen if the negotiation fails. In other words, Mullen is implicitly saying to Shaker, "This will hurt you more than it will hurt me." We often hear such statements made in arguments and disputes; now we see the strategic role they play in negotiations.[7]

For scenarios in which Shaker values his life very highly, with $L > 6$ in our figure, the line from T^* would meet the bargaining frontier at a point southeast of the status quo point, Q. The Nash cooperative solution in that case would correspond to a negative payoff for Shaker, and he would end up with wealth below his original level. We could interpret such an outcome as equivalent to a situation in which Shaker is paying Mullen to take his son back! This kind of outcome may seem impossible in practice. If Mullen's threat, T^*, is that severe, the bargaining outcome may actually be at Q. However, it is not outside the realm of possibility that kidnappers would agree to pay to be rid of their hostage, at least in some other branches of fiction.[8]

The variable-threat strategy of stating, "This will hurt you more than it will hurt me" has been used in real-world bargaining situations, too. For example, smart labor unions threaten or launch strikes at times when they will deliver the biggest hit to firms' profits. The British coal miners' union did this consistently in the 1970s, although, in that case, the hit was not so much to the nationalized coal industry's profits as to the political popularity of the government affected by the disruption of power supplies. Conversely, Prime Minister Margaret Thatcher's strategy of provoking the union's leader, Arthur Scargill, into striking in the spring and summer of 1984, when the demand for coal was lower and the disruption caused by a strike correspondingly smaller, was instrumental in the collapse of the strike and led to a collapse of the union itself.

The same strategy was used in the Major League Baseball strike of 1980.[9] The strike started during the exhibition games of the preseason. The players returned to work (actually, to play) at the start of the regular season, but resumed the strike after Memorial Day. This curious discontinuous strike can be better understood when we examine the time-varying costs of the strike to the two sides. During the exhibition games, the players are not paid salaries, but the team owners earn substantial revenues from fans who combine a vacation in a warmer clime with following their favorite team’s stars and prospects. Once the regular season starts, the players get salaries, but attendance at games, and therefore the owners’ revenues, are low; those numbers grow substantially only after Memorial Day. Therefore the discontinuous strike was the players’ correct strategy to maximize the owners’ losses relative to their own.

Endnotes

- Movies have their own requirements of dramatic tension and denouement that override game-theoretic logic. To conform to those demands, *Ransom* does not have an efficient resolution on the bargaining frontier where one of the players accedes to the other's demand, but rather twists that end in chases and gunfights. That spin on the story is not material to the basic bargaining game we want to illustrate. Return to reference 7
- In O. Henry's short story "The Ransom of Red Chief," two small-time crooks kidnap a banker's 10-year-old son. He turns out to be a brat who makes their lives so impossible that they pay the father to take him back. The text is available at http://fiction.eserver.org/short/ransom_of_red_chief.xhtml (accessed February 17, 2016). Return to reference 8
- Lawrence M. DeBrock and Alvin E. Roth, "Strike Two: Labor-Management Negotiations in Major League Baseball," *Bell Journal of Economics*, vol. 12, no. 2 (Autumn, 1981), pp. 413–25. Return to reference 9

Glossary

variable-threat bargaining

A two-stage game where at the first stage you can take an action that will alter the BATNAs of both bargainers (within certain limits), and at the second stage bargaining results in the Nash solution on the basis of these BATNAs.

3 ALTERNATING-OFFERS MODEL I: TOTAL VALUE DECAYS

In this section, we move back to the more realistic noncooperative-game approach and think about the process of individual strategizing that may produce an equilibrium in a bargaining game. Our standard picture of this process is one of alternating offers. One player—say, A—makes an offer. The other—say, B—either accepts it or makes a counteroffer. If he does the latter, then A can either accept it or come back with another offer of his own. And so on. Thus, we have a sequential-move game, and we can look for its rollback equilibrium.

To find a rollback equilibrium, we must start at the end of the game and work backward. But where is the end? Why should the process of offers and counteroffers ever terminate? Perhaps more drastic is the question, Why would it ever start? Why would the two bargainers not stick to their original positions and refuse to budge? It is costly to both if they fail to agree at all, but the benefit of an agreement is likely to be smaller to the one who makes the first, or the larger, concession. The reason why anyone concedes must be that continuing to stand firm would cause an even greater loss of benefit. This loss takes one of two broad forms. The available surplus may decay (shrink) with each offer, a possibility that we consider in this section. The alternative possibility is that time has value and impatience is important, and so a delayed agreement is worth less; we examine this possibility in Section 5.

Consider the following story of bargaining over a shrinking surplus. A fan arrives at a professional football (or basketball) game without a ticket. He is willing to pay as much as $25 to watch each quarter of the game. He finds a scalper, who states a price for a ticket. If the fan is not willing to pay this price, he goes to a nearby bar to watch the first quarter on its big-screen TV. At the end of the quarter, he comes out, finds the scalper still there, and makes a counteroffer for the ticket. If the scalper does not agree, the fan goes back to the bar. He comes out again at the end of the second quarter, when the scalper makes him yet another offer. If that offer is not acceptable to the fan, he goes back into the bar, emerging at the end of the third quarter to make yet another counteroffer. The value of watching the rest of the game is declining as the quarters go by.[10]

Rollback analysis enables us to predict the outcome of this alternating-offers bargaining process. At the end of the third quarter, the fan knows that if he does not buy the ticket then, the scalper will be left with a small piece of paper of no value. So the fan will be able to make a very small offer that, for the scalper, will still be better than nothing. Thus, on his last offer, the fan can get the ticket almost for free. Backing up one period, we see that at the end of the second quarter, the scalper has the initiative in making an offer. But he must look ahead and recognize that he cannot hope to extract the whole of the remaining two quarters' value from the fan. If the scalper offers the ticket for more than $25—the value of the *third* quarter to the fan—the fan will turn down the offer because he knows that he can get the fourth quarter later for almost nothing, so the scalper can ask for $25 at most. Now consider the situation at the end of the first quarter. The fan knows that if he does not buy the ticket now, the scalper can expect to get only $25 later, and so $25 is all that the fan needs to offer now to secure the ticket. Finally, before the game even begins, the scalper can look ahead and ask for $50; this $50 includes the $25 value of the *first* quarter to the fan plus the $25 for which the fan can get the remaining three quarters’ worth. Thus, the two will strike an immediate agreement, and the ticket will change hands for $50, but the price is determined by the full forward-looking rollback reasoning process.[11]

This story can be easily turned into a more general theory of negotiation between two bargainers, A and B. Suppose A makes the first offer to split the total surplus, which we call v (measured in some currency—say, dollars). If B refuses the offer, the total surplus available drops by x_1 to $(v - x_1)$. B offers a split of this amount. If A refuses B’s offer, the total drops by a further amount x_2 to $(v - x_1 - x_2)$. A offers a split of this amount. This offer and counteroffer process continues until finally, say, after 10 rounds, $v - x_1 - x_2 - . . . - x_{10} = 0$, so the game ends. As usual with sequential-move games, we begin our analysis at the end.

If the game has gone to the point where only x_{10} is left, B can make a final offer whereby he gets to keep almost all of the surplus, leaving a measly cent or so to A. Left with the choice of that or absolutely nothing, A should accept the offer. To avoid the finicky complexity of keeping track of tiny cents, let us call this outcome “x_{10} to B, 0 to A.” We will do the same in the other (earlier) rounds.

Knowing what is going to happen in round 10, we turn to round 9. Here, A will make the offer, and $(x_9 + x_{10})$ is left. A knows that he must offer at least x_{10} to B, or else B will refuse the offer and take the game to round 10, where he can get that much. Bargainer A does not want to offer any more to B. So, on round 9, A will offer a split where he keeps x_9 and leaves x_{10} to B.

Then, on the round before that, when $(x_8 + x_9 + x_{10})$ is left, B will offer a split where he gives x_9 to A and keeps $(x_8 + x_{10})$. Working backward to the very first round, A will offer a split where he keeps ($x_1 + x_3 + x_5 + x_7 + x_9$) and gives ($x_2 + x_4 + x_6 + x_8 + x_{10}$) to B. This offer will be accepted.

You can remember these formulas by means of a simple trick: *Hypothesize* a sequence in which all offers are refused. (This sequence is *not* what actually happens.) Then add up the amounts of the surplus that would be destroyed by the refusals of one player. This total is what the other player gets in the actual equilibrium. For example, if B refused A’s first offer, the total available surplus would drop by x_1, so x_1 becomes part of what goes to A in the equilibrium of the game.

If each player has a positive BATNA, the analysis must be modified somewhat to take that into account. In the last round, B must offer A at least the BATNA a. If x_{10} is greater than a, B is left with ($x_{10} - a$); if not, the game must terminate before this round is reached. Now, at round 9, A must offer B the larger of two amounts: the ($x_{10} - a$) that B can get in round 10, or the BATNA b that B can get outside the agreement. The analysis can proceed all the way back to round 1 in this way; we leave it to you to complete the rollback reasoning for this case.

We have found the rollback equilibrium of the alternating-offers bargaining game, and in the process of deriving the outcome, we have also described the full strategies (complete contingent plans of action) behind the equilibrium—namely, what each player *would* do if the game reached some later stage. In fact, actual agreement is immediate on the very first offer. The later stages are not reached; they are off-equilibrium nodes and paths of play. But, as is usual with rollback reasoning, the foresight about what rational players would do at those nodes if they were reached is what informs the initial action.

The other important point to note is that *gradual decay* (with several potential rounds of offers) leads to a more even or fairer split of the surplus than does *sudden decay* (with only one round of bargaining permitted). In the latter case, no agreement would result if B turned down A's very first offer; then, in a rollback equilibrium, A would get to keep (almost) the whole surplus, giving B an "ultimatum" to accept a measly cent or else get nothing at all. The subsequent rounds give B the credible ability to refuse a very uneven first offer.

Endnotes

- Just to keep the argument simple, we imagine this process as one-on-one bargaining. Actually, there may be several fans and several scalpers, turning the situation into a *market.* You can access our supplemental chapter on interactions in markets at digital.wwnorton.com/gamesofstrategy5. Return to reference 10
- To keep the analysis simple, we have omitted the possibility that the game might get exciting, and so the value of the ticket might actually increase as the quarters go by. This uncertainty would make the problem much more complex, but also more interesting. The ability of game theory to deal with such possibilities should inspire you to go beyond this book or course to study more advanced game theory. Return to reference 11

Glossary

alternating offers
: A sequential move procedure of bargaining in which, if the offer made by one player is refused by the other, then the refuser gets the next turn to make an offer, and so on.

decay
: Shrinkage over time of the total surplus available to be split between the bargainers, if they fail to reach an agreement for some length of time during the process of their bargaining.

impatience
: Preference for receiving payoffs earlier rather than later. Quantitatively measured by the discount factor.

4 ALTERNATING-OFFERS MODEL II: IMPATIENCE

In this section, we consider a different kind of cost of delay in reaching an agreement. Suppose the actual monetary value of the total surplus available for splitting does not decay, but players prefer to receive money earlier rather than later, and will accept somewhat less now rather than having to wait until later to get a bit more. For concreteness, we will say that both bargainers believe that having only 95 cents right now is as good as having $1 one round later. They make alternate offers as in Section 3.

A player who prefers having something right away to having the same thing later is *impatient*; he attaches less importance to the future than to the present. We came across this idea in Chapter 10, Section 2, and saw two reasons for it. First, Player A may be able to invest his money—say, $1—now and get his principal back along with interest and capital gains at a rate of return r, for a total gain of $(1 + r)$ in the next period (tomorrow, next week, next year, or whatever is the length of the period). Second, there may be some risk that the game will end between now and the next offer (as in the sudden end after 3 to 5 minutes in the classroom game described earlier). If p is the probability that the game continues, then the chance of getting a dollar next period has an expected value of only p now.

Suppose we consider a bargaining process between two players with zero BATNAs. Let us start the process with one of the two bargainers—say, A—making an offer to split $1. If the other player, B, rejects A's offer, then B will have an opportunity to make his own offer one round later. The two bargainers are in identical situations when each makes his offer, because the amount to be split is always \$1. Thus, in equilibrium, the amount that goes to the player currently in charge of making the offer (call it x) is the same, regardless of whether that person is A or B. We can use rollback reasoning to find an equation that can be solved for x.

Suppose A starts the alternating-offers process. He knows that B can get x in the next round when it is B's turn to make the offer. Therefore, A must give B at least an amount that is equivalent, in B's eyes, to getting x in the next round; A must give B at least $0.95x$ now. (Remember that, for B, 95 cents received now is equivalent to \$1 received in the next round; so $0.95x$ now is as good as x in the next round.) Player A will not give B any more than is required to induce B's acceptance. Thus, A offers B exactly $0.95x$ and is left with $(1 - 0.95x)$. But the amount that A gets when making the offer is just what we called x. Therefore, $x = 1 - 0.95x$, or $(1 + 0.95)x = 1$, or $x = 1/1.95 = 0.512$.

Two things about this calculation should be noted. First, even though the process allows for an unlimited sequence of alternating offers and counteroffers, in equilibrium, the very first offer A makes gets accepted, and the bargaining ends. Because time has value, this outcome is efficient. The cost of delay governs how much A must offer B to induce acceptance; it thus affects A's rollback reasoning. Second, the player who makes the first offer gets more than half of the surplus—namely, 0.512 rather than 0.488. Thus, each player gets more when he makes the first offer than when the other player makes the first offer. But this advantage is far smaller than that in an ultimatum game with no future rounds of counteroffers.

Now suppose the two players are not equally patient (or impatient, as the case may be). Player B still regards \$1 in the next round as equivalent to 95 cents now, but A regards it as equivalent to only 90 cents now. Thus, A is willing to accept a smaller amount of money to get his money sooner; in other words, A is more impatient. This inequality in rates of impatience can translate into unequal payoffs from the bargaining process in equilibrium. To find the equilibrium for this example, we write x for the amount that A gets when he starts the process and y for what B gets when he starts the process.

Player A knows that he must give B at least $0.95y$; otherwise B will reject the offer in favor of the y that he knows he can get when it becomes his turn to make the offer. Thus, the amount that A gets, x, must be $1 - 0.95y$; $x = 1 - 0.95y$. Similarly, when B starts the process, he knows that he must offer A at least $0.90x$, and then $y = 1 - 0.90x$. These two equations can be solved for x and y:

$x = 1 - 0.95(1 - 0.9x)$		$y = 1 - 0.9(1 - 0.95y)$
$[1 - 0.95(0.9)]x = 1 - 0.95$	and	$[1 - 0.9(0.95)]y = 1 - 0.9$
$0.145x = 0.05$		$0.145y = 0.10$
$x = 0.345$		$y = 0.690$

Note that x and y do not add up to 1, because each of these amounts is the payoff to a given player when he makes the first offer. Thus, when A makes the first offer, A gets 0.345 and B gets 0.655; when B makes the first offer, B gets 0.69 and A gets 0.31. Once again, each player does better when he makes the first offer than when the other player makes the first offer, and once again the advantage from being able to make the first offer is small.

The outcome of this case with unequal rates of impatience differs from that of the preceding case, with equal rates of impatience, in a major way. With unequal rates of impatience, the more impatient player, A, gets a lot less than B even when he is able to make the first offer. We expect that the person who is willing to accept less to get it sooner ends up getting less, but the difference is very dramatic. The proportion of A's share to B's share is almost 1:2.

As usual, we can now build on these examples to develop the more general algebra. Suppose A regards \$1 immediately as being equivalent to \$$(1 + r)$ one round later or, equivalently, A regards \$$1/(1 + r)$ immediately as being equivalent to \$1 one round later. For brevity, we substitute a for $1/(1 + r)$ in the calculations that follow. Likewise, suppose Player B regards \$1 today as being equivalent to \$$(1 + s)$ one round later; we use b for $1/(1 + s)$. If r is high (or equivalently, if a is low), then Player A is very impatient. Similarly, B is impatient if s is high (or if b is low).

Here we look at bargaining that takes place in alternating rounds, with a total of \$1 to be divided between two players, both of whom have zero BATNAs. (You can solve the even more general case easily once you understand this one.) What is the rollback equilibrium?

We can find the payoffs in equilibrium by extending the simple argument used earlier. Suppose A's payoff in the rollback equilibrium is x when he makes the first offer; B's payoff in the rollback equilibrium is y when he makes the first offer. We look for a pair of equations linking the values x and y and then solve these equations to determine the equilibrium payoffs.[12]

When A is making the offer, he knows that he must give B an amount that B regards as being equivalent to y one period later. This amount is $by = y/(1 + s)$. Then, after making the offer to B, A can keep only what is left: $x = 1 - by$.

Similarly, when B is making the offer, he must give A the equivalent of x one period later—namely, ax. Therefore, $y = 1 - ax$. Solving these two equations is now a simple matter.

We have $x = 1 - b(1 - ax)$, or $(1 - ab)x = 1 - b$. Expressed in terms of r and s, this equation becomes

$$x = \frac{1-b}{1-ab} = \frac{s+rs}{r+s+rs}.$$

Similarly, $y = 1 - a(1 - by)$, or $(1 - ab)y = 1 - a$. This equation becomes

$$y = \frac{1-a}{1-ab} = \frac{r+rs}{r+s+rs}.$$

Although this quick solution might seem like sleight of hand, it follows the same steps used earlier, and we will soon give a different method of reasoning yielding exactly the same answer. First, let us examine some features of this answer.

First, note that, as in our simple unequal-impatience example, the two magnitudes x and y add up to more than 1:

$$x+y = \frac{r+s+2rs}{r+s+rs} > 1.$$

Remember that x is what A gets when he has the right to make the first offer, and y is what B gets when he has the right to make the first offer. When A makes the first offer, B gets $(1 - x)$, which is less than y; this just shows A’s advantage from being the first proposer. Similarly, when B makes the first offer, B gets y and A gets $(1 - y)$, which is less than x.

However, usually r and s are small numbers. When offers can be made at short intervals, such as a week or a day or an hour, the interest that your money can earn from one offer to the next, or the probability that the game ends precisely within the next interval, is quite small. For example, if r is 1% (0.01) and s is 2% (0.02), then the formulas yield x = 0.668 and y = 0.337; so the advantage of making the first offer is only 0.005. (A gets 0.668 when making the first offer, but 1 − 0.337 = 0.663 when B makes the first offer; the difference is only 0.005.) More formally, when r and s are both small compared with 1, their product rs is very small indeed; thus, we can ignore rs to write an approximate solution for the split that does not depend on which player makes the first offer:

$$x = \frac{s}{r+s} \quad \text{and} \quad y = \frac{r}{r+s}.$$

Now $x + y$ is approximately equal to 1.

Most importantly, x and y in the approximate solution are the shares of the surplus that go to the two players, and $y/x = r/s$; that is, the shares of the players are inversely proportional to their rates of impatience as measured by r and s. If B is twice as impatient as A, then B gets half as much as A; so the shares are ⅓ and ⅔, or 0.333 and 0.667, respectively. Thus, we see that patience is an important advantage in bargaining. Our formal analysis supports the intuition that if you are very impatient, the other player can offer you a quick but poor deal, knowing that you will accept it.

This effect of impatience hurts the United States in numerous negotiations that its government agencies and diplomats conduct with other countries. The American political process puts a great premium on speed. The media, interest groups, and rival politicians all demand results and are quick to criticize the administration or the diplomats for any delay.

Under this pressure to deliver, the negotiators are always tempted to come back with results of any kind. Such results are often poor from the long-term U.S. perspective; the other countries' concessions often have loopholes, and their promises are less than credible. The U.S. administration hails the deals as great victories, but they usually unravel after a few years. The financial crisis of 2008 offers another, even more dramatic, example. When the housing boom collapsed, some major financial institutions that held mortgage-backed assets faced bankruptcy. That led them to curtail credit, which in turn threatened to throw the U.S. economy into a severe recession. The crisis exploded in September, in the midst of a presidential election campaign. The Treasury, the Federal Reserve, and political leaders in Congress all wanted to act fast. This impatience led them to offer generous terms of rescue to many financial institutions, when a slower process would have yielded an outcome that cost the taxpayers much less and offered them much better prospects of sharing in future gains on the assets being rescued.

Similarly, individuals who have suffered losses are in a much weaker position when they negotiate with insurance companies on coverage. The companies often make lowball offers of settlement to people who have suffered a major loss, knowing that their clients urgently want to make a fresh start and are therefore very impatient.

As a conceptual matter, the formula $y/x = r/s$ ties our noncooperative-game approach to bargaining to Nash's cooperative solution discussed in Section 1. The formula for shares of the available surplus that we derived there becomes, with zero BATNAs, $y/x = k/h$. In the cooperative approach, the shares of the two players stood in the same proportions as their bargaining strengths, but those strengths were assumed to be imposed somehow from the outside. Now we have an explanation for the bargaining strengths in terms of some more basic characteristics of the players: h and k are inversely proportional to the players' rates of impatience r and s. In other words, Nash's cooperative solution can also be given an alternative, and perhaps more satisfactory, interpretation as the rollback equilibrium of a noncooperative game of offers and counteroffers, if we interpret the abstract bargaining-strength parameters in the cooperative solution correctly in terms of the players' characteristics, such as impatience.

Finally, note that agreement is once again immediate—the very first offer is accepted. As usual, the rollback analysis imposes discipline by making the first proposer recognize that the other would credibly reject a less adequate offer.

To conclude this section, we offer an alternative derivation of the same (precise) formula for the equilibrium offers that we derived earlier. Suppose this time that there are 100 rounds of bargaining; A is the first proposer and B the last. Start the backward induction in the 100th round; B will keep the whole dollar. Therefore, in the 99th round, A will have to offer B the equivalent of \$1 in the 100th round—namely, b, and A will keep $(1 - b)$. Then proceed backward:

In round 98, B offers $a(1 - b)$ to A and keeps

$1 - a(1 - b) = 1 - a + ab.$

In round 97, A offers $b(1 - a + ab)$ to B and keeps

$1 - b(1 - a + ab) = 1 - b + ab - ab^2.$

In round 96, B offers $a(1 - b + ab - ab^2)$ to A and keeps

$1 - a + ab - a^2b + a^2b^2.$

In round 95, A offers $b(1 - a + ab - a^2b + a^2b^2)$ to B and keeps

$1 - b + ab - ab^2 + a^2b^2 - a^2b^3.$

Proceeding in this way and following the established pattern, we see that in round 1, A gets to keep

$1 - b + ab - ab^2 + a^2b^2 - a^2b^3 + \ldots + a^{49}b^{49} - a^{49}b^{50}$

$= (1 - b)[1 + ab + (ab)^2 + \ldots + (ab)^{49}].$

The consequence of allowing more and more rounds is now clear: We just get more and more of these terms, growing geometrically by the factor *ab* for every two offers. To find A's payoff when he is the first proposer in an infinitely long sequence of offers and counteroffers, we have to find the limit of the infinite geometric sum. In the appendix to Chapter 10, we saw how to sum such series. Using the formula obtained there, we get the answer

$$(1 - b)[1 + ab + (ab) + (ab)^2 + \cdots + (ab)^{49} + \cdots] = \frac{1 - b}{1 - ab}.$$

This is exactly the solution for x that we obtained before. By a similar argument, you can find B's payoff when he is the proposer and, in doing so, improve your understanding and technical skills at the same time.

Endnotes

- We are taking a shortcut; we have simply assumed that such an equilibrium exists and that the payoffs are uniquely determined. More rigorous theory proves these conditions. For a step in this direction, see John Sutton, “Non-Cooperative Bargaining: An Introduction,” *Review of Economic Studies*, vol. 53, no. 5 (October 1986), pp. 709–24. The fully rigorous (and quite difficult) theory is given in Ariel Rubinstein, “Perfect Equilibrium in a Bargaining Model,” *Econometrica*, vol. 50, no. 1 (January 1982), pp. 97–109. Return to reference 12

5 EXPERIMENTAL EVIDENCE

The theory of noncooperative bargaining is fairly simple, and many people have staged laboratory or classroom experiments that create such conditions to observe what the experimental subjects actually do. We mentioned some of these experiments briefly in Chapter 3, when considering the validity of rollback reasoning. Here we examine them in more detail in the specific context of bargaining.[13]

The simplest bargaining experiment is the ultimatum game that we introduced in Chapter 3, Section 6. In that game, there is only one round: Player A makes an offer, and if B does not accept it, the bargaining ends and both get nothing. The general structure of these experiments is as follows: A pool of players is brought together, either in the same room or via an online network. The players are paired; one person in each pair is designated the *proposer* (A, the one who makes an offer) and the other is designated the *chooser* (B, the one who accepts or refuses the offer). The pair is given a fixed surplus, usually $1 or some other sum of money, to split.

Rollback reasoning suggests that A should offer B the minimal unit—say, 1 cent out of a dollar—and that B should accept such an offer. Actual results are dramatically different. In the case in which the subjects are together in a room and the assignment of the role of proposer is made randomly, the most common offer is a 50:50 split. Very few offers worse than 75:25 are made (with the proposer to keep 75% and the chooser to get 25%), and if made, they are often rejected.

This finding can be interpreted in two ways: Either the players cannot or do not perform the calculation required for rollback, or the payoffs of the players include something other than the money they get out of this round of bargaining. Surely the calculation in the ultimatum game is simple enough that anyone should be able to do it, and the subjects in most of these experiments are college students. A more likely explanation is the one that we put forth in Chapter 3, Section 6, and Chapter 5, Section 4—that the theory, which assumed payoffs to consist only of the sum earned in the one round of bargaining, is too simplistic.

Participants can have payoffs that include things other than money. First, they may have self-esteem or pride that prevents them from accepting a very unequal split. Even if A does not include this consideration in his own payoff, if he thinks that B might, then it is a good strategy for A to offer enough to make it likely that B will accept. A balances his higher payoff with a smaller offer to B against the risk of getting nothing if B rejects an offer deemed too unequal.

A second possibility is that when the participants in the experiment are gathered in a room, the anonymity of pairing cannot be guaranteed. If the participants come from a group of people, such as college classmates or villagers, who have ongoing relationships outside this game, they may value those relationships. Then the proposers fear that if they offer too unequal a split in this game, those relationships may suffer. Therefore, they will be more generous in their offers than the simplistic theory suggests. If this is the explanation for the results, then ensuring greater anonymity should enable the proposers to make more unequal offers, and experiments do find this to be the case.

Finally, people may have a sense of fairness drilled into them during their nurture and education. This sense of fairness may have evolutionary value for society as a whole and may therefore have become a social norm. Whatever its origin, it may lead the proposers to be relatively generous in their offers, quite irrespective of the fear of their rejection. One of us (Skeath) has conducted classroom experiments with several different ultimatum games. Students who had bargaining partners previously known to them were noticeably "fairer" in their splits. In addition, several students cited specific cultural backgrounds as explanations for behavior that was inconsistent with theoretical predictions.

Experimenters have tried variants of the basic ultimatum game to differentiate between these explanations. The question of ongoing relationships can be addressed by stricter procedures that visibly guarantee anonymity. Doing so by itself has some effect on the outcomes, but still does not produce offers as extreme as those predicted by the purely selfish rollback argument of the theory. The other two explanations—namely, fear of the offer's rejection and an ingrained sense of fairness—remain to be sorted out.

The fear of rejection can be removed by considering a variant called the *dictator game.* Again, the participants are matched in pairs. One person (say, A) is designated to determine the split, and the other (say, B) is simply a passive recipient of what A decides. Now the split becomes decidedly more uneven, but even here a majority of As choose to keep no more than 70%. This result suggests a role for an ingrained sense of fairness.

But such a sense has its limits. In some experiments, a sense of fairness was created when the experimenter randomly assigned the roles of proposer and chooser. In one variant, the participants were given a simple quiz, and those who performed best were made proposers. This procedure created a sense that the role of proposer had been earned, and the outcomes showed more unequal splits. When the dictator game was played with earned dictator roles and with stricter anonymity conditions, most As kept everything, but some (about 5%) still offered a 50:50 split.

One of us (Dixit) carried out a classroom experiment in which students in groups of 20 were gathered together. They all logged into the game system and were matched randomly and anonymously in pairs, and each pair tried to agree on how to split 100 points. Roles of proposer and chooser were not assigned; either could make an offer or accept the other's offer. Offers could be made and changed at any time. The pairs could exchange messages instantly on their screens. The bargaining round ended at a random time after 3 to 5 minutes; if an agreement was not reached in that time by a pair, both got zero. There were 10 such rounds, with different random pairs of opponents each time. Thus, the game itself offered no scope for cooperation through repetition. In a classroom context, the students had ongoing relationships outside the game, but they did not generally know or guess with whom they were playing in any round, even though no great attempt was made to enforce anonymity. Each student's score for the whole game was the sum of his point score for the 10 rounds. The stakes were quite high, because the score accounted for 5% of the course grade!

The highest total of points achieved was 515. Those who quickly agreed on 50:50 splits did the best, and those who tried to hold out for very uneven scores or who refused to split a difference of 10 points or so between the offers and ran the risk of time running out on them did poorly.[14] It seems that moderation and fairness do get rewarded, even as measured in terms of one's own payoff.

Endnotes

- For more details, see Douglas D. Davis and Charles A. Holt, *Experimental Economics* (Princeton, N.J.: Princeton University Press, 1993), pp. 263–69, and *The Handbook of Experimental Economics*, ed. John H. Kagel and Alvin E. Roth (Princeton, N.J.: Princeton University Press, 1995), pp. 255–74. Return to reference 13
- Those who were best at the mathematical aspects of game theory, such as problem sets, did a little worse than the average, probably because they tried too hard to eke out an extra advantage and met resistance. And women did slightly better than men. Return to reference 14

Glossary

ultimatum game
: A form of bargaining where one player makes an offer of a particular split of the total available surplus, and the other has only the all-or-nothing choice of accepting the offer or letting the game end without agreement, when both get zero surplus.

6 MANIPULATING INFORMATION IN BARGAINING

We have seen that the outcomes of bargaining depend crucially on various characteristics of the parties to the bargain, most importantly their BATNAs and their impatience. We have proceeded thus far by assuming that the players know each other’s characteristics as well as their own. In fact, we have assumed that each player knows that the other knows, and so on; that is, that their characteristics are common knowledge. In reality, we often engage in bargaining without knowing the other side’s BATNA or degree of impatience; sometimes we do not even know our own BATNA very precisely.

As we saw in Chapter 9, a game with such uncertainty or asymmetry of information has associated with it an important game of signaling and screening. Bargaining is replete with opportunities for manipulating information using such strategies. A player with a good BATNA or a high degree of patience wants to signal this fact to the other player. However, because someone without these good attributes will want to imitate them, the other player will be skeptical and will examine the signals critically for their credibility. And each player will also try screening, using strategies that induce the other to take actions that will reveal his characteristics truthfully.

In this section, we look at some information-manipulation strategies used by buyers and sellers in the housing market. Most Americans are active in the housing market several times in their lives, and professional real-estate agents or brokers have even more extensive experience in that market. Moreover, housing is one of the few markets in the United States where haggling or bargaining over price is accepted and even expected. Therefore, considerable experience with strategies is available. We draw on this experience for many of our examples and interpret it in the light of our game-theoretic ideas and insights.[15]

When you contemplate buying a house in a new neighborhood, you are unlikely to know the range of prices for the particular type of house in which you are interested. Your first step should be to find out what this range is so that you can then determine your BATNA. And that does not mean looking at newspaper ads or realtors' listings, which indicate only asking prices. Local newspapers and some Internet sites list recent actual transactions and the actual prices; you should check them against the asking prices of the same houses to get an idea of the state of the market and the range of bargaining that might be possible.

Next comes screening: finding out the other side's BATNA and level of impatience. If you are a buyer, you can find out why the house is being sold and how long it has been on the market. If it is empty, why? And how long has it been that way? If the owners are getting divorced, or have moved elsewhere and are financing another house with an expensive bridge loan, it is likely that they have a low BATNA or are rather impatient.

You should also find out other relevant things about the other side's preferences, even though those preferences may seem irrational to you. For example, some people consider an offer too far below the asking price an insult and will not sell at any price to someone who makes such an offer. Norms of this kind vary across regions and times. It pays to find out what the common practices are.

Most importantly, the *acceptance* of an offer more accurately reveals a player's true willingness to pay than anything else and therefore is open to exploitation by the other player. A brilliant game-theorist friend of ours tried just such a ploy. He was bargaining for a floor lamp. Starting with the seller’ s asking price of $100, the negotiation proceeded to a point where our friend made an offer to buy the lamp for $60. The seller said yes, at which point our friend thought, “This guy is willing to sell it for $60, so his true rock-bottom price must be even lower. Let me try to find out whether it is.” So our friend said, “How about $55?” The seller got very upset, refused to sell for any price, and asked our friend to leave the store and never come back.

The seller’ s behavior conformed to the norm that considers it the utmost in bad faith to renege on an offer once it is accepted. This norm makes good sense in the context of all bargaining games that take place in society. If an offer on the table cannot be accepted in good faith by the other player without fear of the kind of exploitation attempted by our friend, then each bargainer will wait to get the other to accept an offer, thereby revealing his true rock-bottom acceptance level, and the whole process of bargaining will grind to a halt. Therefore, such behavior has to be disallowed. Establishing a social norm to which people adhere instinctively, as the seller in the example did, is a good way for society to achieve this aim.

An offer may explicitly state that it is open for only a specified and limited time; this stipulation can be part of the offer itself. Job offers usually specify a deadline for acceptance; stores have sales for limited periods. But in that case, the offer is truly a *package* of price and time, and reneging on either dimension violates the norms of bargaining. For example, customers get quite angry if they arrive at a store in the sale period and find an advertised item unavailable. The store must offer a rain check, which allows the customer to buy the item at its sale price when it is next available at the regular price, but even this offer causes some inconvenience to the customer and risks some loss of goodwill. The store can specify “limited quantities, no rain checks” very clearly in its advertising of the sale; even then, many customers instinctively get upset if they find that the store has run out of the item.

Next on our list of strategies to use in one-on-one bargaining, as in the housing market, comes signaling your own high BATNA or patience. The best way to signal patience is to *be* patient. Do not come back with counteroffers too quickly, “let the sellers think they’ve lost you.” This signal is credible because someone not genuinely patient would find it too costly to mimic the leisurely approach. Similarly, you can signal a high BATNA by starting to walk away, a tactic that is common in negotiations at bazaars in other countries and some flea markets and tag sales in the United States.

Even if your BATNA is low, you may commit yourself to not accepting an offer below a certain level. This constraint acts just like a high BATNA, because the other party cannot hope to get you to accept anything less. In the housing context, you can claim your inability to concede any further by inventing (or creating) a tightwad parent who is providing the down payment or a spouse who does not really like the house and will not let you offer any more. Sellers can try similar tactics. A parallel in wage negotiations is the *mandate.* A meeting is convened at which the workers pass a resolution—the mandate—authorizing the union leaders to represent them in the negotiations, but with the constraint that the negotiators must not accept an offer below a certain level specified in the resolution. Then, at the meeting with the management, the union leaders can say that their hands are tied; there is no time to go back to the membership to get their approval for any lower offer.

Most of these strategies entail some risk. While you are signaling patience by waiting, the seller of the house you want may find another willing buyer. As employer and union wait for one another to concede, tensions may mount so high that a strike that is costly to both sides cannot be prevented. In other words, many strategies of information manipulation are instances of brinkmanship. We saw in Chapter 13 how such games can have an outcome that is bad for both parties. The same is true in bargaining. *Threats* of breakdown of negotiations or of strikes are strategic moves intended to achieve quicker agreement or a better deal for the player making the move; however, an *actual* breakdown or strike is an instance of a threat gone wrong. The player making the threat—initiating the brinkmanship—must assess the risk and the potential rewards when deciding whether and how far to proceed down this path.

Endnotes

- We have taken the insights of practitioners from Andrée Brooks, “Honing Haggling Skills,” *New York Times*, December 5, 1993. Return to reference 15

7 BARGAINING WITH MANY PARTIES AND ISSUES

Our discussion thus far has been confined to the classic situation where two parties are bargaining about the split of a given total surplus. But many real-life negotiations include several parties or several issues simultaneously. Although the games get more complicated in these cases, often the enlargement of the group or the set of issues actually makes it easier to arrive at a mutually satisfactory agreement. In this section, we take a brief look at such matters.[16]

A. Multi-Issue Bargaining

In a sense, we have already considered multi-issue bargaining. The negotiation over price between a seller and a buyer always involves *two* things: (1) the object offered for sale or considered for purchase and (2) money. The potential for mutual benefit arises when the buyer values the object more than the seller does—that is, when the buyer is willing to give up more money in return for getting the object than the seller is willing to accept in return for giving up the object. Both players can be better off as a result of their bargaining agreement.

The same principle applies more generally. International trade is the classic example. Consider two hypothetical countries, Freedonia and Ilyria. If Freedonia can convert 1 loaf of bread into 2 bottles of wine (by using less of its resources, such as labor and land, in the production of bread and using those resources to produce more wine instead), and if Ilyria can convert 1 bottle of wine into 1 loaf of bread (by switching its resources in the opposite direction), then between them, they can create more goods “out of nothing.” For example, suppose that Freedonia can produce 200 more bottles of wine if it produces 100 fewer loaves of bread, and that Ilyria can produce 150 more loaves of bread if it produces 150 fewer bottles of wine. These switches in resource use create an extra 50 loaves of bread and an extra 50 bottles of wine relative to what the two countries produced originally. This extra bread and wine is the surplus that they can create if they can agree on how to divide it between them. For example, suppose Freedonia gives 175 bottles of wine to Ilyria and gets 125 loaves of bread. Then each country will have 25 more loaves of bread and 25 more bottles of wine than it did before. But there is a whole range of possible exchanges corresponding to different divisions of the gain. At one extreme, Freedonia could give up all the 200 extra bottles of wine that it has produced in exchange for 101 loaves of bread from Ilyria, in which case Ilyria would get almost all the gain from trade. At the other extreme, Freedonia could give up only 151 bottles of wine in exchange for 150 loaves of bread from Ilyria, so that Freedonia would get almost all the gain from trade.[17] Between these limits lies the frontier where the two can bargain over the division of the gains from trade.

The general principle should now be clear: When two or more issues are on the bargaining table at the same time and the two parties are willing to trade more of one thing against less of the other at different rates, then a mutually beneficial deal exists. The mutual benefit can be realized by trading at a rate somewhere between the two parties' different rates of willingness to trade. The division of gains depends on the choice of the rate of trade. The closer it is to one side's willingness ratio, the less that side gains from the deal.

Now you can also see how the possibilities for mutually beneficial deals can be expanded by bringing more issues to the table at the same time. With more issues, you are more likely to find divergences in the ratios of valuation between the two parties and are thereby more likely to locate possibilities for mutual gain. In regard to a house, for example, many of the fittings or furnishings may be of little use to the seller in the new house to which he is moving, but they may be of sufficiently good fit and taste that the buyer values having them. Then, if the seller cannot be induced to lower the price of the house, he may be amenable to including these items in the original price to close the deal.

However, the expansion of issues is not an unmixed blessing. If you value something greatly, you may fear putting it on the bargaining table; you may worry that the other side will extract big concessions from you, knowing that you want to protect that one item of great value. At the worst, a new issue on the table may make it possible for one side to deploy threats that lower the other side’s BATNA. For example, a country engaged in diplomatic negotiations may be vulnerable to an economic embargo; then it would much prefer to keep the political and economic issues distinct.

B. Multiparty Bargaining

Having many parties simultaneously engaged in bargaining may also facilitate agreement, because instead of having to look for pairwise deals, the parties can seek a circle of concessions. International trade is again the prime example. Suppose the United States can produce wheat very efficiently, but is less productive in cars; Japan is very good at producing cars, but has no oil; and Saudi Arabia has a lot of oil, but cannot grow wheat. In pairs, they can achieve little, but the three together have the potential for a mutually beneficial deal.

As with multiple issues, expanding the bargaining to multiple parties is not simple. In our example, the deal would be that the United States would send an agreed amount of wheat to Saudi Arabia, which would send an agreed amount of oil to Japan, which would in turn ship an agreed number of cars to the United States. But suppose that Japan reneges on its part of the deal. Saudi Arabia cannot retaliate against the United States because, in this scenario, it is not offering anything to the United States that it can potentially withhold. Saudi Arabia can only break its deal to send oil to Japan. Thus, enforcement of multilateral agreements may be problematic. The General Agreement on Tariffs and Trade (GATT) between 1946 and 1994, as well as the World Trade Organization (WTO) since then, have indeed found it difficult to enforce their members’ agreements and to levy punishments on countries that violate the rules.

Endnotes

- Mutually-beneficial agreements are possible whenever there are “gains from trade,” a topic typically covered in microeconomics courses; see, for example, Robert S. Pindyck and Daniel L. Rubenfeld, *Microeconomics*, 9th ed. (Upper Saddle River, NJ: Pearson, 2017), Chapter 16. For a more thorough treatment, see Howard Raiffa, *The Art and Science of Negotiation* (Cambridge, MA: Harvard University Press, 1982), Parts III and IV. Return to reference 16
- Economics uses the concept *ratio of exchange*, or price, which here is expressed as the number of bottles of wine that trade for each loaf of bread. The crucial point is that the possibility of gain for both countries exists with any ratio that lies between the 2:1 rate at which Freedonia can convert bread into wine and the 1:1 rate at which Ilyria can do so. At a ratio of exchange close to 2:1, Freedonia gives up almost all of its 200 extra bottles of wine and gets little more than the 100 loaves of bread that it sacrificed to produce the extra wine; thus Ilyria has almost all of the gain. Conversely, at a ratio close to 1:1, Freedonia has almost all of the gain. The issue in the bargaining is the division of gain, and therefore the ratio or the price at which the two should trade. Return to reference 17

SUMMARY

Bargaining negotiations attempt to divide the *surplus* (excess value) that is available to the parties if an agreement can be reached. Bargaining can be analyzed as a *cooperative* game in which parties find and implement a solution jointly or as a (structured) *noncooperative* game in which parties choose strategies separately and attempt to reach an equilibrium.

Nash’s cooperative solution is based on three principles: the outcome’s invariance with linear changes in the payoff scale, its *efficiency*, and its invariance with removal of irrelevant outcomes. The solution is a rule that states the proportions of division of surplus, beyond the backstop payoff levels (also called *BATNAs* or *best alternatives to a negotiated agreement*) available to each party, based on the parties’ relative bargaining strengths. Strategic manipulation of the backstops can be used to increase a party’s payoff.

In a noncooperative setting of *alternating offers* and counteroffers, rollback reasoning is used to find an equilibrium; this reasoning generally includes a first-round offer that is immediately accepted. If the surplus value *decays* with refusals, the sum of the (hypothetical) amounts destroyed owing to the refusals of a single player is the payoff to the other player in equilibrium. If delay in agreement is costly owing to *impatience*, the equilibrium agreement shares the surplus roughly in inverse proportion to the parties’ rates of impatience. Experimental evidence indicates that players often offer more than is necessary to reach an agreement in such games; this behavior is thought to be related to lack of player anonymity as well as beliefs about fairness.

The presence of information asymmetries in bargaining games makes signaling and screening important. Some parties will wish to signal their high BATNAs or extreme patience; others will want to screen to obtain truthful revelation of such characteristics. When more issues are on the table or more parties are participating, agreements may be easier to reach, but bargaining may be riskier or the agreements more difficult to enforce.

KEY TERMS

alternating offers (684)

best alternative to a negotiated agreement (BATNA) (675)

decay (685)

efficient (678)

efficient frontier (679)

impatience (685)

Nash cooperative solution (676)

surplus (675)

ultimatum game (692)

variable-threat bargaining (681)

Glossary

alternating offers

A sequential move procedure of bargaining in which, if the offer made by one player is refused by the other, then the refuser gets the next turn to make an offer, and so on.

best alternative to a negotiated agreement (BATNA)

In a bargaining game, this is the payoff a player would get from his other opportunities if the bargaining in question failed to reach an agreement.

decay

Shrinkage over time of the total surplus available to be split between the bargainers, if they fail to reach an agreement for some length of time during the process of their bargaining.

efficient

An outcome of a bargaining game is called efficient if there is no feasible alternative that would leave one bargainer with a higher payoff without reducing the payoff of the other.

efficient frontier

This is the northeast boundary of the set of feasible payoffs of the players, such that in a bargaining game it is not possible to increase the payoff of one person without lowering that of another.

impatience

Preference for receiving payoffs earlier rather than later. Quantitatively measured by the discount factor.

Nash cooperative solution

This outcome splits the bargainers' surpluses in proportion to their bargaining powers.

ultimatum game

A form of bargaining where one player makes an offer of a particular split of the total available surplus, and the other has only the all-or-nothing choice of accepting the offer or letting the game end without agreement, when both get zero surplus.

variable-threat bargaining

A two-stage game where at the first stage you can take an action that will alter the BATNAs of both bargainers (within certain limits), and at the second stage bargaining results in the Nash solution on the basis of these BATNAs.

surplus

A player's surplus in a bargaining game is the excess of his payoff over his BATNA.

SOLVED EXERCISES

1. Ali and Baba are bargaining to split a total that starts out at $100. Ali makes the first offer, stating how the $100 will be divided between them. If Baba accepts this offer, the game is over. If Baba rejects it, a dollar is withdrawn from the total, so it is now only $99. Then Baba gets the second turn to make an offer of a division. The turns alternate in this way, a dollar being removed from the total after each rejection. Ali’s BATNA is $2.25 and Baba’s BATNA is $3.50. What is the rollback equilibrium outcome of the game?
2. William, a worker at Acme Enterprises, is responsible for managing Acme’s relationship with Must Buy, an important supplier. William has learned how to work well with Must Buy, allowing Acme to get a benefit of $280,000 per year from the relationship (not counting William’s salary), whereas Acme would only get only $100,000 per year from the relationship if William left and someone else had to take his role. Currently, everyone at Acme (including William) earns a salary of $100,000 per year, but William knows that he’s worth a lot more to the company. On the other hand, William also knows that his boss has three times as much bargaining power as he does.
   1. One day, William drops by his boss’s office and demands to renegotiate his salary. His boss counters, “That’s fine. But you should understand that if we can’t reach an agreement, you’re gone—and I’ll make sure that you never work again!” Should William back down (and keep his current salary) or press forward with the negotiation? If he presses forward, what does Nash’s theory of bargaining predict will happen? Assume that the boss’s threat “You’ll never work again” is credible.

2. What if the boss’s threat isn’t credible? Yes, William will lose his job if the negotiation fails, but if so, William will be able to find another job. However, this new job will have a lower salary of $80,000 per year. What is the Nash cooperative solution now?
3. Before going into the boss’s office, William realizes that he has even more to offer the company. If his boss agrees to give him a raise, William can teach his coworker, Jill, how to increase the profitability of her supplier relationship by $60,000 per year. However, if the negotiation fails, William will still get fired and have to take that $80,000 per year job. What is the Nash cooperative solution now?

3. Two hypothetical countries, Euphoria and Militia, are holding negotiations to settle a dispute. They meet once a month, starting in January, and take turns making offers. Suppose the total at stake is 100 points. The government of Euphoria is facing reelection in November. Unless that government produces an agreement at the October meeting, it will lose the election, an outcome it regards as being just as bad as getting zero points from an agreement. The government of Militia does not really care about reaching an agreement; it is just as happy to prolong the negotiations, or even to fight, rather than settle for anything significantly less than 100.
   1. What will be the outcome of the negotiations? What difference will the identity of the first mover make?
   2. In light of your answer to part (a), discuss why actual negotiations often continue right down to the deadline.
4. In 1974, the U.S. Congress passed the Budget Act, creating the House Budget Committee and requiring that all budget resolutions pass through the Committee before coming to the House floor for a vote. This act gives the chair of the Budget Committee great negotiating power,

since the chair controls what the committee can consider and the committee controls what Congress can consider. Actual budget negotiations are very complex, but for our purposes here, suppose that the committee is considering a budget with total expenditures of $4 trillion, of which $3.88 trillion are for “must-fund programs” that will be funded even if Congress fails to reach an agreement. This leaves $120 billion per year that can be allocated to “pork” projects that, while not essential, benefit voters in the places where those projects are located. From this perspective, the substance of budget negotiations is to determine where the pork flows.

Suppose that each round of this negotiation takes one month, with the chair deciding how much pork to give to his own district (or to allies who will now owe him a favor) and how much to give to others. If the first offer is rejected, one month’s worth of pork ($120/12 = $10 billion) will be lost and the whole process will start over, with the chair making another offer of how to divide the remaining $110 billion, and so on until $10 billion remains in the final month. What is the rollback equilibrium outcome of the game?

UNSOLVED EXERCISES

1. Recall the variant of the pizza pricing game in Exercise U3, part (b), in Chapter 10, in which one store (Donna’s Deep Dish) was much larger than the other (Pierce’s Pizza Pies). The payoff table for that version of the game is

		PIERCE’S PIZZA PIES	
		High	Low
DONNA’S DEEP DISH	High	156, 60	132, 70
	Low	150, 36	130, 50

The noncooperative Nash equilibrium is (High, Low), yielding profits of 132 to Donna’s and 70 to Pierce’s, for a total of 202. If the two could achieve (High, High), their total profit would be 156 + 60 = 216, but Pierce’s would not agree to this pricing.

Suppose the two stores can reach an enforceable agreement whereby both charge High and Donna’s pays Pierce’s a sum of money. The alternative to this agreement is simply the noncooperative Nash equilibrium. Donna’s has 2.5 times as much bargaining power as Pierce’s. In the resulting agreement, what sum will Donna’s pay to Pierce’s?

2. This exercise considers a variation of William’s negotiation game (from Exercise S2). As before, William’s current salary is $100,000 per year, he creates $280,000 per year value for his boss at Acme Enterprises, and his boss has three times as much bargaining strength as William. Also, as in part (a) of

Exercise S2, William initially believes that he has no hope of ever getting another job. But now William gets a call from another company, Wily Industries, inquiring whether William would like to work for them. William won't be quite as productive at Wily, creating $240,000 of value there rather than his $280,000 at Acme; also, like his current boss at Acme, the boss at Wily has three times as much bargaining strength as William. The question then arises: Does it matter who William negotiates with first?

1. Suppose that William first negotiates with Wily Industries and that the Wily negotiation fails. William then has the option to negotiate with his boss at Acme. Verify that, in this scenario, William will choose not to negotiate for a higher wage with Acme and hence will continue earning $100,000 per year. [Feel free to consult the solution to Exercise S2, part (a).]
2. Given your finding in (a), what is the predicted Nash cooperative solution of this sequential-move bargaining game if William negotiates first with Wily Industries? Will William switch jobs, and what will his new wage be?
3. What if William first negotiates with his boss at Acme, explaining that he has received interest from Wily but has not yet entered negotiations with them? As in Exercise S2, assume that William loses his job if the negotiation with Acme fails. The difference now is that if William is tossed out of Acme, he can then negotiate with Wily. Verify that in this scenario, the predicted Nash cooperative outcome of negotiation between William and Wily will be for William to accept an offer at Wily at a salary of $60,000 per year.
4. Given your finding in (c), what is the predicted Nash cooperative outcome of this sequential-move bargaining game if William negotiates first with Acme

Enterprises, his current employer? Will William switch jobs, and what will his new wage be?
5. Which is better for William: to negotiate first with Acme or to negotiate first with Wily?

3. Consider two players who bargain over a surplus initially equal to a whole-number amount V, using alternating offers. That is, Player 1 makes an offer in round 1; if Player 2 rejects this offer, she makes an offer in round 2; if Player 1 rejects this offer, she makes an offer in round 3; and so on. Suppose that the available surplus decays by a constant value of $c = 1$ each period. For example, if the players reach agreement in round 2, they divide a surplus of $V - 1$; if they reach agreement in round 5, they divide a surplus of $V - 4$. This means that the game will be over after V rounds, because at that point there will be nothing left to bargain over. (For comparison, remember the football ticket example in Section 3, in which the value of the ticket to the fan started at \$100 and declined by \$25 per quarter over the four quarters of the game.) In this problem, we first solve for the rollback equilibrium of this game, and then solve for the equilibrium of a generalized version of this game in which the two players can have BATNAs.
   1. Let's start with a simple version. What is the rollback equilibrium when $V = 4$? In which period will the players reach agreement? What payoff x will Player 1 receive, and what payoff y will Player 2 receive?
   2. What is the rollback equilibrium when $V = 5$?
   3. What is the rollback equilibrium when $V = 10$?
   4. What is the rollback equilibrium when $V = 11$?
   5. Now we're ready to generalize. What is the rollback equilibrium for any whole-number value of V? (Hint: You may want to consider even values of V separately from odd values.)

Now consider BATNAs. Suppose that if no agreement is reached by the end of round *V*, Player A gets a payoff of *a* and Player B gets a payoff of *b*. Assume that *a* and *b* are whole numbers satisfying the inequality $a + b < V$, so that the players can get higher payoffs by reaching agreement than they can by not reaching agreement.

1. (f) Suppose that $V = 4$. What is the rollback equilibrium for any possible values of *a* and *b*? [Hint: You may need to write down more than one formula, just as you did in part (e). If you get stuck, try assuming specific values for *a* and *b*, and then change those values to see what happens. In order to roll back, you'll need to figure out the turn at which the value of *V* has declined to the point where a negotiated agreement would no longer be profitable for the two bargainers.]
2. (g) Suppose that $V = 5$. What is the rollback equilibrium for any possible values of *a* and *b*?
3. (h) For any whole-number values of *a*, *b*, and *V*, what is the rollback equilibrium?
4. (i) Relax the assumption that *a*, *b*, and *V* are whole numbers: Let them be any nonnegative numbers such that $a + b < V$. Also relax the assumption that the value of *V* decays by exactly 1 each period: Let the value decay each period by some constant amount $c > 0$. What is the rollback equilibrium to this general problem?

4. Let *x* be the amount that Player A asks for, and let *y* be the amount that Player B asks for, when making the first offer in an alternating-offers bargaining game with impatience. Their rates of impatience are *r* and *s*, respectively.
   1. If we use the approximate formulas $x = s/(r + s)$ for *x* and $y = r/(r + s)$ for *y*, and if B is twice as impatient as A, then A gets two-thirds of the surplus

and B gets one-third. Verify that this result is correct.

2. Let $r = 0.01$ and $s = 0.02$, and compare the x and y values found by using the approximation method with the more exact solutions for x and y found by using the formulas $x = (s + rs)/(r + s + rs)$ and $y = (r + rs)/(r + s + rs)$ derived in the chapter.

Glossary

■

Here we define the key terms that appear in the text. We aim for verbal definitions that are logically precise but not mathematical or detailed like those found in more advanced textbooks.

acceptability condition An upper bound on the probability of fulfillment in a brinkmanship threat, expressed as a function of the probability of error, showing the upper limit of risk that the player making the threat is willing to tolerate.

action node A node at which one player chooses an action from two or more that are available.

addition rule If the occurrence of *X* requires the occurrence of *any one* of several disjoint *Y*, *Z*, . . . , then the probability of *X* is the sum of the separate probabilities of *Y*, *Z*, . . .

adverse selection A form of information asymmetry where a player's type (available strategies, payoffs . . .) is his private information, not directly known to others.

affirmation A response to the *favored action* that is a best response, as specified within a strategic move. Only the second mover in a game can feasibly make an affirmation. However, credibility is not required since the specified action is already the player's best response. The strategic move referred to as "making a threat" entails declaring both a threat and an affirmation.

agenda paradox A voting situation where the order in which alternatives are paired when voting in multiple rounds determines the final outcome.

agent The agent is the more-informed player in a principal–agent game of asymmetric information. The principal (less-informed) player in such games attempts to design a mechanism that aligns the agent’s incentives with his own.

all-pay auction An auction in which each person who submits a bid must pay her highest bid amount at the end of the auction, even if she does not win the auction.

alternating offers A sequential move procedure of bargaining in which, if the offer made by one player is refused by the other, then the refuser gets the next turn to make an offer, and so on.

amendment procedure A procedure in which any amended version of a proposal must win a vote against the original version before the winning version is put to a vote against the status quo.

antiplurality method A positional voting method in which the electorate is asked to vote against one item on the slate (or to vote for all but one).

approval voting A voting method in which voters cast votes for all alternatives of which they approve.

ascending-price auction An open-outcry auction in which prices are announced in increasing order either by an auctioneer (in the case of an English auction) or by bidders themselves (in the case of jump bidding). The last person to bid or accept the announced price wins the auction and pays that price.

assurance game A game where each player has two strategies, say, Cooperate and Not, such that the best response of each is to Cooperate if the other cooperates, Not if not, and the outcome from (Cooperate, Cooperate) is better for both than the outcome of (Not, Not).

asymmetric information Information is said to be asymmetric in a game if some aspects of it—rules about what actions are permitted and the order of moves if any, payoffs as functions of the players strategies, outcomes of random choices by "nature," and of previous actions by the actual players in the game—are known to some of the players but are not common knowledge among all players.

auction A game in which multiple players (called bidders) compete for a scarce resource.

auction designer A player who sets the rules of an auction game.

babbling equilibrium In a game where communication among players (which does not affect their payoffs directly) is followed by their choices of actual strategies, a babbling equilibrium is one where the strategies are chosen ignoring the communication, and the communication at the first stage can be arbitrary.

backward induction Same as **rollback**.

battle of the sexes A game where each player has two strategies, say, Hard and Soft, such that [1] (Hard, Soft) and (Soft, Hard) are both Nash equilibria, [2] of the two Nash equilibria, each player prefers the outcome where he is Hard and the other is Soft, and [3] both prefer the Nash equilibria to the other two possibilities, (Hard, Hard) and (Soft, Soft).

Bayesian Nash equilibrium A Nash equilibrium in an asymmetric information game where players use Bayes' theorem and draw correct inferences from their observations of other players' actions.

Bayes' theorem An algebraic formula for estimating the probabilities of some underlying event by using knowledge of some consequences of it that are observed.

belief The notion held by one player about the strategy choices of the other players and used when choosing his own optimal strategy.

best alternative to a negotiated agreement (BATNA) In a bargaining game, this is the payoff a player would get from his other opportunities if the bargaining in question failed to reach an agreement.

best response The strategy that is optimal for one player, given the strategies actually played by the other players, or the belief of this player about the other players' strategy choices.

best-response analysis Finding the Nash equilibria of a game by calculating the best-response functions or curves of each player and solving them simultaneously for the strategies of all the players.

best-response curve A graph showing the best strategy of one player as a function of the strategies of the other player(s) over the entire range of those strategies.

best-response rule A function expressing the strategy that is optimal for one player, for each of the strategy combinations actually played by the other players, or the belief of this player about the other players' strategy choices.

bidder A player in an auction game.

binary method A class of voting methods in which voters choose between only two alternatives at a time.

Black's condition Same as the condition of **single-peaked preferences**.

Borda count A positional voting method in which the electorate indicates its order of preference over a slate of alternatives. The winning alternative is determined by allocating points based on an alternative’s position on each ballot.

branch Each branch emerging from a node in a game tree represents one action that can be taken at that node.

brinkmanship A threat that creates a risk but not certainty of a mutually bad outcome if the other player defies your specified wish as to how he should act, and then gradually increases this risk until one player gives in or the bad outcome happens.

cheap talk equilibrium In a game where communication among players (which does not affect their payoffs directly) is followed by their choices of actual strategies, a cheap talk equilibrium is one where the strategies are chosen optimally given the players’ interpretation of the communication, and the communication at the first stage is optimally chosen by calculating the actions that will ensue.

chicken A game where each player has two strategies, say Tough and Weak, such that [1] both (Tough, Weak) and (Weak, Tough) are Nash equilibria, [2] of the two, each prefers the outcome where she plays Tough and the other plays Weak, and [3] the outcome (Tough, Tough) is worst for both.

coercion In this context, forcing a player to accept a lower payoff in an asymmetric equilibrium in a collective action game, while other favored players are enjoying higher payoffs. Also called **oppression** in this context.

collective action A problem of achieving an outcome that is best for society as a whole, when the interests of some or all individuals will lead them to a different outcome as the equilibrium of a noncooperative game.

combinatorial auction An auction of multiple dissimilar objects in which bidders are able to bid on and win combinations of objects.

commitment An action taken at a pregame stage, stating what action you would take unconditionally in the game to follow.

common value An auction is called a common-value auction when the object up for sale has the same value to all bidders, but each bidder knows only an imprecise estimate of that value.

compellence An attempt to induce the other player(s) to act to change the status quo in a specified manner.

compound interest When an investment goes on for more than one period, compound interest entails calculating interest in any one period on the whole accumulation up to that point, including not only the principal initially invested but also the interest earned in all previous periods, which itself involves compounding over the period previous to that.

Condorcet method A voting method in which the winning alternative must beat each of the other alternatives in a round-robin of pairwise contests.

Condorcet paradox Even if all individual voter preference orderings are transitive, there is no guarantee that the social preference ordering generated by Condorcet's voting method will also be transitive.

Condorcet terms A set of ballots that would generate the Condorcet paradox and that should together logically produce a tied vote among three possible alternatives. In a three-candidate election among A, B, and C, the Condorcet terms are three ballots that show A preferred to B preferred to C; B preferred to C preferred to A; C preferred to A preferred to B.

Condorcet winner The alternative that wins an election run using the *Condorcet method*.

constant-sum game A game in which the sum of all players' payoffs is a constant, the same for all their strategy combinations. Thus, there is a strict conflict of interests among the players—a higher payoff to one must mean a lower payoff to the collectivity of all the other players. If the payoff scales can be adjusted to make this constant equal to zero, then we have a *zero-sum game*.

contingent strategy In repeated play, a plan of action that depends on other players' actions in previous plays. (This is implicit in the definition of a strategy; the adjective "contingent" merely reminds and emphasizes.)

continuation The continuation of a strategy from a (noninitial) node is the remaining part of the plan of action of that strategy, applicable to the subgame that starts at this node.

continuous distribution A probability distribution in which the random variables may take on a continuous range of values.

continuous strategy A choice over a continuous range of real numbers available to a player.

contract In this context, a way of achieving credibility for one's strategic move by entering into a legal obligation to perform the committed, threatened, or promised action in the specified contingency.

convention A mode of behavior that finds automatic acceptance as a focal point, because it is in each individual's interest to follow it when others are expected to follow it too (so the game is of the assurance type). Also called **custom**.

convergence of expectations A situation where the players in a noncooperative game can develop a common understanding of the strategies they expect will be chosen.

cooperative game A game in which the players decide and implement their strategy choices jointly, or where joint-action agreements are directly and collectively enforced.

coordination game A game with multiple Nash equilibria, where the players are unanimous about the relative merits of the equilibria, and prefer any equilibrium to any of the nonequilibrium possibilities. Their actions must somehow be coordinated to achieve the preferred equilibrium as the outcome.

Copeland index An index measuring an alternative's record in a round-robin of contests where different numbers of points are allocated for wins, ties, and losses.

credibility A strategy is credible if its continuation at all nodes, on or off the equilibrium path, is optimal for the subgame that starts at that node.

credibility device A means by which a player acquires credibility, for instance, when declaring a promise or threat as part of a strategic move.

custom Same as **convention**.

decay Shrinkage over time of the total surplus available to be split between the bargainers, if they fail to reach an agreement for some length of time during the process of their bargaining.

decision node A decision node in a decision or game tree represents a point in a game where an action is taken.

decision tree Representation of a sequential decision problem facing one person, shown using nodes, branches, terminal nodes, and their associated payoffs.

default action In the context of strategic moves, the action that the other player (the player not making a strategic move) will take in the absence of a strategic move, as opposed to the *favored action*.

descending-price auction An open-outcry auction in which the auctioneer announces possible prices in descending order. The first person to accept the announced price wins the auction and pays that price. Also called **Dutch auction**.

deterrence An attempt to induce the other player(s) to act to maintain the status quo.

diffusion of responsibility A situation where action by one or a few members of a large group would suffice to bring about an outcome that all regard as desirable, but each thinks it is someone else's responsibility to take this action.

discount factor In a repeated game, the fraction by which the next period's payoffs are multiplied to make them comparable with this period's payoffs.

discrete distribution A probability distribution in which the random variables may take on only a discrete set of values such as integers.

disjoint Events are said to be disjoint if two or more of them cannot occur simultaneously.

distribution function A function that indicates the probability that a variable takes on a value less than or equal to some number.

dominance solvable A game where iterated elimination of dominated strategies leaves a unique outcome, or just one strategy for each player.

dominant strategy A strategy X is dominant for a player if the outcome when playing X is always better than the outcome when playing any other strategy, no matter what strategies other players adopt.

dominated strategy A strategy X is dominated by another strategy Y for a player if the outcome when playing X is always worse than the outcome when playing Y, no matter what strategies other players adopt.

doomsday device An automaton that will under specified circumstances generate an outcome that is very bad for all players. Used for giving credibility to a severe threat.

drop-out price In an English auction, the price at which a bidder drops out of the bidding.

Dutch auction Same as a **descending-price auction**.

dynamic chicken A game of chicken in which the choice to play Weak may be made at any time, the game ends as soon as either player chooses Weak, and the risk of the mutually worst outcome increases gradually over time if neither player has played Weak; a special case of the *war of attrition*.

effectiveness condition A lower bound on the probability of fulfillment in a brinkmanship threat, expressed as a function of the probability of error, showing the lower limit of risk that will induce the threatened player to comply with the wishes of the threatener.

effective rate of return Rate of return corrected for the probability of noncontinuation of an investment to the next period.

efficiency wage A higher-than-market wage paid to a worker as a means of incentivizing him to exert effort. If the worker shirks and is detected, he will be fired and will have to get a lower-wage job in the general labor market.

efficient An outcome of a bargaining game is called efficient if there is no feasible alternative that would leave one bargainer with a higher payoff without reducing the payoff of the other.

efficient frontier This is the northeast boundary of the set of feasible payoffs of the players, such that in a bargaining game it is not possible to increase the payoff of one person without lowering that of another.

English auction A type of *ascending-price auction* in which the auctioneer calls out a sequence of increasing prices, bidders decide when to drop out of the bidding, and the last bidder remaining pays the last announced price.

equilibrium A configuration of strategies where each player's strategy is his best response to the strategies of all the other players.

equilibrium path of play The *path of play* actually followed when players choose their rollback equilibrium strategies in a sequential game.

evolutionary game A situation where the strategy of each player in a population is fixed genetically, and strategies that yield higher payoffs in random matches with others from the same population reproduce faster than those with lower payoffs.

evolutionary stability A population is evolutionarily stable if it cannot be successfully invaded by a new mutant phenotype.

evolutionarily stable strategy (ESS) A phenotype or strategy that can persist in a population, in the sense that all the members of a population or species are of that type; the population is evolutionarily stable (static criterion). Or, starting from an arbitrary distribution of phenotypes in the population, the process of selection will converge to this strategy (dynamic criterion).

expected payoff The probability-weighted average (statistical mean or expectation) of the payoffs of one player in a game, corresponding to all possible realizations of a random choice of nature or mixed strategies of the players.

expected utility The probability-weighted average (statistical mean or expectation) of the utility of a person, corresponding to all possible realizations of a random choice of nature or mixed strategies of the players in a game.

expected value The probability-weighted average of the outcomes of a random variable, that is, its statistical mean or expectation.

extensive form Representation of a game by a game tree.

external effect When one person's action alters the payoff of another person or persons. The effect or spillover is *positive* if one's action raises others' payoffs (for example, network effects) and *negative* if it lowers others' payoffs (for example, pollution or congestion). Also called **externality** or **spillover effect**.

externality Same as **external effect**.

external uncertainty A player's uncertainty about external circumstances such as the weather or product quality.

favored action The action that a player making a *strategic move* wants the other player to take, as opposed to the *default action*.

feasibility Possibility within the (unchangeable) physical and/or procedural restrictions that apply in a given game.

first-mover advantage This exists in a game if, considering a hypothetical choice between moving first and moving second, a player would choose the former.

first-price auction A sealed-bid auction in which the highest bidder wins and pays the amount of her bid.

fitness The expected payoff of a phenotype in its games against randomly chosen opponents from the population.

focal point A configuration of strategies for the players in a game, which emerges as the outcome because of the convergence of the players' expectations on it.

free rider A player in a collective-action game who intends to benefit from the positive externality generated by others' efforts without contributing any effort of his own.

game (game of strategy) An action situation where there are two or more mutually aware players, and the outcome for each depends on the actions of all.

game table A spreadsheetlike table whose dimension equals the number of players in the game; the strategies available to each player are arrayed along one of the dimensions (row, column, page, . . .); and each cell shows the payoffs of all the players in a specified order, corresponding to the configuration of strategies that yield that cell. Also called **payoff table**.

game tree Representation of a game in the form of nodes, branches, and terminal nodes and their associated payoffs.

genotype A gene or a complex of genes, which give rise to a phenotype and which can breed true from one generation to another. (In social or economic games, the process of breeding can be interpreted in the more general sense of teaching or learning.)

Gibbard–Satterthwaite theorem With three or more alternatives to consider, the only voting method that prevents strategic voting is dictatorship; one person is identified as the dictator and her preferences determine the outcome.

grim strategy A strategy of noncooperation forever in the future, if the opponent is found to have cheated even once. Used as a threat of punishment in an attempt to sustain cooperation.

hawk–dove game An evolutionary game where members of the same species or population can breed to follow one of two strategies, Hawk and Dove, and depending on the payoffs, the game between a pair of randomly chosen members can be either a prisoners’ dilemma or chicken.

histogram A bar chart; data are illustrated by way of bars of a given height (or length).

impatience Preference for receiving payoffs earlier rather than later. Quantitatively measured by the discount factor.

imperfect information A game is said to have perfect information if each player, at each point where it is his turn to act, knows the full history of the game up to that point, including the results of any random actions taken by nature or previous actions of other players in the game, including pure actions as well as the actual outcomes of any mixed strategies they may play. Otherwise, the game is said to have imperfect information.

impossibility theorem A theorem that indicates that no preference aggregation method can satisfy the six critical principles identified by Kenneth Arrow.

incentive-compatibility condition (constraint) A constraint on an incentive scheme or screening device that makes it optimal for the agent (more-informed player) of each type to reveal his true type through his actions.

incentive design The process that a *principal* uses to devise the best possible incentive scheme (or mechanism) in a *principal–agent problem* to motivate the agent to take actions that benefit the principal. By design, such incentive schemes take into account that the agent knows something (about the world or about herself) that the principal does not know. Also called **mechanism design**.

independent events Events Y and Z are independent if the actual occurrence of one does not change the probability of the other occurring. That is, the conditional probability of Y occurring given that Z has occurred is the same as the ordinary or unconditional probability of Y.

infinite horizon A repeated decision or game situation that has no definite end at a fixed finite time.

information set A set of nodes among which a player is unable to distinguish when taking an action. Thus, his strategies are restricted by the condition that he should choose the same action at all points of an information set. For this, it is essential that all the nodes in an information set have the same player designated to act, with the same number and similarly labeled branches emanating from each of these nodes.

initial node The starting point of a sequential-move game. (Also called the **root** of the tree.)

instant-runoff voting (IRV) Same as **single transferable vote**.

intermediate valuation function A rule assigning payoffs to nonterminal nodes in a game. In many complex games, this must be based on knowledge or experience of playing similar games, instead of explicit rollback analysis.

internalize the externality To offer an individual a reward for the external benefits he conveys on the rest of society, or to inflict a penalty for the external costs he imposes on the rest, so as to bring his private incentives in line with social optimality.

intransitive ordering A preference ordering that cycles and is not *transitive*. For example, a preference ordering over three alternatives A, B, and C is intransitive if A is preferred to B and B is preferred to C but it is not true that A is preferred to C.

invasion The appearance of a small proportion of mutants in the population.

irreversible Cannot be undone by a later action. In a sequential-move game, the first mover's action must be irreversible and *observable* before the second mover's action is irreversible.

iterated elimination of dominated strategies Considering the players in turns and repeating the process in rotation, eliminating all strategies that are dominated for one at a time, and continuing doing so until no such further elimination is possible. Also called **successive elimination of dominated strategies**.

jump bidding Submitting a bid that is significantly higher than the previous bid and well beyond whatever minimum bid increment exists.

leadership In a prisoners' dilemma with asymmetric players, this is a situation where a large player chooses to cooperate even though he knows that the smaller players will cheat.

locked in A situation where the players persist in a Nash equilibrium that is worse for everyone than another Nash equilibrium.

majority rule A voting method in which the winning alternative is the one that garners a majority (more than 50%) of the votes.

majority runoff A two-stage voting method in which a second round of voting ensues if no alternative receives a majority in the first round. The top two vote-getters are paired in the second round of voting to determine a winner.

marginal private gain The change in an individual's own payoff as a result of a small change in a continuous-strategy variable that is at his disposal.

marginal social gain The change in the aggregate social payoff as a result of a small change in a continuous-strategy variable chosen by one player.

mechanism design Same as **incentive design**.

median voter The voter in the middle—at the 50th percentile—of a distribution.

median voter theorem If the political spectrum is one-dimensional and every voter has single-peaked preferences, then [1] the policy most preferred by the median voter will be the Condorcet winner, and [2] power-seeking politicians in a two-candidate election will choose platforms that converge to the position most preferred by the median voter. (This is also known as the **principle of minimum differentiation.**)

mixed method A multistage voting method that uses plurative and binary votes in different rounds.

mixed strategy A mixed strategy for a player consists of a random choice, to be made with specified probabilities, from his originally specified pure strategies.

monomorphism All members of a given species or population exhibit the same behavior pattern.

moral hazard A situation of information asymmetry where one player's actions are not directly observable to others.

move An action at one node of a game tree.

multiplication rule If the occurrence of *X* requires the simultaneous occurrence of *all* the several independent *Y, Z,* . . . , then the probability of *X* is the *product* of the separate probabilities of *Y, Z,* . . .

multistage procedure A voting procedure in which there are multiple rounds of voting.

multi-unit auction An auction in which multiple identical objects are sold.

mutation Emergence of a new genotype.

Nash cooperative solution This outcome splits the bargainers' surpluses in proportion to their bargaining powers.

Nash equilibrium A configuration of strategies (one for each player) such that each player's strategy is best for him, given those of the other players. (Can be in pure or mixed strategies.)

negatively correlated Two random variables are said to be negatively correlated if, as a matter of probabilistic average, when one is above its expected value, the other is below its expected value.

never a best response A strategy is never a best response for a player if, for each list of strategies that the other players choose (or for each list of strategies that this player believes the others are choosing), some other strategy is this player's best response. (The other strategy can be different for different lists of strategies of the other players.)

node This is a point from which branches emerge, or where a branch terminates, in a decision or game tree.

noncooperative game A game where each player chooses and implements his action individually, without any joint-action agreements directly enforced by other players.

nonexcludable Benefits that are available to each individual, regardless of whether he has paid the costs that are necessary to secure the benefits.

nonrival Benefits whose enjoyment by one person does not detract anything from another person's enjoyment of the same benefits.

norm A pattern of behavior that is established in society by a process of education or culture, to the point that a person who behaves differently experiences a negative psychic payoff.

normal distribution A commonly used statistical distribution for which the *distribution function* looks like a bell-shaped curve.

normal form Representation of a game in a game matrix, showing the strategies (which may be numerous and complicated if the game has several moves) available to each player along a separate dimension (row, column, etc.) of the matrix and the outcomes and payoffs in the multidimensional cells. Also called **strategic form**.

objective value An auction is called an objective-value auction when the object up for sale has the same value to all bidders and each bidder knows that value.

observable Known to other players before they make their responding actions. Together with irreversibility, this is an important condition for a game to be sequential-move.

off-equilibrium path A path of play that does not result from the players' choices of strategies in a subgame-perfect equilibrium.

off-equilibrium subgame A subgame starting at a node that does not lie on the equilibrium path of play.

open-outcry auction An auction mechanism in which bids are made openly for all to hear or see.

opponent's indifference property An equilibrium mixed strategy of one player in a two-person game has to be such that the other player is indifferent among all the pure strategies that are actually used in her mixture.

oppression In this context, same as **coercion**.

ordinal payoffs Each player's ranking of the possible outcomes in a game.

pairwise voting A voting method in which only two alternatives are considered at the same time.

participation condition (constraint) A constraint on an incentive scheme or a screening device that should give the more-informed player an expected payoff at least as high as he can get outside this relationship.

path of play A route through the game tree (linking a succession of nodes and branches) that results from a configuration of strategies for the players that are within the rules of the game. (See also *equilibrium path of play*.)

payoff The objective, usually numerical, that a player in a game aims to maximize.

payoff matrix Same as **payoff table** and **game table**.

payoff table Same as **game table**.

penalty We reserve this term for one-time costs (such as fines) introduced into a game to induce the players to take actions that are in their joint interests.

penny auction An auction format in which each bidder may pay a bidding fee (say, 60 cents) to advance the price by one cent. The auction continues until no one is willing to advance the price any longer, at which point the last bidder wins and pays the final price.

perfect Bayesian equilibrium (PBE) An equilibrium where each player's strategy is optimal at all nodes given his beliefs, and beliefs at each node are updated using Bayes' rule in the light of the information available at that point including other players' past actions.

perfect information A game is said to have perfect information if players face neither strategic nor external uncertainty.

phenotype A specific behavior or strategy, determined by one or more genes. (In social or economic games, this can be interpreted more generally as a customary strategy or a rule of thumb.)

playing the field A many-player evolutionary game where all animals in the group are playing simultaneously, instead of being matched in pairs for two-player games.

pluralistic ignorance A situation of collective action where no individual knows for sure what action is needed, so everyone takes the cue from other people’s actions or inaction, possibly resulting in persistence of wrong choices.

plurality rule A voting method in which two or more alternatives are considered simultaneously and the winning alternative is the one that garners the largest number of votes; the winner needs only gain more votes than each of the other alternatives and does not need 50% of the vote as would be true in *majority rule*.

plurative method Any voting method that allows voters to consider a slate of three or more alternatives simultaneously.

polymorphism An evolutionarily stable equilibrium in which different behavior forms or phenotypes are exhibited by subsets of members of an otherwise identical population.

pooling Same as **pooling of types**.

pooling equilibrium A perfect Bayesian equilibrium in a game of asymmetric information, where the actions in the equilibrium cannot be used to distinguish type.

pooling of types An outcome of a signaling or screening game in which different types follow the same strategy and get the same payoffs, so types cannot be distinguished by observing actions.

positional method A voting method that determines the identity of the winning alternative using information on the position of alternatives on a voter’s ballot to assign points used when tallying ballots.

positive feedback When one person’s action increases the payoff of another person or persons taking the same action, thus increasing their incentive to take that action too.

positively correlated Two random variables are said to be positively correlated if, as a matter of probabilistic average, when one is above its expected value, the other is also above its expected value, and vice versa.

present value (PV) The total payoff over time, calculated by summing the payoffs at different periods each multiplied by the appropriate discount factor to make them all comparable with the initial period’s payoffs.

price discrimination Perfect, or first-degree, price discrimination occurs when a firm charges each customer an individualized price based on willingness to pay. In general, price discrimination refers to situations in which a firm charges different prices to different customers for the same product.

primary criterion Comparison of the fitness of a mutant with that of a member of the dominant population, when each plays against a member of the dominant population.

principal The principal is the less-informed player in a principal–agent game of asymmetric information. The principal in such games wants to design a mechanism that creates incentives for the more-informed player (agent) to take actions beneficial to the principal.

principal–agent (agency) problem A situation in which the less-informed player (principal) wants to design a mechanism that creates incentives for the more-informed player (agent) to take actions beneficial to himself (the principal).

principle of minimum differentiation Same as part [2] of the *median voter theorem*.

prisoners' dilemma A game where each player has two strategies, say Cooperate and Defect, such that [1] for each player, Defect dominates Cooperate, and [2] the outcome (Defect, Defect) is worse for both than the outcome (Cooperate, Cooperate).

private information Information known by only one player.

private value An auction is called a private-value auction when each bidder has private information about their own valuation of the object up for sale, but knowing others' private information would not change any bidder's own willingness to pay for the object. An important special case is when each bidder knows their own valuation but others do not.

probabilistic threat A strategic move in the nature of a threat, but with the added qualification that if the event triggering the threat (the opponent's action in the case of deterrence or inaction in the case of compellence) comes about, a chance mechanism is set in motion, and if its outcome so dictates, the threatened action is carried out. The nature of this mechanism and the probability with which it will call for the threatened action must both constitute prior commitments.

probability The probability of a random event is a quantitative measure of the likelihood of its occurrence. For events that can be observed in repeated trials, it is the long-run frequency with which it occurs. For unique events or other situations where uncertainty may be in the mind of a person, other measures are constructed, such as subjective probability.

procurement auction An auction in which multiple bidders compete to supply an item. Bids in a procurement auction are prices that bidders are willing to receive to supply the good. The lowest bidder wins and is paid her bid.

promise A response to the *favored action* that benefits the other player and that is not a best response, as specified within a *strategic move*. It is only feasibly made by the second mover in a game. Credibility is required because the specified action is not a best response. The strategic move referred to as "making a promise" entails declaring both a promise and a warning, whereas "making a combined threat and promise" entails declaring both a promise and a threat.

proportional representation This voting system requires that the number of seats in a legislature be allocated in proportion to each party's share of the popular vote.

pruning Using rollback analysis to identify and eliminate from a game tree those branches that will not be chosen when the game is rationally played.

punishment We reserve this term for costs that can be inflicted on a player in the context of a repeated relationship (often involving termination of the relationship) to induce him to take actions that are in the joint interests of all players.

pure coordination game A coordination game where the payoffs of each player are the same in all the Nash equilibria. Thus, all players are indifferent among all the Nash equilibria, and coordination is needed only to ensure avoidance of a nonequilibrium outcome.

pure public good A good or facility that benefits all members of a group, when these benefits cannot be excluded from a member who has not contributed efforts or money to the provision of the good, and the enjoyment of the benefits by one person does not significantly detract from their simultaneous enjoyment by others.

pure strategy A rule or plan of action for a player that specifies without any ambiguity or randomness the action to take in each contingency or at each node where it is that player’s turn to act.

quantal-response equilibrium (QRE) Solution concept that allows for the possibility that players make errors, with the probability of a given error smaller for more costly mistakes.

ranked-choice voting Another name for **single transferable vote**.

rational behavior Perfectly calculating pursuit of a complete and internally consistent objective (payoff) function.

rational irrationality Adopting a strategy that is not optimal after the fact, but serves a rational strategic purpose of lending credibility to a threat or a promise.

rationalizability A solution concept for a game. A list of strategies, one for each player, is a rationalizable outcome of the game if each strategy in the list is rationalizable for the player choosing it.

rationalizable A strategy is called rationalizable for a player if it is his optimal choice given some belief about what (pure or mixed strategy) the other player(s) would choose, provided this belief is formed recognizing that the other players are making similar calculations and forming beliefs in the same way. (This concept is more general than that of the Nash equilibrium and yields outcomes that can be justified on the basis only of the players' common knowledge of rationality.)

refinement A restriction that narrows down possible outcomes when multiple Nash equilibria exist.

repeated play A situation where a one-time game is played repeatedly in successive periods. Thus, the complete game is mixed, with a sequence of simultaneous-move games.

reputation Relying on the effect on payoffs in future or related games to make threats or promises credible, when they would not have been credible in a one-off or isolated game.

reserve price The minimum price set by the seller of an item up for auction; if no bids exceed the reserve, the item is not sold.

response rule A rule that specifies how you will act in response to various actions of other players.

revenue equivalence theorem (RET) A famous result in auction theory specifying conditions under which two auctions will generate the same expected revenue for the seller.

reversal paradox This paradox arises in an election with at least four alternatives when one of these is removed from consideration after votes have been submitted and the removal changes the identity of the winning alternative.

reversal terms A set of ballots that would generate the *reversal paradox* and that should together logically produce a tied vote between a pair of alternatives. In a three-candidate election among A, B, and C, the reversal terms are two ballots that show a reversal in the location of a pair of alternatives. For example, one ballot with A preferred to B preferred to C and another with B preferred to A preferred to C should produce a tie between A and B.

robustness A measure of the number of sets of voter preference orderings that are nondictatorial, satisfy independence of irrelevant alternatives and the Pareto property, and also produce a transitive *social ranking*.

rollback Analyzing the choices that rational players will make at all nodes of a game, starting at the terminal nodes and working backward to the initial node. Also called **backward induction**.

rollback equilibrium The strategies (complete plans of action) for each player that remain after rollback analysis has been used to prune all the branches that can be pruned.

root Same as **initial node**.

round A single vote within a larger *multistage procedure* that consists of multiple sequentially held votes.

salami tactics A method of defusing threats by taking a succession of actions, each sufficiently small to make it nonoptimal for the other player to carry out his threat.

sanction Punishment approved by society and inflicted by others on a member who violates an accepted pattern of behavior.

screening Strategy of a less-informed player to elicit information credibly from a more-informed player.

screening devices Methods used for screening.

sealed-bid auction An auction mechanism in which bids are submitted privately in advance of a specified deadline, sometimes in sealed envelopes.

secondary criterion Comparison of the fitness of a mutant with that of a member of the dominant population, when each plays against a mutant.

second-mover advantage A game has this if, considering a hypothetical choice between moving first and moving second, a player would choose the latter.

second-price auction A sealed-bid auction in which the highest bidder wins the auction but pays a price equal to the value of the second-highest bid; a special case of the *Vickrey auction*.

selection The dynamic process by which the proportion of fitter phenotypes in a population increases from one generation to the next.

self-selection Where different types respond differently to a screening device, thereby revealing their type through their own action.

semiseparating equilibrium A perfect Bayesian equilibrium in a game of asymmetric information, where the actions in the equilibrium convey some additional information about the players' types, but some ambiguity about these types remains.

separating equilibrium A perfect Bayesian equilibrium in a game of asymmetric information, where the actions in the equilibrium reveal player type.

separation of types An outcome of a signaling or screening game in which different types follow different strategies and get the different payoffs, so types can be identified by observing actions.

sequential moves The moves in a game are sequential if the rules of the game specify a strict order such that at each action node only one player takes an action, with knowledge of the actions taken (by others or himself) at previous nodes.

shading A strategy in which bidders bid slightly below their true valuation of an object.

shill bidder A fake bidder created by sellers at an auction to place fictitious bids for an object they are selling.

signaling Strategy of a more-informed player to convey his "good" information credibly to a less-informed player.

signals Devices used for signaling.

simultaneous moves The moves in a game are simultaneous if each player must take his action without knowledge of the choices of others.

sincere voting Voting at each point for the alternative that you like best among the ones available at that point, regardless of the eventual outcome. Also called **truthful voting**.

single-object auction An auction in which a single indivisible object is sold.

single-peaked preferences A preference ordering in which alternatives under consideration can be ordered along some specific dimension and each voter has a single ideal or most-preferred alternative with alternatives farther away from the most-preferred point providing steadily lower payoffs. Also called **Black's condition**.

single transferable vote A voting method in which each voter indicates her preference ordering over all candidates on a single initial ballot. If no alternative receives a majority of all first-place votes, the bottom-ranked alternative is eliminated and all first-place votes for that candidate are "transferred" to the candidate listed second on those ballots; this process continues until a majority winner emerges. Also called **instant-runoff voting (IRV)** or **ranked-choice voting**.

sniping Waiting until the last moment to make a bid.

social optimum In a collective-action game where payoffs of different players can be meaningfully added together, the social optimum is achieved when the sum total of the players' payoffs is maximized.

social ranking The preference ordering of a group of voters that arises from aggregating the preferences of each member of the group.

spillover effect Same as **external effect**.

spoiler Refers to a third candidate who enters a two-candidate race and reduces the chances that the leading candidate actually wins the election.

strategic form Same as **normal form**.

strategic misrepresentation of preferences Refers to strategic behavior of voters when they use rollback to determine that they can achieve a better outcome for themselves by not voting strictly according to their preference orderings.

strategic move Action taken at a pregame stage that changes the strategies or the payoffs of the subsequent game (thereby changing its outcome in favor of the player making the move).

strategic order The order of moves from a game-theoretic point of view, determined by considerations of observability and irreversibility. It may differ from the chronological order of actions and, in turn, determine whether the game has sequential or simultaneous moves.

strategic uncertainty A player’s uncertainty about an opponent’s moves made in the past or made at the same time as her own.

strategic voting Voting in conformity with your optimal rational strategy found by doing rollback analysis on the game tree of the voting procedure.

strategy A complete plan of action for a player in a game, specifying the action he would take at all nodes where it is his turn to act according to the rules of the game (whether these nodes are on or off the equilibrium path of play). If two or more nodes are grouped into one information set, then the specified action must be the same at all these nodes.

subgame A game comprising a portion or remnant of a larger game, starting at a noninitial node of the larger game.

subgame-perfect equilibrium (SPE) A configuration of strategies (complete plans of action) such that their continuation in any subgame remains optimal (part of a rollback equilibrium), whether that subgame is on- or off-equilibrium. This ensures credibility of all the strategies.

successive elimination of dominated strategies Same as **iterated elimination of dominated strategies**.

superdominant A strategy is superdominant for a player if the worst possible outcome when playing that strategy is better than the best possible outcome when playing any other strategy.

surplus A player' s surplus in a bargaining game is the excess of his payoff over his BATNA.

terminal node This represents an end point in a game tree, where the rules of the game allow no further moves, and payoffs for each player are realized.

threat A response to the default action that harms the other player and that is not a best response, as specified within a strategic move. It can only feasibly be made by the second mover in a game. Credibility is required because the specified action is not a best response. The strategic move referred to as "making a threat" entails declaring both a threat and an affirmation, whereas "making a combined threat and promise" entails declaring both a threat and a promise.

tit-for-tat (TFT) In a repeated prisoners' dilemma, this is the strategy of [1] cooperating on the first play and [2] thereafter doing each period what the other player did the previous period.

transitive ordering A preference ordering for which it is true that if option A is preferred to B and B is preferred to C, then A is also preferred to C.

trigger strategy In a repeated game, this strategy cooperates until and unless a rival chooses to defect, and then switches to noncooperation for a specified period.

truthful bidding A practice by which bidders in an auction bid their true valuation of an object.

truthful voting Same as **sincere voting.**

type Players who possess different private information in a game of asymmetric information are said to be of different types.

ultimatum game A form of bargaining where one player makes an offer of a particular split of the total available surplus, and the other has only the all-or-nothing choice of accepting the offer or letting the game end without agreement, when both get zero surplus.

uniform distribution A common statistical distribution in which the *distribution function* is horizontal; data are distributed uniformly at each location along the range of possible values.

valuation The benefit that a bidder gets from winning the object in an auction.

variable-threat bargaining A two-stage game where at the first stage you can take an action that will alter the BATNAs of both bargainers (within certain limits), and at the second stage bargaining results in the Nash solution on the basis of these BATNAs.

Vickrey auction An auction design proposed by William Vickrey in which truthful bidding is a weakly dominant strategy for each bidder. When a single object is sold, the Vickrey auction is the same as the *second-price auction*.

warning A response to the default action that is a best response, as specified within a strategic move. Only the second mover in a game can feasibly make a warning. However, credibility is not required since the specified action is already the player’s best response. The strategic move referred to as “making a promise” entails declaring both a promise and a warning.

war of attrition A contest between multiple players in which each player decides when to retreat, the victor is whoever remains the longest, and choosing to remain longer is costly for each player.

weakly dominant A strategy is weakly dominant for a player if the outcome when playing that strategy is never worse than the outcome when playing any other strategy, no matter what strategies other players adopt.

winner's curse A situation in a common-value auction where the winner fails to take account of the fact that when she wins, she is likely to have made an overly optimistic estimate of the object's value. Bidders who correctly anticipate this possibility can avoid the winner's curse by lowering their bids appropriately.

zero-sum game A game where the sum of the payoffs of all players equals zero for every configuration of their strategy choices. (This is a special case of a *constant-sum game*, but in practice no different because adding a constant to all the payoff numbers of any one player makes no difference to his choices.)