From the book *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*.
By David Easley and Jon Kleinberg. Cambridge University Press, 2010.
Complete preprint on-line at http://www.cs.cornell.edu/home/kleinber/networks-book/

Chapter 22

Markets and Information

In this final part of the book, we build on the principles developed thus far to consider the design of institutions, and how different institutions can produce different forms of aggregate behavior. By an *institution* here, we mean something very general — any set of rules, conventions, or mechanisms that synthesizes individual behavior across a population into an overall outcome. In the next three chapters, we will focus on three fundamental classes of institutions: markets, voting, and property rights.

We begin by discussing markets, and specifically their role in aggregating and conveying information across a population. Each individual participant in the market arrives with certain beliefs and expectations — about the value of assets or products, and about the likelihood of events that may affect these values. The markets we study will be structured so as to combine this set of beliefs into an overall outcome — generally in the form of market prices — that represents a kind of synthesis of the underlying information.

This is part of a broad issue we have seen several times so far: the fact that individuals' expectations affect their behavior. For example, we saw this in Chapter 8 on Braess's Paradox, where the optimal route depends on which routes others are expected to choose; in Chapter 16 on information cascades, where people draw inferences about the unknown desirability of alternatives (restaurants or fashions) from the behavior of others; and in Chapter 17 on network effects, where the unknown value of a product (a fax machine or a social-networking site) depends on how many others are also expected to use the product. In each of these cases, individuals have to decide what to do without knowing exactly what will happen. Will the route be crowded or not? Is the restaurant good or bad? Will others also join the social networking site? In all of these situations, individuals' expectations about payoffs matter for how they will choose.

Along with the similarities among these settings, there is also an important difference

Draft version: June 10, 2010

that will be fundamental to our discussion here: whether the unknown desirability of the different alternatives is *exogenous* or *endogenous*. Exogenous desirability means that a given alternative is inherently a good idea or a bad idea, regardless of how the individuals make their decisions. Thus, for example, in our model of information cascades, people decided whether to accept or reject an option that was in fact fundamentally either good or bad, and the desirability of the option wasn't affected by whether people accepted it or not. Endogenous desirability is different, and somewhat more subtle: it means that the desirability of an alternative depends on the actual decisions people make about it. In our model of network traffic and Braess's Paradox, no particular route is *a priori* crowded or not; a route becomes crowded if many people choose it. Similarly, we can't tell whether a product with network effects — like a fax machine — is worth purchasing or not until we know whether many people in fact purchase it.

We will consider both types of cases in this chapter. First we will look at what happens in asset markets where the desirability of the assets is exogenous, but unknown. We will begin this analysis by focusing on betting markets as a simple, stylized domain with exogenous but uncertain outcomes. We describe how individuals behave and how prices are set in betting markets, and then discuss how the ideas we develop about betting markets provide insight into more complex settings like stock markets. After this, we will consider what happens in markets where the desirability of the items is endogenous. The issue we focus on in this case is the role of asymmetric information.

The next two chapters in this part of the book will discuss voting and the role of property rights. Markets and voting mechanisms are alternative institutions that aggregate individual behavior into outcomes for the group. One important difference between them is that voting mechanisms are typically used to produce a single group decision while in markets each individual may choose a different outcome. The final chapter discusses the role of property rights in influencing what outcomes are possible.

22.1 Markets with Exogenous Events

In this section we begin by examining how markets aggregate opinions about events in settings where the underlying events are exogenous — the probabilities of the events are not affected by the outcomes in the market. *Prediction markets* are one basic example of this setting. These are markets for (generally very simple) assets which have been created to aggregate individuals' predictions about a future event into a single group, or market, opinion. In a prediction market, individuals bet on the outcome of some event by trading claims to monetary amounts that are conditional on the outcome of the event.

One of the most well-known uses of prediction markets has been for the forecasting of election results. For example, the Iowa Electronic Markets[1] ran a market (one of many with this structure) in which individuals could buy or sell a contract that would pay $1 in the event that a Democrat won the 2008 U.S. Presidential election, and would pay nothing if this event did not occur. An individual who bought this contract was betting that a Democrat would win the election. The corresponding contract that paid $1 if a Republican won was also available. In Figure 1.13 from Chapter 1, we saw a plot of the prices for these two contracts over time, and we saw how the movement of the prices followed the — exogenous — course of events affecting the perceived likelihood of the election outcome.

In a prediction market, or any other market, there are two sides to any trade: what someone buys, someone else sells. So a trade in a prediction market means that two people disagree about which side of the bet they want to take. But note that at the price where the trade actually occurs, both the buyer and the seller find the trade desirable. In a sense that we will make precise later, the price *separates their beliefs*: their beliefs are on opposite sides of the price, and we can view the price as an average of their beliefs. This is the motivation for the usual interpretation that the price in a prediction market is an average prediction about the probability of the event occurring. So, if the price of the Democrat-wins contract is 60 cents, the usual interpretation is that "the market" believes that the probability of a Democrat winning the election is 0.6. Of course, the market itself does not have beliefs — it's simply an institution, a place where trade is conducted under a particular set of rules. So when we say that the market believes something about a future event, this phrase really means that the market price represents an average belief.

Betting markets for sporting events such as horse races are also markets that aggregate diverse opinions into a price. As is the case with prediction markets, the outcome of the sporting event is independent of the betting behavior of the participants. Of course, some bettors may have more accurate beliefs than other bettors. But, assuming that there is no cheating, what happens in the betting market does not affect the outcome of the sporting event.

Markets for stocks are similar to prediction markets or betting at horse races, and we will use the understanding we develop for these betting markets to help us understand how the stock market works. In both betting markets and the stock market, individuals make decisions under uncertainty about the value of a contract, bet, or stock, and the market aggregates their diverse opinions about the value of the asset. But there is also an important difference between the stock market and gambling. The price set in a gambling market, and who holds what bets, are both interesting, but they do not affect the allocation of real capital. On the other hand the stock market allocates the available shares of stock in a company. The market price for these shares determines the cost of equity capital for the company; it is the expected rate of return that investors demand in order to hold the existing shares of

[1] www.biz.uiowa.edu/iem/

stock or to buy new shares of stock. The financial capital that the company receives for its shares of stock affects its real investment decisions and thus the future value of the stock. So there is an indirect link between the aggregate opinion in the market about the company, its stock price, and the actual value of the company. But this link is really quite indirect, and for a first pass at understanding stock market prices it is reasonable to ignore it. (In fact, much of the academic literature on asset pricing also ignores this effect.)

As we will see, we can interpret the price of the asset being traded, whether it's a stock, a contract that pays out if a Democrat wins, or a betting ticket at a race-track, as a market prediction about some event. In the next section, we will examine how these markets work and we will build an understanding of the circumstances under which they do a good or bad job of producing a useful aggregate prediction.

22.2 Horse Races, Betting, and Beliefs

It is easiest to understand what goes on in these markets if we begin with the simple example of betting on a two-horse race [64]. Suppose that two horses, whom we'll call A and B, will run a race which one of them will win (we will ignore the possibility of a tie). How should a bettor who has w dollars available to bet allocate his wealth between bets on the two horses?

We will make the assumption that the bettor plans to bet all of this money w on the two horses in some fashion: We will let r be a number between 0 and 1 representing the fraction of his wealth that he bets on horse A; the remaining $1 - r$ fraction of his wealth will be bet on horse B. The bettor could bet all of his money on horse A ($r = 1$), all of it on horse B ($r = 0$), or he could split it up and bet some on each horse (by choosing r strictly between 0 and 1). The only thing the bettor cannot do is save some of the money and not bet it. We will see later that in our model there is a betting strategy that returns his wealth for sure, so this lack of a direct way to not bet is not really a constraint.

It seems reasonable to expect a bettor's choice of bet will depend on what the bettor believes about the likelihood of each horse winning the race. Let's suppose that the bettor believes that horse A will win with probability a, and that horse B will win with probability $b = 1 - a$. It seems sensible to suppose that the fraction of wealth r bet on horse A won't decrease if the probability of A winning increases, and r should be equal to one if $a = 1$ because then betting on horse A is a sure thing. But if neither horse is a sure thing, what should the bet look like? The answer to this question depends on more than the probability of A or B winning the race; in particular, it may depend on two other factors.

First, the bettor's choice of bet may depend on the odds. If the odds on horse A are for example three-to-one, then a one-dollar bet on horse A will pay three dollars if horse A wins, and will pay nothing if horse A loses. More generally, if the odds on horse A are o_A, and the odds on horse B are o_B, then a bet of x dollars on horse A will pay $o_A x$ dollars if A

wins, and a bet of y dollars on horse B will pay $o_B y$ dollars if horse B wins. A bettor might find high odds attractive and bet a lot on a horse with high odds in the hope of winning a large sum of money. But if he does this then he will have little left to bet on a horse with low odds, and if that horse wins the race he will be left with very little money. How a bettor evaluates the different levels of risk in these options is the topic we turn to next.

Modeling Risk and Evaluating the Utility of Wealth. A bettor's reaction to risk is the second factor that influences his choice of bets. It seems reasonable to suppose that a bettor who is very risk-averse will bet so as to have some money left no matter which horse wins, by betting some money on each horse. Someone who does not care as much about risk may place a bet more skewed toward one of the horses, and a person who does not care about risk at all might even bet everything on one horse. This issue of characterizing risk will become even more important when we move from the simple example of betting on horse races to investing in financial markets. Individuals invest significant amounts of their wealth in a wide variety of assets, all of which are subject to some risk, and it is very natural to assume that most people will not want to choose investment strategies where there is a plausible scenario in which their savings are reduced to zero. We can formulate the same issues in our current example, considering the bets on horse A or B as the alternatives that contain risk, while also keeping in mind that the whole example is a simply-formulated metaphor for markets with exogenous events in general.

How do we model the bettor's attitude toward risk? We saw a simple version of this question in Chapter 6 in which we asked how a player in a game evaluates the payoff of a strategy with random payoffs. Our answer was that the player evaluates each strategy according to the expected value of its payoff, and we will use the same idea here. We assume that the bettor evaluates a bet according to the expected value of the payoff on the bet. But here we need to be a bit careful. What really is the payoff on the bet? Is it the amount of money won or lost, or is it how the bettor feels about the amount of money?

Presumably the bettor prefers outcomes in which he obtains larger amounts of money, but how does his actual evaluation of the outcome depend on the amount of wealth he acquires? To make this precise, we need to define a numerical way of specifying the bettor's evaluation of the outcome as a function of his wealth, and then use this numerical measure as the bettor's payoff. We will do this with a *utility function* $U(\cdot)$: when a bettor has a wealth w, his evaluation of the outcome — i.e., his payoff — is equal to the quantity $U(w)$.

The simplest example of a utility function is the linear function $U(w) = w$, in which a bettor's utility for wealth is exactly its value. We could also consider more general linear utility functions, of the form $U(w) = aw+b$ for some positive number a. With such functions, the bettor's utility increase from gaining a dollar is precisely equal to his utility decrease from losing a dollar. At first glance, it might seem strange to use any other utility function, but in fact linear utility functions predict behavior that doesn't align well with either empirical evidence or common sense.

An easy way to see this is to ask whether a bettor would accept a fair gamble. We consider a particular instance of a fair gamble to illustrate the issue; other examples show similar effects. Suppose that a bettor's current total wealth is w, and he is offered a gamble in which he gains w dollars with probability $\frac{1}{2}$, and loses w dollars with probability $\frac{1}{2}$. In other words, with probability $\frac{1}{2}$, his wealth after the bet will be $2w$, and with probability $\frac{1}{2}$, his wealth after the bet will be 0. We call this "fair" in the sense that the expected value of the bettor's wealth after the bet is $\frac{1}{2} \cdot (2w) + \frac{1}{2} \cdot 0 = w$. Now, a bettor with the linear utility function $U(w) = w$ would be indifferent between accepting and rejecting this gamble, since his expected utility from accepting the gamble would be

$$\frac{1}{2}U(2w) + \frac{1}{2}U(0) = \frac{1}{2} \cdot (2w) + \frac{1}{2} \cdot 0 = w,$$

which is the same as his expected utility from passing up the opportunity to gamble. The same would hold for any linear utility function. But if we imagine our scenario serving as a model for the behavior of an individual investing his wealth in a financial market, then our calculation corresponds to the following premise: an investor whose total net worth is \$1 million would be indifferent between accepting and rejecting an investment strategy in which he gains or loses a million dollars with equal probability. This isn't a good model for investor behavior; we'd much prefer a model in which the investor views the downside of such a strategy — having your net worth reduced to zero — as far outweighing the upside of potentially doubling your net worth. In other words, we want to model the idea that the investor, at least to some extent, views this strategy as highly risky despite its expected net change of zero to his wealth.

We can capture this kind of behavior by assuming that the bettor's utility grows at a decreasing rate as a function of his wealth w. Examples of utility functions that have this property are $U(w) = w^{1/2}$ and $U(w) = \ln(w)$, the natural logarithm of wealth. In these cases utility still increases as a function of wealth; it is simply that the rate of increase slows down as wealth increases: the appeal of gaining each extra dollar goes down as you become wealthier. With a bit of analysis, we can check that a bettor with this type of utility would reject a fair gamble. For example, if $U(w) = w^{1/2}$, then he would evaluate the expected utility from accepting the bet discussed above as

$$\frac{1}{2}U(2w) + \frac{1}{2}U(0) = \frac{1}{2} \cdot (2w)^{1/2} + \frac{1}{2} \cdot 0 = 2^{-1/2} \cdot w^{1/2}.$$

This is less than his current utility of $w^{1/2}$, which he could maintain simply by refusing to bet. Thus, with this kind of utility function, which grows at a decreasing rate in the wealth w, the bettor would reject the gamble.

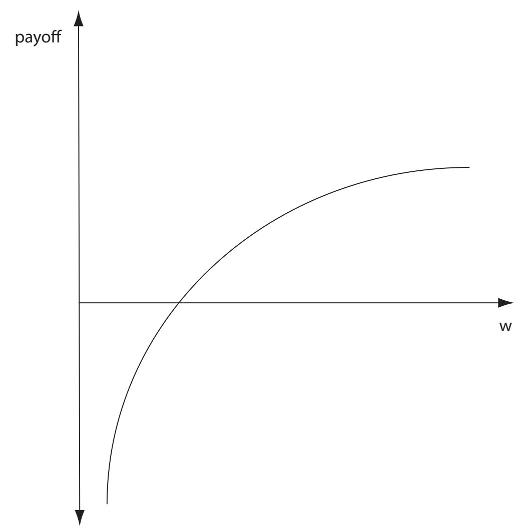

Figure 22.1: When we assume that an individual's utility is logarithmic in his wealth, this means that utility grows at a decreasing rate as wealth increases.

We will be using these types of utility functions in our analysis here. It's important to remember, however, that regardless of how we model utility as a function of wealth, we assume that the bettor evaluates bets according to the expected value of the utility; the difference is simply in the shape of the bettor's utility function.

Logarithmic Utility. To build a simple model of how a bettor might behave, we'll suppose in particular that the bettor's utility function is the natural logarithm of wealth, $\ln(w)$, where $w > 0$ is the bettor's wealth. This utility function is plotted in Figure 22.1; as noted above, it grows with the wealth, but the rate of growth slows down with increasing wealth. This logarithmic form for utility has a simple intuitive property: the bettor receives the same utility benefit from doubling his wealth, regardless of how much he currently has. In other words, the value of each additional dollar declines as wealth increases, but the value of doubling one's wealth is always the same. To see why this is true, we use the following basic fact about logarithms:

$$\ln(x) - \ln(y) = \ln(x/y) \tag{22.1}$$

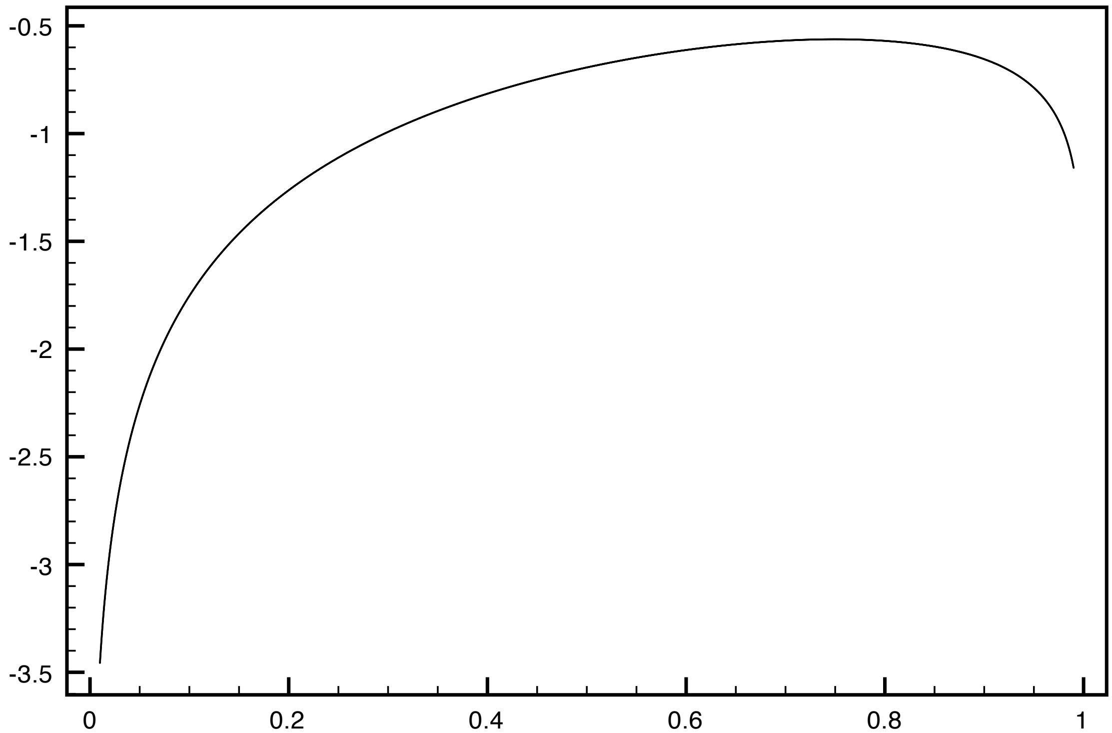

Figure 22.2: A plot of $a \ln(r) + b \ln(1-r)$ as a function of r, when $a = 0.75$. The maximum is achieved when $r = a$.

for any x and y. Given this fact, the increase in utility from doubling your money, when your current wealth is w, is equal to

$$\ln(2w) - \ln(w) = \ln(2w/w) = \ln(2),$$

where the first equality follows by plugging $x = 2w$ and $y = w$ into Equation (22.1). A similar argument would hold for any multiplicative increase or decrease in the bettor's wealth: the change in utility doesn't depend on the current wealth.

The logarithmic utility function will make our analysis particularly clean, although the general ideas of the analysis apply equally well to any utility function that grows at a decreasing rate in wealth.

The Optimal Strategy: Betting Your Beliefs. Let's now figure out the optimal strategy for our bettor, given the logarithmic utility of wealth, the odds being offered, and the bettor's beliefs about the respective probabilities that horses A and B will win.

Recall that the odds on horse A are o_A and the odds on horse B are o_B. Suppose that the bettor bets a fraction r of his wealth on horse A. Then since the amount bet on A is rw, the bettor's wealth will be rwo_A if horse A wins. The amount bet on B is $(1-r)w$, so the bettor's wealth will be $(1-r)wo_B$ if horse B wins. The bettor believes there is a probability a that horse A will win, and a probability $b = 1 - a$ that horse B will win. So with the given betting strategy r and with these probabilities, the bettor ends up with a utility of $\ln(rwo_A)$ with probability a (in the event horse A wins), and a utility of $\ln((1-r)wo_B)$ with probability $1 - a$ (in the event horse B wins). Adding these up, the expected utility after the bet is

$$a \ln(rwo_A) + (1-a) \ln((1-r)wo_B). \tag{22.2}$$

The bettor wants to choose r to maximize the value in this expression.

As a step toward maximizing this, we can use another basic fact about logarithms, closely related to Equation (22.1):

$$\ln(x) + \ln(y) = \ln(xy) \tag{22.3}$$

for any x and y. As a result, we can unpack the products of variables inside the logarithms in Formula (22.2), arriving at an equivalent way to write the expected utility that the bettor wants to maximize:

$$a \ln(r) + (1-a) \ln(1-r) + a \ln(wo_A) + (1-a) \ln(wo_B). \tag{22.4}$$

Something interesting is already happening here. The third and fourth terms of this expression do not contain the value r, and this value r is the only thing the bettor has control over. So the bettor's maximization problem is really just to maximize the first two terms: he needs to choose r to maximize

$$a \ln(r) + (1-a) \ln(1-r). \tag{22.5}$$

This leads to a counter-intuitive conclusion, but one that follows directly from our assumption of a logarithmic utility function: the formula (22.5) does not contain the odds o_A and o_B, and so when the bettor determines the optimal choice of r by maximizing this formula, it will not depend on the value of the odds. This makes sense once we think further about what logarithmic utility really means. We can interpret the odds of o_A on horse A as follows: In the event that horse A wins, you will first be paid rw dollars, and then your wealth will be further increased by a factor of o_A. But we just argued that with logarithmic utilities, the benefit from a multiplicative increase in your wealth is a fixed amount independent of how much you have. So while this final multiplicative boost of o_A is a nice bonus, the value you assign to it is independent of how much money you have at the time, and hence it does not affect your choice of r.

So let's go back to the bettor's problem of maximizing Formula (22.5). Its typical shape as a function of r is shown in Figure 22.2: it drops steeply near $r = 0$ and $r = 1$, and it assumes a unique maximum in between. With some very simple calculus, one can show that it is maximized at $r = a$. In what follows, we'll use this result in a "black-box" fashion, without worrying how it is obtained, but the argument is very short. The derivative of the expression in Formula (22.5) with respect to r is

$$\frac{a}{r} - \frac{1-a}{1-r} \tag{22.6}$$

If we set this to zero, we get an equation that is solved simply by setting $r = a$, and this is the maximum point.

This result has a nice interpretation: *the bettor bets his beliefs.* The fraction of wealth bet on horse A is the bettor's belief about the probability that horse A wins. Note also that the optimal bet has the sensible property that the amount bet on A increases with the probability of A winning, and it approaches the bettor's full wealth as this probability approaches 1.

We now use this basic result to study what happens in markets with many participants. Throughout this analysis we will continue to assume that all bettors have logarithmic utility. If we were to use a different utility function, the overall results to follow would still have the same qualitative behavior, but the analysis would become much more elaborate and some of the specific facts we use would no longer hold. In particular, with a different utility function, the bettors' decisions would no longer necessarily be independent of the odds, and this would lead to more complex reasoning about bettor behavior.

22.3 Aggregate Beliefs and the "Wisdom of Crowds"

When there is only one bettor, we can learn about the bettor's beliefs by observing his optimal strategy, but with only one bettor the race-track cannot be said to be aggregating multiple opinions. To understand how aggregation works, we now consider systems where there are multiple bettors.

Let's suppose that there are N bettors named $1, 2, 3, \ldots, N$, and that each bettor n believes there is a probability of a_n that horse A will win, and thus a probability of $b_n = 1 - a_n$ that B will win. We will allow the bettors to disagree about the probability of winning, but we don't require that they actually do disagree.[2] As we will see, if they agree, then although they wouldn't be willing to take opposite sides of a bet with respect to each other, we can determine the market odds and the aggregate opinion — it will simply be the commonly held opinion.

[2]We take bettors' disagreement about the probabilities as exogenously given. It is also interesting to ask what happens if any disagreements are generated by differing information. But this case is much more complex as here bettors also need to make inferences about the information of others from whatever market statistics they can observe [345].

There is no reason to assume that the bettors all have the same wealth, and once they begin betting there will be winners and losers, so eventually their wealths will have to differ. So we might as well allow for different wealths at the beginning. Suppose that bettor n has wealth w_n and thus the total wealth is the sum of all w_n, which we write

$$w = w_1 + w_2 + \cdots + w_N.$$

We will assume that all of the bettors evaluate wealth using the same utility function, and we will continue to use the natural logarithm of wealth, $\ln(w)$, for this utility.

In Section 22.2, we saw that the optimal betting strategy for bettor n with belief a_n is $r_n = a_n$: that is, bettor n will bet $a_n w_n$ on horse A. Correspondingly, bettor n will bet $(1 - a_n)w_n$ on horse B. So the amount that all the bettors together bet on horse A is the sum

$$a_1 w_1 + a_2 w_2 + \cdots + a_N w_N,$$

and the total amount bet on horse B is the sum

$$b_1 w_1 + b_2 w_2 + \cdots + b_N w_N.$$

Since each bettor bets all of his wealth, the total amount bet is the aggregate wealth w.

The Odds Determined by the Race-Track. Now we would like to determine the odds that the race-track should offer on horses A and B if it wants to break even — that is, if it wants to pay out to the bettors exactly the total amount bet, no matter which horse wins. We assume that the race-track collects the bets from the bettors, and so it has w in total bets. It then uses this money to pay off the winning bets. We will assume that no matter which horse wins the race, the race-track pays out everything that it collects. That is, it has no cost and makes no profit.

If horse A wins, the amount that is owed to bettor n is $a_n w_n o_A$. The total amount owed to the bettors is the sum of their winnings, which is equal to

$$a_1 w_1 o_A + \cdots + a_N w_N o_A.$$

In order to have the amount that is paid out to bettors in the event that horse A wins equal the amount of money w that the track collected, the odds on horse A must solve

$$a_1 w_1 o_A + \cdots + a_N w_N o_A = w. \tag{22.7}$$

That is, the equilibrium odds on horse A are determined so that the track just breaks even if A wins. Solving for the inverse of the equilibrium odds on horse A (which will produce a somewhat nicer expression than the formula for the actual odds), we get

$$\frac{a_1 w_1}{w} + \cdots + \frac{a_N w_N}{w} = o_A^{-1}. \tag{22.8}$$

If we write $f_n = w_n/w$ for the share of the total wealth held by bettor n, then this can be written as

$$a_1 f_1 + \cdots + a_N f_N = o_A^{-1}. \tag{22.9}$$

An analogous calculation shows that the equilibrium inverse odds on horse B are

$$b_1 f_1 + \cdots + b_N f_N = o_B^{-1}. \tag{22.10}$$

These inverse odds have a nice interpretation. If the odds on horse A are 4 (i.e. four-to-one, with a one-dollar bet paying four-dollars), then in order to have a one-dollar payout in the event that horse A wins the bettor would need to bet one-fourth of a dollar on horse A. This amount is the inverse of the odds for A: that is, a bet of o_A^{-1} dollars on horse A will result in a payment of \$1 in the event that horse A wins the race. Thus the inverse odds on A are the price of a dollar to be paid in the event that A wins, and similarly the inverse odds on B are the price of a dollar to be paid in the event that B wins.

State Prices. Let's denote these "prices of a dollar" by $\rho_A = o_A^{-1}$ for the event that horse A wins and $\rho_B = o_B^{-1}$ for the event that horse B wins. These prices are usually called *state prices*, as they are the price of a dollar in the event that a certain future state of the world is reached [24].

There is one more feature of equilibrium odds that is important. Let's ask how much a bettor would have to pay now to get one dollar for sure after the race. To do this the bettor needs to bet enough on horse A so as to receive one dollar if A wins, and enough on horse B so as to receive one dollar if horse B wins. These amounts are, as we saw before, o_A^{-1} and o_B^{-1}. So the amount needed to guarantee receiving one dollar in either case — i.e., regardless of the outcome — is $o_A^{-1} + o_B^{-1}$. Let's use the equilibrium values of odds to figure out how much money this takes.

$$\begin{aligned}
o_A^{-1} + o_B^{-1} &= (a_1 f_1 + \cdots + a_N f_N) + (b_1 f_1 + \cdots + b_N f_N) \\
&= (a_1 + b_1) f_1 + \cdots + (a_N + b_N) f_N \\
&= (1 \cdot f_1) + \cdots + (1 \cdot f_N) \\
&= 1.
\end{aligned}$$

This calculation shows there is a betting strategy that will turn one dollar before the race into one dollar for sure after the race. This is the sense in which our assumption that the bettors actually bet all of their wealth is not a constraint. Any portion of their wealth that they don't want to risk they can bet according to the inverse odds. This calculation also gives us the very useful property that inverse odds, or state prices, sum to one.

Now having done all of these calculations we are in a position to interpret the state prices. First, note that if every bettor believes that the probability that horse A will win is a, then

$\rho_A = a$. That is, if the bettors agree about the probabilities then the market accurately reflects these beliefs, with the state price equal to the common belief. Second, since the wealth shares sum to one, *the state prices are weighted averages of the bettors' beliefs*. The weight on bettor n's beliefs is bettor n's share, f_n, of the aggregate wealth. In particular, if a bettor has no wealth then the state price is not influenced by his beliefs as he cannot bet. Alternatively, if a bettor controls all of the wealth then the state price is his probability. More generally, how much influence a bettor's beliefs have on the state price depends on how much of the aggregate wealth is controlled by that bettor.

So it really does make sense to think of the state prices as the market's averaging of individual beliefs — or, in the more typical phrasing, they can be interpreted as the *market's beliefs*. For our horse-race market (with the logarithmic utility function) this market probability is the weighted average of the investors' beliefs, with each investor's weight determined by his share of the wealth.[3]

The Relationship to the "Wisdom of Crowds." What does this analysis say about the intuition popularized by recent books such as James Surowiecki's *The Wisdom of Crowds* [383]? The basic argument there, drawing on a long history of intuition about markets, is that the aggregate behavior of many people, each with limited information, can produce very accurate beliefs.

Our results on state prices illustrate some of the technical basis for this intuition. In particular, we found that the crowd at the racetrack determines the odds, or the state prices, and these odds are an average of the opinions in the crowd. If the opinions in the crowd about the probability of horse A winning are independently drawn from a distribution whose mean is equal to the true probability of horse A winning, and if wealth shares are equal, then the state prices actually do converge to the true probabilities as the size of the crowd grows. This occurs because the state prices are actually the average belief in the crowd, and this average converges to the truth with the size of the crowd.[4]

But these claims have two important qualifications embedded in them, both of which are important for understanding the limitations of the wisdom of crowds. First, it is important that the opinions are independent. We explored the subtleties of non-independent opinions in Chapter 16, noting the difficulties they pose in reasoning about the behavior of crowds, and the fact that they can lead to poor aggregate predictions even when many people participate. Second, it is important that all beliefs are equally weighted. If some bettors have more wealth than others, then state prices place more weight on their beliefs than on the beliefs of

[3]The logarithmic utility function is important for the exact form of this relationship. With other utility functions, state prices also depend on individuals' attitudes toward risk.

[4]In this discussion we treat bettors' beliefs as fixed and exogenously given. If instead, bettors' beliefs differ because they have differing information, then bettors should learn from prices. In the case in which beliefs are independent draws from a distribution with mean equal to the true probability, the market price reveals the average belief, and all bettors should use it to update their individual beliefs via the market price.

those who have little wealth. Whether this reduces or improves the accuracy of state prices depends on whether the beliefs of these wealthy bettors are more or less accurate than those of the poorer bettors. One might expect that over time those bettors with more accurate beliefs would become rich as they tend to make better bets than do those with less accurate beliefs. If this occurs then more accurate beliefs will have more weight placed on them and the market price will become a better and better predictor. We investigate this idea more fully in Section 22.10.

It is also interesting to think about what happens over time to state prices as the bettors watch horse races and learn about the likelihood of each horse winning. Suppose, for example, that horses A and B run a race against each other every week, and that the outcomes of the races are independent. If the true probability of horse A winning is a, then the fraction of times that A wins the race will converge to a. A bettor who watches the races, and who initially forecasts a winning probability for A that is not a, should modify his beliefs in light of his experience. In our examination of Bayesian learning in Chapter 16 we argued that over time an observer who watches independent events and employs Bayes' Rule learns the true probability. (This result is reviewed and expanded on in Section 22.10.) So over time, each bettor's belief about the probability of A winning will converge to a, and similarly each bettor's belief about the probability of B winning will converge to b. The state prices are weighted averages of these beliefs, so they too converge to a and b.

22.4 Prediction Markets and Stock Markets

Thus far we have been telling a story about horse races, but there is a direct analogy to any market where participants purchase assets whose future value depends on the outcome of uncertain events. Two specific examples are prediction markets and — by far the most consequential application of these ideas — stock markets. In both cases, we will see that state prices play a key role in how we reason about what takes place in the market.

Prediction Markets. In a prediction market, individuals trade claims to a one-dollar return conditional on the occurrence of some event. For example, as we discussed at the beginning of the chapter, participants might trade claims to a one-dollar return in the event that a Democrat wins the next U.S. Presidential election. The institutional structure of prediction markets differs from the structure of a betting market at a race-track. In a prediction market individuals trade with each other through the market, while at a race track individuals place bets directly with the track. Nonetheless, prices play the same role in both markets. The inverse odds on a horse are the cost of a one-dollar return in the event that the horse wins the race. Similarly, the price of a contract in a prediction market is the price of a one-dollar return in the event specified in the contract (such as a particular outcome of an election). In both cases the prices reflect an averaging of the beliefs of the participants in the market.

Here, we will ignore the various institutional structures of prediction markets and instead see how much we can discover about them by applying our analysis of horse races via state prices. Consider, for example, the prediction market for the 2008 U.S. Presidential election with two possible outcomes: a Democrat wins or a Republican wins. (The same analysis can handle prediction markets with many plausible outcomes, such as the earlier prediction market for the identity of the Democratic and Republican nominees for President in 2008.)

Let f_n be the share of the total wealth bet in the prediction market that is bet by trader n. Let a_n and b_n be trader n's probabilities of a Democrat and a Republican winning, respectively. Then, just as was the case for horse races, the market price ρ^D for the contract on a Democratic winner will be the wealth-share weighted average of the investors' probabilities of a Democrat winning the election. That is,

$$\rho^D = a_1 f_1 + \cdots + a_N f_N. \tag{22.11}$$

Similarly, the price of the contract on a Republican winner will be the wealth-share weighted average of the investors' probabilities of a Republican winning the election. In this discussion we are looking at a snapshot of the prediction market at one point in time, which we can think of as a point just before the event occurs. We can also examine the dynamic behavior of the market over periods of time during which beliefs and wealth shares are both likely to change; such dynamic questions are the focus of Section 22.10 at the end of this chapter.

Are the prices of these contracts good predictors of the outcome of the election? They are weighted averages of the beliefs of the investors in the markets, but as we saw with horse races this doesn't necessarily make them either good or bad predictors. It depends on the dispersion of beliefs in the population of investors and on how wealth shares are distributed across those investors. One way to address this question empirically is to look at the predictions made by real prediction markets and ask how well they have done at predicting the outcome of actual events. An interesting paper by Berg, Nelson, and Rietz [51] shows that for the period 1988–2004, the Iowa Electronic Markets did a significantly better job of predicting the outcome of U.S. Presidential elections than was done by an average of the major national polls.

Stock Markets. Stock markets also provide individuals with the opportunity to bet on future states of the world, but these are more complicated bets, since stocks don't just offer a one-dollar return in the event that a particular state occurs. Instead a share of stock in a company offers a monetary amount that will vary depending on which of possibly many states occurs. These states may be things like "the company's current investment in research and development succeeds", "a strong new competitor enters the market", "the demand for the goods produced by the company grows more rapidly than expected", or "the workers go on strike". Each of these states would have an impact on the future value of the stock in the company, and so it's reasonable to think of the stock as providing a different amount conditional on each state. The difference between a bet on a two-horse race and a share of stock, then, is that there are many states, as there would be in a many-horse race, and the amount of money that the owner of a share of stock has a claim to is not given by explicit odds but instead is determined by the value of the stock in each of these states.[5]

If we knew the price of a dollar in each of these states (the state prices), and the value of the stock in each state, we would know how much investors should be willing to pay for the stock conditional on each state: it would be the value of the stock in that state times the price of a dollar in that state. The price of the stock today would be the sum of these terms across all the states. As long as the collection of stocks traded are rich enough, it is possible to determine the state prices from the prices of stocks, and conversely to determine stock prices from state prices. We now give a sense for how this works by means of a simple example.

State Prices in the Stock Market. The general framework for determining state prices from stock prices and vice versa, and for specifying when a set of stocks is "rich enough" to be able to perform this determination, is complex. Here is a streamlined example that conveys the central idea.

Suppose there are two companies, named 1 and 2, whose stock is being traded. There are also two possible states, which we'll call s_1 and s_2. To be concrete, we can imagine that state s_1 is "Company 1 does well," and state s_2 is "Company two does well". Suppose that stock in Company 1 is worth one dollar in state s_1 and worth nothing in state s_2; and stock in Company 2 is worth one dollar in state s_2 and nothing in state s_1. Then the stocks are equivalent to the contracts traded in a prediction market and, just as we saw in a prediction market, their prices are the market probabilities of the states.

Now let's suppose, more realistically, that each stock is worth something in each state. Suppose that stock in Company 1 is worth two dollars in state s_1 and one dollar in state s_2; and stock in Company 2 is worth one dollar in state s_1 and two dollars in state s_2. If we know the state prices we can determine the price of each stock. Let's call these prices v_1 for stock in Company 1 and v_2 for stock in Company 2, and let's write ρ_1 and ρ_2 to denote the state prices for states s_1 and s_2. The price of a share of stock in Company 1 is the value now of the future worth of the company, which is $2\rho_1 + 1\rho_2$. Intuitively, this is because we can think of a share of stock in Company 1 as simply a "package deal" that offers two dollars in state 1 and one dollar in state 2; the price of this package is just the price of its

[5]Here, just as in the case of horse races we are taking the value of the stock in each event as being given exogenously. This is an important simplification, since in reality the value of the stock in each event is endogenous and is determined in the market equilibrium that arises in that state.

constituent ingredients, which if sold separately would require the purchase of two one-dollar contracts for state 1, at a price of $2\rho_1$, plus the purchase of a single one-dollar contract for state 2, at a price of ρ_2. (Of course, these contracts based on the states themselves are not sold separately; the point is that they only come implicitly "bundled" into the price of the stock.) Similarly, the price of a share of Company 2 is $1\rho_1 + 2\rho_2$.

Conversely, if we know the price of each stock we can determine the state prices by solving the system of equations

$$
\begin{aligned}
v_1 &= 2\rho_1 + 1\rho_2 \\
v_2 &= 1\rho_1 + 2\rho_2
\end{aligned}
$$

for the state prices ρ_1 and ρ_2. The solutions are

$$
\begin{aligned}
\rho_1 &= \frac{2v_1 - v_2}{3} \\
\rho_2 &= \frac{2v_2 - v_1}{3}.
\end{aligned}
$$

With these examples in mind we can now get some idea of what is meant by "a rich enough set of stocks". Essentially, we need a set of stocks such that when we write down their prices as functions of the underlying state prices, we get a system of equations like the one above that can be solved by a unique set of state prices. If we can do this, then stock prices determine state prices and investors will be able to use trades in stocks to move money across states in any way that they want. Indeed, in this case, we can essentially imagine that there is a big prediction market with a contract for each state, find the equilibrium state prices, and then determine stock prices from these state prices.

The conclusion of our analysis here is that stock markets, prediction markets, and betting markets are all essentially the same. They each give individuals the opportunity to place bets, and they all produce prices which can be interpreted as aggregate predictions about the likelihood of future states. This point of view also provides us with some intuition about what causes prices to change: They change if the distribution of wealth changes (and individuals' beliefs differ) or if individuals' beliefs about the probability of states change. If individuals suddenly believe, for whatever reason, that states with high payoffs are less likely, then prices will fall. Conversely, if individuals become more optimistic then prices will rise. Of course, this leaves us with the question of why individuals' beliefs change. One possibility is that they are learning about the likelihood of states using Bayes' Rule.[6] If their observations cause them to predict a less optimistic probability, then prices will fall. If they are in an environment in which information cascades can occur, then as we saw in Chapter 16, even small events may be able to cause large changes in predictions.

[6] This topic is explored in more detail in Section 22.10 at the end of this chapter.

22.5 Markets with Endogenous Events

As we noted at the beginning of this chapter, sometimes the events that individuals have beliefs about are endogenous — that is, whether they come true depends on the aggregate behavior of the individuals themselves. To take an example from our discussion of network effects, if no one expects anyone else to join a particular social-networking site, then no one expects a positive payoff from joining the site; consequently, no one joins and the payoff to joining is indeed low. Alternately, if many people expect a large membership at the social-networking site, then they expect a large payoff; as a result, many people join and the payoff is indeed high.

Here's a different example, from a setting in which there is a market with buyers and sellers. Suppose that people expect used cars offered for sale to be of uniformly low quality. Then no one will be willing to pay a high price for a used car. As a result, no one with a good used car will offer it for sale (since it would get a price below what it's worth), and so in fact the market will contain only low-quality used cars for sale. On the other hand, if people expect used cars to have some reasonable average quality, then they may be willing to pay a price high enough to induce sellers with both good and bad used cars to put them on the market.

An important common theme in these two stories is the notion of *self-fulfilling expectations* — and in particular, the presence of multiple different self-fulfilling expectations equilibria. This concept was central to our discussion of network effects in Chapter 17, and we see it again in the example of the used-car market. With one set of expectations, the world turns out in a certain way that makes the expectations come true; but with a different set of expectations, the world would have turned out in a different way that would have made *those* expectations come true.

Asymmetric Information. There is, however, an important difference between these two stories. In the case of the social-networking site, it seems reasonable to suppose that most people have similar information about the payoffs to joining the site. This information may be more or less accurate, but there is no *a priori* reason to suppose that some large fraction of the population is intrinsically more well-informed than some other large fraction. In the case of the used-car market, however, each seller of a used car knows something about his or her own car — its quality — that potential buyers do not know. This is an inherent feature of how the market works: there is *asymmetric information.*

In Chapter 17 we studied self-fulfilling expectations equilibria in settings without asymmetric information. In the rest of this chapter, we add asymmetric information to the picture; this turns out to be a basic ingredient in the way that beliefs about endogenous events can manifest themselves in markets. There is a fundamental reason for this: in many settings where buyers and sellers interact, one side of the market has better information about the goods or services being traded than the other side does. In the market for used cars, sellers know more than buyers do about the sellers' cars. On electronic markets for goods, such as eBay, sellers often know more than buyers do about the goods they are offering for sale. In the market for health insurance, on the other hand, buyers of insurance often know more than sellers do about the value of the good (health insurance) being purchased, since a buyer may well know more about his or her inherent health risks than the company offering the insurance does. In the stock market, either side of a transaction could have information about the future value of the stock that is not known to the other side of the transaction (a feature that we ignored in our earlier discussion of the stock market). In all of these cases, uninformed traders need to form expectations about the value of the good being traded, and these expectations should take into account the behavior of better-informed traders.

22.6 The Market for Lemons

At the beginning of the chapter, we started with a simple scenario involving horse-racing and then showed how the resulting principles extended to much larger and more complex systems such as the stock market. For considering the role of asymmetric information, we'll follow a similar strategy, first developing the case of the used-car market as a simple, stylized example, and then showing how the same principles apply to a range of more complex and fundamental markets.

In focusing first on used cars, we're following the rhetorical lead of the economist George Akerlof, who published a foundational paper on asymmetric information [9] for which he shared the 2001 Nobel Prize in Economics. His leading example in the paper was the market for used cars — or, as he called it, the "market for lemons." (A used car that is particularly bad is called a lemon.) The idea behind this phrase is old, probably as old as trading itself, but Akerlof was the first to clearly articulate the underlying principle and its implications for how markets work — or, in some cases, how they fail to work. Once we develop the basic ideas for the example of the used car market, we will then discuss how to apply these ideas to other markets.

Let's suppose that there are two types of used cars: *good cars*, and *bad cars*. Each seller knows the quality of his or her own car. Buyers do not know the quality of any seller's car, but they are aware of the fact that sellers know the quality of their own cars. Market participants — the buyers and sellers — value used cars differently. To keep the analysis simple we will pick the following specific values for used cars.

- Sellers value good cars at 10 and bad cars at 4. (We can imagine these as multiples of a thousand dollars, for example.) These values can be interpreted as sellers' reservation prices for their cars. That is, a seller with a good car would be willing to sell it for a price of at least 10, but at any lower price would prefer to hold onto it. Similarly, a a seller with a bad car would be willing to sell it for a price of at least 4, but at any lower price would prefer to hold onto it.

- Buyers value good cars at 12 and bad cars at 6. These values can be interpreted as buyers' reservation prices for cars. Thus, a buyer would be willing to buy a car that is known to be a good car if and only if the price is no more than 12, and a bad car if and only if the price is no more than 6.

Note that we have assumed that for any type of car, all buyers place the same value on it and this value is more than the common value that all sellers place on it. This simplification is not necessary for our analysis, but it will make our example of market failure more striking.

Let's suppose that a fraction g of used cars are good cars, and hence a fraction $1 - g$ are bad cars. We'll also assume that everyone knows g. Finally, let's suppose that there are more buyers than used cars (or more buyers than sellers, since each seller has one used car).

The Market with Symmetric Information. As an initial baseline, let's consider the simple case in which the type of each car is known to everyone. In this case, since there are more buyers than sellers, every car could be sold to some buyer.

How would the market work under this assumption? We would expect to observe different prices for good cars and bad cars. The price of good cars will clearly be between 10 and 12 because only at prices in this range can all good cars be sold. Similarly, the price of bad cars will be between 4 and 6. Since there are more buyers than sellers, some buyers will not be able to purchase a car, and so we would expect prices of each type of car to be bid up to the upper limit of the range of prices. That is, the price of good cars will be 12 and the price of bad cars will be 6.

The Market with Asymmetric Information. But what happens if buyers cannot tell in advance of a purchase what type of car they are buying? Since cars are indistinguishable to buyers, there can only be one price for a used car — all cars that trade will have to trade at that price. Furthermore, because a buyer can't tell the quality of the car she is buying, the quality of the car she gets is random, based on the mixture of qualities on the market. Given this random aspect to the outcome, we need to consider how buyers evaluate risk, just as we did when considering bettors and horse races. To keep the analysis simple in the current case of the used-car market, we will assume that buyers do not care about risk: they simply evaluate the expected value of the used cars they are considering. We could also introduce utility functions that capture the notion that buyers are concerned with risk, as we did in the earlier parts of this chapter, but in the present case it would complicate the model without significantly changing our qualitative conclusions.

Let's consider what the market looks like in the case that buyers can't distinguish among different types of cars. First, the fraction of good cars in the population of used cars for sale is some number h. This fraction h may be the same as g — the fraction of good cars in the overall used-car population — but it might not, since not all sellers of good used cars will necessarily put their cars up for sale. Given this fraction h, the value that any buyer places on a used car is

$$12h + 6(1 - h) = 6 + 6h. \qquad (22.12)$$

Thus, in order for buyers to know how much they should be willing to pay for a used car, they need to have a prediction for the value of h.

This puts us into the domain of self-fulfilling expectations equilibria, similarly to what we saw in discussing network effects in Chapter 17 (but here with the added issue of information asymmetry). We will look for a shared expectation h by the buyers that is self-fulfilling, in that if each buyer expects a fraction h of the cars on the market to be good, then indeed an h fraction of the cars on the market will be good.

Characterizing the Self-Fulfilling Expectations Equilibria. One candidate for an equilibrium of this form is $h = g$. This would be a correct prediction by the buyers if all sellers indeed choose to offer their car for sale. If this occurs, then we can plug $h = g$ into Equation (22.12) and see that buyers would be willing to pay $6 + 6g$ for a car. Let's call this price p^*. For the prediction $h = g$ to be correct, it must be the case that at a price of p^* both sellers who own good cars and sellers who own bad cars offer them for sale. A seller with a good car would offer it for sale at p^* provided that

$$p^* = 6 + 6g \geq 10,$$

or, equivalently, if $g \geq 2/3$. It is easy to see that if a seller with a good car would sell it at p^* then a seller with a bad car would also be happy to make a sale. So, if $g \geq 2/3$, there is a self-fulfilling expectations equilibrium in which all cars are offered for sale.

Now let's consider what happens if $g < 2/3$. Could there be a self-fulfilling expectations equilibrium in which $h = g$ — that is, in which all cars are offered for sale? We can analyze this as follows. When $g < 2/3$, the price that buyers would be willing to pay if they believe that all cars will be offered for sale, using Equation (22.12), is

$$p^* = 6 + 6g < 10.$$

However, owners of good cars would not be willing to sell for a price below 10, and so when $g < 2/3$ they will keep them off the market — meaning that h, the fraction of good cars on the market, would not be equal to g. So in this case, there cannot be a self-fulfilling expectations equilibrium in which $h = g$.

However, for any value of g, there is always a self-fulfilling expectations equilibrium in which $h = 0$ — that is, in which only bad cars are sold. To see why, note that if buyers expect there to be only bad cars on the market, then they are willing to pay 6 for a car.

At this price, sellers of bad cars would be willing to sell, but sellers of good cars would not, and so the market would consist only of bad cars. Thus, this is a self-fulfilling expectations equilibrium with $h = 0$.

So to summarize, the value $g = 2/3$ is a critical point in this example. If $g \geq 2/3$, there are two possible self-fulfilling expectations equilibria: one in which all cars are sold, and one in which only bad cars are sold. If $g < 2/3$, on the other hand, then the only equilibrium is the one in which only bad cars are sold. In this latter case, the abundance of bad cars, combined with buyers' inability to distinguish good cars from bad cars, has driven the good cars out of the market.

Complete Market Failure. Our example with good and bad cars illustrates the basic idea of how equilibria with asymmetric information work, but it doesn't fully capture the possible extent of market failure, or how bad the effect can get. To explore this, let's consider an example in which there are now three types of used cars: good cars, bad cars, and *lemons*. Good cars and bad cars still play the same basic role as in our previous example, whereas lemons are completely worthless to both buyers and sellers: a market in which only lemons are offered for sale is in fact not a functioning market at all, since it consists only of opportunities to trade items of value 0.

For this example with three types of used cars, let's suppose:

- One-third of the cars are good, one-third are bad and one-third are lemons.
- Sellers value good cars at 10, bad cars at 4, and lemons at 0.
- Buyers value good cars at 12, bad cars at 6, and lemons at 0.
- There are more buyers than there are used cars.

So if there were complete information, we would expect all good and bad cars to be sold at prices of 12 and 6 respectively, since there are more buyers than sellers, and the buyers value each type of car at least as much as the sellers do, Whether the lemons are sold is a question of buyer and seller indifference, since they are worth 0 to everyone.

But with asymmetric information, we need to consider what the possible self-fulfilling expectations equilibria are. There are three candidates for an equilibrium: (a) all cars are offered for sale; (b) only bad cars and lemons are offered for sale; or (c) only lemons are offered for sale. Again, note that option (c) represents the complete failure of the market, since all items on the market would have value 0. Let's consider each of these in turn and see which are actually possible.

(a) First, suppose buyers expect all cars to be on the market. Then the expected value of a car to a buyer would be
$$\frac{12 + 6 + 0}{3} = 6.$$

This is less than the value that sellers of good cars places on their cars, and so they would not put them on the market, meaning that this expectation would not be borne out by what happens. Hence this is not an equilibrium.

(b) Alternately, suppose buyers expect bad cars and lemons to be on the market. Then the expected value of a car to a buyer would be

$$\frac{6+0}{2} = 3.$$

But this is less than the value that sellers of bad cars places on bad cars, and so they would not put them on the market, meaning again that this expectation would not be borne out by what happens. So this too is not an equilibrium.

(c) Finally, as in our previous example with two types of cars, it is clearly an equilibrium if buyers expect only lemons to be sold. In this case, their expected value for a car is 0, and if this is what they are willing to pay, then the market will consist completely of lemons.

Notice how the market has been subverted by a kind of chain reaction: good cars can't survive on the market because of the frequency of bad cars and lemons; and even without the good cars, the bad cars can't survive on the market because of the frequency of lemons. It is not hard to produce this effect with even larger numbers of types of cars: things can chain together so that, in Akerlof's words, "it is ... possible to have the bad driving out the not-so-bad driving out the medium driving out the not-so-good driving out the good in such a sequence of events that no market exists at all" [9].

Summary: Ingredients of the Market for Lemons. In the next section, we'll take the lessons from our used-car examples and apply them to markets that are much larger and more fundamental. To do this, it's useful to review the key features of the current examples that led to market failure:

(i) The items that can be offered for sale have varying qualities.

(ii) For any given level of quality, the buyers value the items of that quality at least as much as the sellers do – so with complete information, the market would succeed in allocating items from sellers to buyers, potentially with different prices for different levels of quality.

(iii) There is asymmetric information about the quality of the items — only one side in a transaction can reliably determine the quality what is being sold. (In the used-car example, the seller in a potential buyer-seller transaction could tell the quality of what he was selling. In the next section, we'll also talk about other basic markets in which it is the buyer who has this power relative to the seller.)

(iv) Because of (iii), the items all must be sold for the same uniform price, and sellers will only put their items up for sale if they value them at or below this uniform price.

The market does not necessarily fail when these ingredients are present. It depends on whether there is an equilibrium where the buyers expect a mixture of qualities, and hence a price they are willing to pay, that induces the sellers to put their items up for sale. Market failure becomes more likely when the fraction of low-quality items is higher, and also when the difference between buyer and seller values is smaller.

In our discussion we have implicitly compared the market outcome with an outcome that can be achieved only with full information about each seller's car. But only the seller knows the value of his car, so any allocation procedure — not just a market — has to deal with this problem. Any procedure has to at least implicitly reward sellers for revealing their information, and this reward drives a wedge between the compensation needed to convince sellers to participate and the amount that buyers are willing to pay. Determining exactly which assignments are possible is complicated, but the optimal allocation that would be possible with full information cannot always be achieved.

22.7 Asymmetric Information in Other Markets

The ideas behind the market for lemons turn out to be fundamental to some of society's most important markets. Once you start thinking about interactions in which one party to the transaction knows something that the other party cares about, you realize that the phenomenon we're discussing is far from exceptional; in fact, it occurs all the time.

The Labor Market. One example where these ideas apply very naturally is to the labor market, in which people seeking jobs play the role of the sellers, and companies seeking employees play the role of the buyers. That is, we think of the process of employment as a market where people offer their skills for sale to employers, who pay them wages in return. Let's consider the basic assumptions of the market for lemons — numbered (i)–(iv) at the end of the previous section — in the context of the labor market.

(i) There are different qualities of workers — some are very productive while others are less productive, and this affects the value they will produce for the company that hires them.

(ii) It is natural to consider a setting where there are different kinds of jobs at different levels of wages, and where companies would be willing to hire any given prospective employee if they could accurately determine which jobs and wage levels were appropriate for them.

(iii) There is asymmetric information: a person generally has a better sense for how productive they are than a prospective employer does.

(iv) If we take a strong but plausible version of (iii), where employers can't reliably determine the quality of the people they are hiring, then employers can't hire only productive workers and wages can't directly depend on the quality of the person being hired. Rather, a uniform wage will be offered, and only applicants who believe this wage acceptably values their skills will take the job.

In this analysis we are assuming that although workers have differing productivities, each individual's productivity is a fixed, given amount, and not affected by anything that the worker chooses to do. It is plausible that workers can affect their productivity by varying the amount of effort that they put into their job, but we will ignore this issue for the sake of the present formulation. Thus, the key issue is point (iv) above, which — as in the case of the used-car market — can be viewed as a problem of *adverse selection.* The firm cannot select for a population consisting only of high productivity workers; instead, if it hires any workers at all, the only thing it can be sure of is getting those with low productivity.

It's useful to work through the consequences of information asymmetry in the labor market through a simple example whose structure closely parallels our used-car example. Suppose a firm hires workers from a large pool of potential employees. Suppose further that workers come in two types, productive and unproductive, and that half the workers in the population are of each type. Each productive worker hired by the firm will generate \$80,000 of revenue per year for the firm, while each unproductive worker will generate \$40,000 of revenue per year for the firm.

Each worker knows his own type. Also, each worker could choose not to work for the firm, and instead generate an alternate income by being self-employed. Workers who are more productive will get more value from being self-employed: suppose that each productive worker could produce an income of \$55,000 per year through self-employment, while each unproductive worker could produce an income of \$25,000 per year through self-employment. So if the firm could accurately determine the type of each job applicant, the situation would be straightforward: the firm could offer a salary between \$55,000 and \$80,000 to each productive applicant, and a salary between \$25,000 and \$40,000 to each unproductive applicant, all job offers would be accepted, and both workers and the firm would benefit from each job that is taken at the firm.

Unfortunately for the firm, it cannot reliably determine the type of each worker. So the firm has to offer a uniform wage of w, and simply hire workers who are willing to work at wage w. The firm is willing to offer a given wage w if and only if the average revenue it receives from the workers it hires at this wage is at least w.

Equilibria in the Labor Market. In our example, what wages can be offered, and which workers will be willing to work for the firm? The reasoning is very similar to the case of used cars. We start by looking for a self-fulfilling expectations equilibrium. If the firm expects all workers to be on the job market, then — since there are equal numbers of the two types of workers — its expected revenue per employee will be

$$\frac{80,000 + 40,000}{2} = 60,000,$$

and hence it can offer a uniform wage of \$60,000 per year. At this wage, both types of workers will be willing to accept the firm's offers, and so the expectation will be confirmed by what happens — we have an equilibrium in which all workers are hired.

By analogy with the used-car example, there is also another — less socially desirable — equilibrium. If the firm expects only unproductive workers to be on the job market, then it expects to make only \$40,000 per year per employee, and so this is the maximum wage it will offer. At this wage, only unproductive workers will be willing to accept jobs, and so again the firm's expectations are confirmed. So there are two possible equilibria here — a high one and a low one, with different mixtures of workers in the applicant pool in the two equilibria. Essentially, the firm's *a priori* level of confidence in the quality of its job applicants is self-fulfilling.

Things change if we shift the relative fractions of productive and unproductive workers in the population. Suppose that only 1/4 of the workers are productive and 3/4 are unproductive. There is still an equilibrium in which only unproductive workers are hired, but is there also one in which all workers are hired? If the firm were to expect all workers to be on the market, then its expected revenue per employee would be

$$\frac{1}{4} \cdot 80,000 + \frac{3}{4} \cdot 40,000 = 50,000,$$

and so this is the highest wage it could offer. But at this wage, the productive workers wouldn't be willing to accept the firm's offers, and so in fact not all workers would be on the market. In other words, there is no equilibrium in which the productive workers apply for jobs at the firm — just as with good used cars, they have been driven out of the market by the high frequency of unproductive workers.

The Market for Insurance. There are many markets that we can analyze in a similar fashion. For example, asymmetric information plays an important role in the market for health insurance. Health insurance companies generally know much less about the health of those they insure than the insured know about their own health. Health insurance companies are very good at predicting the average cost of insuring a pool of people, but it is difficult for them to predict the cost of insuring any particular individual. They group people into risk categories based on their medical history, but within any group each individual knows more about his or her own history, and about how he or she will behave in the future, than the insurance company knows.

So we have all the ingredients of the market for lemons: individuals in a given risk category can be more or less costly to insure, but the insurance company cannot reliably make these fine-grained distinctions. We should also notice an interesting twist in the case of health insurance: it is the *buyers* of health insurance, rather than the sellers, who have the additional information. But the consequence is the same. For any risk category, the insurance company has to essentially charge a uniform price for the insurance that is sufficient to cover the average cost of providing health care for the group. This means that the healthiest individuals in the group are being charged a price that is greater than the expected cost of providing care for them, and so they may be unwilling to buy insurance. Then, because these relatively healthy people do not participate, the average quality of the remaining pool goes down; the insurance company would need to set a higher price for this less healthy pool. Now the healthiest people in this remaining pool may be unwilling to pay this higher price, they too may chose not to buy insurance, and the average quality in the pool goes down further. As in the case of the used-car market, the market for health insurance can unravel to the point that no one buys insurance. Of course, whether this actually happens depends on the actual numbers: how much it costs to provide the insurance, and how much people value the insurance compared to their alternatives. But just as in our earlier examples, we see how socially undesirable outcomes can occur in the market when there are imbalances in information.

The information asymmetry we have focused on in the market for health insurance leads, just as in the cases of used cars or employment, to a type of adverse selection. Insurance companies cannot select for a population consisting only of healthy individuals; rather, if anyone buys insurance at all, the only thing one can be sure of is that it will be bought by those who are less healthy. There is another type of information asymmetry that occurs in the market for health insurance that we have so far ignored in our discussion. As in the previous examples, we have treated the health status of each individual, and thus his or her cost to insure, as fixed and given. But individuals can take actions which affect their health. If these actions are not observable to the insurance company then we have a new source of information asymmetry, since each individual knows more about his future behavior than the insurance company does. Once an individual purchases health insurance, his incentive to undertake costly actions to maintain his health is reduced, since he no longer bears the full cost of poor health. This introduces an effect known as *moral hazard*: when you're shielded from the full cost of your potential bad behavior, you have less incentive to avoid engaging in it.

Information Asymmetry in Trading and the Stock Market. It is useful to reflect further on these examples in light of one of the basic lessons of this chapter: that in any trade, each trader should ask why whoever is on the other side of the trade wants to make the trade. As we noted at the beginning of this chapter, if one trader is buying then the other trader is selling, and vice-versa. So the actions of the two traders are exactly the opposite of each other. Understanding the motivation behind the other trader's action may be crucial to understanding whether the trade is actually a good idea. For example, in the used-car market, a buyer should ask why any seller wants to sell. The same question can be asked by sellers when they are at a potential information disadvantage: for example, companies selling health insurance cannot be sure exactly why any one individual is seeking to buy their insurance.

All these issues play an important role in another market we discussed earlier in the chapter — the market for financial assets such as stocks or bonds. Here too, for every buyer there is a seller, and each should be curious about the other's motivation. A seller of a stock could be selling because of a desire to adjust their portfolio or a need for cash. A seller of a stock could be selling because their opinion differs from the opinion reflected in the market price (the market belief), even though they do not have private information. Alternatively, a seller could be selling because they have some private information that suggests that the price of the stock will fall in the future. Similarly, a buyer could want to buy because they have extra cash to invest, because they simply happen to have a different belief about the market, or because they are taking advantage of information that suggests the price of the stock will rise in the future.

When one side in such stock trades has better information, the other party would value the stock differently if they too had the information. Determining what the other party to the transaction knows is often impossible, but understanding that sometimes the other party knows something is not impossible. Once each party to the transaction takes this into account it is possible that no trade occurs, just as we saw with the example of used cars [299].

22.8 Signaling Quality

Given how powerfully information asymmetry can affect the operation of a market, it is natural to consider methods for alleviating it. One fundamental approach, useful in a number of settings, is to create a kind of certification mechanism: a way for a seller to provide a signal about the quality of the good that he or she is offering for sale.

To return to the case of used cars, for example, we can identify a variety of such possible signals. One that dealers sometimes offer is a guarantee that a given car is a "certified used car." Dealers certify that these cars have been checked for a number of possible defects and that any problems have been repaired. Another signaling mechanism is to offer a warranty promising that if the car needs to be repaired during some period after the sale, then the seller will pay for, or provide for free, the needed repair. Both of these quality assurances are directly valuable to buyers, but their value is more than you might imagine. It is less expensive for sellers who have good cars to provide these guarantees or warranties than it is for sellers who have bad cars. Either fewer repairs are needed before the car is sold or fewer repairs are expected to be needed after the car is sold. If it is too expensive for sellers of bad cars to provide these signals, then only good cars will have the signals — or at least, applying a milder form of this reasoning, the population of cars with these signals will contain a higher proportion of good cars compared to the population as a whole. Thus, buyers can make inferences about the quality of the car from the existence of the signal. These inferences raise the expected value of the car to the buyer even more than the direct value of the completed or promised future repairs.

Thus, the overall system of warranties might be crucial for breaking down information asymmetries that could otherwise cause the market to fail.

Signaling in the Labor Market. This idea of signaling applies to many settings other than just the used-car market. Perhaps its most important application is to the labor market, in which education can serve as signal; Michael Spence developed this idea and shared the 2001 Nobel Prize in Economics (with George Akerlof and Joseph Stiglitz) for his work on this topic [379].

Spence's idea is easy to understand in the context of our earlier labor market example, where firms cannot initially distinguish productive workers from unproductive ones. Suppose that it is easier for productive workers to obtain education than it is for unproductive workers. (Perhaps productive workers also perform better in school, and they can obtain a degree with less effort.) In this case, education provides a credible signal of productivity, and employers would be willing to pay higher wages to workers with more education than to workers with less education.

Notice that this signaling mechanism works even if education has no direct effect on a worker's productivity. Of course, education is also intrinsically valuable, but when we take information asymmetry into account, we see that education has a kind of two-fold power in the market. It trains workers for future employment; but beyond this, it also reduces information asymmetry about worker quality in a way that can potentially be necessary for the labor market to function effectively at all.

22.9 Quality Uncertainty On-Line: Reputation Systems and Other Mechanisms

Once we adopt the perspective that the availability of information is crucial in many markets, we can begin to appreciate that many of the standard mechanisms used in Web sites for on-line commerce are in fact motivated by considerations of asymmetric information and signaling. In this section we will discuss two of these mechanisms: reputation systems, and the role of ad-quality measures in sponsored-search advertising.

Reputation Systems. One of the clearest examples of these ideas at work in an on-line setting is the development of *reputation systems* for sites like eBay [171]. Since eBay is designed to facilitate trade between arbitrary people who have never met and may never meet again, a buyer faces a risk that he is dealing with a bad seller (like a seller of a lemon) who will provide an item of lower quality than advertised, or possibly fail to provide a promised item at all. Thus we have a situation that corresponds closely to the market for lemons: if buyers believe that the chance of receiving bad products (or of being cheated outright) is too high, then the price they will be willing to pay for an arbitrary item on eBay will be so low that no seller of reasonable items will want to participate. In this case, eBay's market could fail completely.

Reputation systems are a kind of feature provided by Web sites like eBay to offer a certification mechanism for alleviating this problem. After each purchase, the buyer can provide an evaluation of the seller, reporting whether the transaction and the item they received met the expectations that were conveyed. The evaluations received by a seller are synthesized by an algorithm at the heart of the system to provide an overall *reputation score* for the seller. A seller's reputation score evolves over time: favorable evaluations cause the score to go up while unfavorable ones cause it to go down. Thus, a good reputation score serves as a signal — in principle, it is costly to obtain, since it requires engaging in a sequence of transactions that cause the respective buyers to be satisfied. If it's cheaper for good sellers to acquire a good reputation than it is for bad sellers to acquire the same reputation, then reputation can serve as a signal of seller quality, just as a seller certifying his used car or a worker paying for education serves as a signal. In this way, if a site like eBay can convince buyers to have confidence in the reliability of the reputation system, then the resulting scores can reduce some of the strong information asymmetries inherent in the site.

There are many challenges in creating a reputation system that functions effectively, and a number of these challenges arise from the on-line nature of the application itself [171]. In particular, participants on a site like eBay can generally create multiple identities by registering for multiple user accounts on the system, and this leads to several approaches for subverting the goals of the reputation system. First, a seller interested in misbehaving can build up the reputation of a particular identity so that buyers will trust it, then behave badly until its reputation gets seriously damaged, discard the identity in favor of a freshly created one, and start the process again. In other words, the reputational consequences of bad behavior can be mitigated on-line if there is an easy way to "start over" by simply registering a new identity on the site. This ability to start over adds a severe moral-hazard feature to the on-line transaction problem, just like the ability of an individual to affect his health status adds a moral-hazard component to health insurance. This makes the problem of creating a reliable reputation system more difficult than it would be if there were only an adverse-selection problem. In addition, the design of a reputation system is further complicated by the potential for other kinds of misleading seller behavior. In particular, a seller can operate several identities simultaneously, and have the different identities engage in transactions with one another purely for the purpose of having them lavish positive feedback on each other. The seller can thereby obtain identities with high reputation scores despite no genuine history of good behavior.

In spirit, these types of strategies are reminiscent of what we saw in our discussion of link analysis for Web search — they are extreme versions of the general principle that when people's behavior is being evaluated by an algorithm, we should expect that many people will react and adapt to the criteria of the algorithm in ways that benefit them. Designing reputation systems that are robust in the presence of these kinds of difficulties is an ongoing research question.

Ad Quality in Keyword-Based Advertising. The ideas behind the market for lemons also show up clearly in the systems that search engines use for keyword-based advertising, and in fact this makes for an interesting case study in how these ideas have influenced a large on-line market. Specifically, we talked in Chapter 15 about the problem of ad quality: how the ranking of an ad on a page should not be based purely on the bid offered by the advertiser, but also on an estimate of the true clickthrough rates that this ad will have in a given position, compared to other ads. Otherwise, an unappealing ad based on a high bid could end up clogging the top slot on the page, generating very little revenue for the search engine because almost no one clicks on it.

But when you look at how the search industry actually runs the market for advertising, you quickly appreciate that the notion of "ad quality" is not just a proxy for the estimated rate of clicks the ad will get — it is a more subtle concept that is based on a broader estimate of overall user satisfaction with the ad. A common scenario here is as follows. There can be an advertiser that bids very highly for an ad on a certain query, *and* this ad has enticing text that causes it to get clicked on at a rapid rate by users from the search results page. A high bid per click multiplied by a high rate of clicks generates significant revenue for the search engine. However, the actual page the ad links to (the *landing page* that users reach when they click on the ad) is of low quality — not fraudulent, just not actually very relevant to the query with which the ad is associated. Consider for example an advertiser that places a high bid on an ad for the query "Caribbean vacations," and includes ad text on the Google results page saying "Dream vacations here" — but when you click on the ad, you get to a page that's trying to rent vacation properties in some completely different part of the world. It would be natural for most users to be disappointed when they click on this ad.

In such scenarios, the current strategy of search engines is to significantly lower the placement of such an ad on the page, or not to display it, *even though* this apparently causes them to lose the high rate of price-per-click revenue that the ad would generate in a high position. Their reason for this is the following: If users learn from experience that clicking on ads often takes them to low-quality landing pages, then they won't click on ads as much in general, and this overall effect on user behavior will have a huge negative effect on revenue in the long run. Essentially, the short-term gain in revenue from high-clickthrough low-quality ads is being traded off against the long-term losses due to user perceptions of quality.

The problem of asymmetric information is a fundamental issue behind this trade-off, and in fact the market for search advertising exhibits the basic ingredients of the market for lemons. Although clicking on a single search ad is a much less significant action than purchasing a car or hiring a new employee — as in our earlier examples of the used-car market and the labor market — it is still an activity that a user (the buyer) will undertake only if she believes that what she will find at the other end of the resulting link (the item being offered by the seller) is worth her effort. And just as a buyer can't tell the true quality of a used car until she purchases it, a user can't tell how well the ad text reflects the true quality of the landing page until she clicks on the ad. In this respect, advertisers have more information about the quality of their landing pages than the users of search engines do, and users implicitly form a mental estimate of how much they expect the ad text to reflect the quality of the landing page.

Notice, therefore, that while the analogy to the market for lemons is quite natural, it is also a bit subtle. In particular, it is not about the relationship between the advertisers and the search engine (though one can look for information asymmetry there too), but between the users and the advertisers, with user effort in clicking on ads as the quantity being valued. In aggregate, of course, all these user decisions to click are crucial, since they add up to a large portion of the search industry's revenue.

Now, we have seen in our earlier analyses that there can be multiple self-fulfilling expectations equilibria in these types of markets: some where buyers predict high average quality and there are in fact high-quality items for sale; and some where buyers predict low average quality and only low-quality items are on the market. Such equilibria are based on the assumption that consumers make correct predictions, which makes sense in the case of search advertising if users have time to learn the distribution of ad quality. Since the search engines have control over the ads they display, they are trying to maintain a mixture of ads of reasonable quality, thereby selecting an equilibrium for the overall market in which users expect high-quality ads, and advertisers with high-quality content are correspondingly willing to advertise via search engines.

22.10 Advanced Material: Wealth Dynamics in Markets

When we considered markets for assets such as stocks, shares in a prediction market, or bets in a horse race, we observed that market prices serve to aggregate the beliefs of the market participants — essentially, the market produces a weighted average of the participants' beliefs, with the weights determined by the participants' relative shares of wealth. Now, if we were to watch the market as it runs over time, certain participants would do better than others, their wealth shares would increase, and as a result their overall effect on the aggregate market price would increase. If we expect that people with more accurate beliefs will do better in the market, then this re-weighting as wealth shifts toward them should in fact produce more accurate market prices.

This intuition about market evolution over time was developed in the writings of a number of economists in the mid-20th-century [11, 157, 172]. The basic argument is that markets impose a kind of "natural selection" favoring traders whose decisions are closest to optimal. Early writers used this idea to argue that one tends to find rational investors in markets (because the others will have been driven out of the market) and that stock markets tend to be efficient (because prices are determined by the traders who have survived over the long run).

It has only been relatively recently, however, that this general idea has been explored more precisely, and its scope and limitations have begun to be understood. In this section, we describe a basic mathematical analysis that formalizes the intuition at the heart of these ideas [64]. The analysis will work by developing a close analogy between wealth dynamics in a market and the use of Bayes' Rule by an individual who learns over time. Recall from Chapter 16 that Bayes' Rule provides a systematic way to make use of new observations in decision-making. We will see that as wealth moves between participants in a market, their contributions to the aggregate market price change over time exactly like the probabilities on different hypotheses would change according to Bayes' Rule.

So in a precise sense, although the market is simply an institution that facilitates trade, it can also be viewed as acting like an artificially intelligent Bayesian agent that aggregates information. Moreover, if there is a set of traders who have correct beliefs, then over time their fraction of the wealth converges to one, and the market price converges to reflect their (correct) beliefs. This provides a concrete expression of the general idea that markets for assets tend to work well at synthesizing the information held by groups of people.

A. Bayesian Learning in a Market

We begin the analysis by considering how a Bayesian learner — that is, someone applying Bayes' Rule — would update his beliefs over time in a market. Once we've done this, we'll draw an analogy to how the wealth of participants changes over time.

We discussed Bayes's Rule in Chapter 16, but here we'll cast it in the notation of this chapter, and also extend some of the conclusions. Since horse races have served as a useful example to suggest the phenomena at work in more complex settings like the stock market, we'll continue to use horse races in the discussion here. Thus, suppose that two horses A and B will run a race every week; suppose that the outcomes of these races are independent; and suppose that A wins each one with probability a (and hence B wins with probability $b = 1 - a$).

Now, our Bayesian learner does not know the values of a and b; rather, he wants to learn them over time by watching the outcomes of races. He begins with a set of N possible hypotheses for the probabilities, which we will denote by

$$(a_1, b_1), (a_2, b_2), \ldots, (a_N, b_N).$$

For now, let's assume in fact that one of these hypotheses is correct (although the learner does not know which it is); re-labeling them if necessary, we'll suppose that $(a_1, b_1) = (a, b)$.

The learner begins with a *prior probability* on each hypothesis; let f_n be the prior probability on hypothesis (a_n, b_n). We will assume that each prior probability f_n is greater than zero, indicating that the learner considers it to be a possible description of the true probability. Since these prior probabilities form an initial weighted average over the hypotheses, the learner's initial predicted probability of A winning is

$$a_1 f_1 + a_2 f_2 + \cdots + a_N f_N.$$

Now, suppose that T races are run, and we observe a sequence S of outcomes of these races, in which horse A wins a total of k times and horse B wins a total of ℓ times. Then using Bayes' Rule as in Chapter 16, we can compute the posterior probability of the hypothesis (a_n, b_n), conditional on the sequence S, as follows.

$$\begin{aligned}
\Pr\left[(a_n, b_n) \mid S\right] &= \frac{f_n \cdot \Pr\left[S \mid (a_n, b_n)\right]}{\Pr\left[S\right]} \\
&= \frac{f_n \cdot \Pr\left[S \mid (a_n, b_n)\right]}{f_1 \cdot \Pr\left[S \mid (a_1, b_1)\right] + f_2 \cdot \Pr\left[S \mid (a_2, b_2)\right] + \cdots + f_N \cdot \Pr\left[S \mid (a_N, b_N)\right]}.
\end{aligned}$$

Now, the probability of S given a hypothesis (a_n, b_n) is simply the probability it assigns to the sequence of wins in S: since horse A wins a total of k times and horse B wins a total of

ℓ times, this probability is just $a_n^k b_n^\ell$. Therefore, we have

$$\Pr[(a_n, b_n) \mid S] \quad = \quad \frac{f_n a_n^k b_n^\ell}{f_1 a_1^k b_1^\ell + f_2 a_2^k b_2^\ell + \cdots + f_N a_N^k b_N^\ell}. \tag{22.13}$$

Moreover, after this sequence S of observed outcomes, the learner's predicted probability on horse A is

$$a_1 \Pr[(a_1, b_1) \mid S] + a_2 \Pr[(a_2, b_2) \mid S] + \cdots + a_N \Pr[(a_N, b_N) \mid S]. \tag{22.14}$$

This is the sense in which the learner is Bayesian: as he observes outcomes, he updates his predicted probability according to Bayes' Rule.

Convergence to the Correct Hypothesis. Now, let's consider how the posterior probabilities of the different hypotheses fare as horse races are run over a long period of time. The easiest way to do this is to consider the ratios of these probabilities. After a sequence S of observed outcomes, with k wins by A and ℓ wins by B, the ratio of the posterior probability on hypothesis (a_m, b_m) to the posterior probability on hypothesis (a_n, b_n) can be computed simply by taking the ratios of the respective expressions given by Equation (22.13), and noticing that the two expressions have the same denominator:

$$\frac{\Pr[(a_m, b_m) \mid S]}{\Pr[(a_n, b_n) \mid S]} = \frac{f_m a_m^k b_m^\ell}{f_n a_n^k b_n^\ell} \tag{22.15}$$

We will be particularly interested in ratios that compare the correct hypothesis (a_1, b_1) to an alternate hypothesis (a_n, b_n):

$$\frac{\Pr[(a_1, b_1) \mid S]}{\Pr[(a_n, b_n) \mid S]} = \frac{f_1 a_1^k b_1^\ell}{f_n a_n^k b_n^\ell} \tag{22.16}$$

We will call this ratio $R_n[S]$. Taking logarithms yields something known as the *log odds ratio* of the two hypotheses, given the sequence of observed outcomes S:

$$\ln(R_n[S]) = \ln\left(\frac{f_1}{f_n}\right) + k \ln\left(\frac{a_1}{a_n}\right) + \ell \ln\left(\frac{b_1}{b_n}\right).$$

Now, let's divide both sides by the total number of observations T, obtaining

$$\frac{1}{T}\ln(R_n[S]) = \frac{1}{T}\ln\left(\frac{f_1}{f_n}\right) + \frac{k}{T}\ln\left(\frac{a_1}{a_n}\right) + \frac{\ell}{T}\ln\left(\frac{b_1}{b_n}\right). \tag{22.17}$$

We're interested in what happens as T goes to infinity, and in this case the right-hand side of this equation can be simplified as follows. The first term is just a fixed constant $\ln(f_1/f_n)$ divided by T, so it converges to 0 as T grows. To analyze the second and third terms, we observe, using the Law of Large Numbers, that k/T converges almost surely to the true probability of horse A winning, which is a, and ℓ/T converges almost surely to the true probability of horse B winning, which is b. So the entire right-hand side of Equation (22.17) converges almost surely to

$$a \ln\left(\frac{a_1}{a_n}\right) + b \ln\left(\frac{b_1}{b_n}\right) \quad = \quad a \ln(a_1) + b \ln(b_1) - [a \ln(a_n) + b \ln(b_n)] \qquad (22.18)$$

We'd like to know whether this limit is positive, negative, or zero, since that will let us reason about what's happening to the left-hand side of Equation (22.17). Here is how we can think about this limit. The first two terms have the form $a \ln(x) + (1-a)\ln(1-x)$, with $x = a_1$, and the third and fourth terms have this form as well, with $x = a_n$. But we know from Section 22.2, and specifically the discussion around Equation (22.5), that the expression $a \ln(x) + (1-a)\ln(1-x)$ is maximized when $x = a$, and it is strictly smaller than this maximum for all other values of x. Since $a_1 = a$, the sum of the first two terms therefore achieves this maximum, and since $a_n \neq a$, the third and fourth terms that are being subtracted off don't achieve the maximum. Therefore, the expression in (22.18) is strictly positive (since the first two terms outweigh the latter two), and so returning to Equation (22.17), we conclude that

$$\frac{1}{T} \ln(R_n[S]) > 0$$

almost surely as T goes to infinity.

It follows that as T goes to infinity, $\ln(R_n[S])$, and hence $R_n[S]$ itself, must be diverging to positive infinity. Moreover, this takes place for every $n > 1$ — that is, for every incorrect hypothesis. How can this happen? Each $R_n[S]$ is the ratio of two probabilities, so in order for one of the probabilities (on (a_1, b_1)) to become larger than all the others by an arbitrary factor, it must be that the probability on hypothesis (a_1, b_1) is converging to one while the probability on each of the others is converging to zero.

The conclusion from this analysis is that the Bayesian learner will, in the limit, assign a posterior probability of one to the correct hypothesis. Moreover, this means that his predicted probability on horse A, as computed by Equation (22.14), is converging to $a_1 = a$.

Convergence without a Correct Hypothesis. If we think about it, the analysis above in fact shows something stronger than we've claimed. In order for the learner to converge to a posterior probability of one on the hypothesis (a_1, b_1), it is not necessary that (a_1, b_1) actually be correct. We simply need that the expression in Equation(22.18) is positive for all competing hypotheses $n > 1$.

The interpretation of this stronger claim is usually expressed in terms of a notion of "distance" between hypotheses, as follows. For a given hypothesis (a_n, b_n), we define the *relative entropy* $D_{(a,b)}(a_n, b_n)$ between (a_n, b_n) and the true hypothesis (a, b) to be

$$D_{(a,b)}(a_n, b_n) = a \ln(a) + b \ln(b) - [a \ln(a_n) + b \ln(b_n)]. \qquad (22.19)$$

By our earlier observation about the maximization of $a\ln(x)+(1-a)\ln(1-x)$, we see that the contribution of the first two terms must always outweigh the negative effect of the third and fourth terms, and so $D_{(a,b)}(a_n,b_n)$ is always a non-negative number, and it is zero when $(a_n,b_n)=(a,b)$. We can therefore interpret the relative entropy as a non-linear measure of how far a given hypothesis is from the truth: smaller relative entropies indicate better agreement with the true hypothesis.

Going back to Equations (22.17) and (22.18), we see that even when (a_1,b_1) is not the correct hypothesis, the quantity $\ln(R_n[S])/T$ is converging almost surely to

$$D_{(a,b)}(a_n,b_n)-D_{(a,b)}(a_1,b_1).$$

Therefore, suppose that $a_1\neq a$, but that the hypothesis (a_1,b_1) is closer than any other (a_n,b_n) to the true hypothesis in relative entropy: that is,

$$D_{(a,b)}(a_1,b_1)<D_{(a,b)}(a_n,b_n)$$

for all $n>1$. Then just as before, we have

$$\frac{1}{T}\ln(R_n[S])>0$$

almost surely as T goes to infinity. And from this, we draw the same conclusion as before: that the posterior probability the learner places on (a_1,b_1) converges to one.

In other words, when no hypothesis is correct but some hypothesis is uniquely closest to the truth in relative entropy, a Bayesian learner will assign a posterior probability of one to this hypothesis in the limit.

B. Wealth Dynamics

We have now seen how a Bayesian learner aggregates information about events taking place in a market: the learner maintains a weighted average of the probabilities assigned by different hypotheses, updating the weights using Bayes' Rule. Earlier in the chapter, we saw that the odds computed by the market are also a weighted average, in that case an average of bettors' beliefs weighted by their wealth shares. As time runs forward, the weights in this weighted average are updated simply because the bettors are gaining and losing money. We now show that this updating works exactly the way Bayes' Rule does, which is why the aggregate behavior of the market itself can be viewed as that of a Bayesian learner.

Evolution of Wealth Shares. Let's use the framework of betting markets from Sections 22.2 and 22.3 (again as a stand-in for more complex settings like stock markets). There are N bettors; each bettor n has a fixed belief that horse A will win with probability a_n (and hence that horse B will win with probability $b_n=1-a_n$). Bettor n has an initial wealth of

w_n; if the total wealth of all bettors is w, then this corresponds to a share $f_n = w_n/w$ of the total wealth.

Now, horses A and B race each other in each of time steps $t = 1, 2, 3, \ldots$. At the start of each time step t, before the t^{th} race, the market determines odds $o_A^{\langle t \rangle}$ and $o_B^{\langle t \rangle}$ on horses A and B; note that the odds may be different in each step, and as we saw in Section 22.3, they may depend on who the bettors are and how much they are betting. Also at the start of each step t, each bettor n has a current wealth $w_n^{\langle t \rangle}$; he bets this wealth optimally given his beliefs (a_n, b_n). As we saw in Section 22.2, this corresponds to putting a bet of $a_n w_n^{\langle t \rangle}$ on horse A and a bet of $b_n w_n^{\langle t \rangle}$ on horse B. Consequently, bettor n's new wealth $w_n^{\langle t+1 \rangle}$ after this race is equal to $a_n w_n^{\langle t \rangle} o_A^{\langle t \rangle}$ if A wins, and it is equal to $b_n w_n^{\langle t \rangle} o_B^{\langle t \rangle}$ if B wins.

Let's consider two bettors m and n, with initial wealth shares f_m and f_n, and suppose that by step t, due to the results of their bets on the first $t-1$ races, their wealth shares are $f_m^{\langle t \rangle}$ and $f_n^{\langle t \rangle}$ respectively. Let's consider the two possible outcomes of race t.

- If horse A wins race t, the wealth of bettor m is multiplied by $a_m o_A^{\langle t \rangle}$ and the wealth of bettor n is multiplied by $a_n o_A^{\langle t \rangle}$. Therefore, in this case, the ratio of the wealth shares of m and n changes from $f_m^{\langle t \rangle}/f_m^{\langle t \rangle}$ to $a_m f_m^{\langle t \rangle}/a_n f_m^{\langle t \rangle}$. (Notice that the odds cancel out of this ratio, since they apply equally to both bettors.) In other words, the ratio is multiplied by a_m/a_n.

- If horse B wins race t, the wealth of bettor m is multiplied by $b_m o_B^{\langle t \rangle}$ and the wealth of bettor n is multiplied by $b_n o_B^{\langle t \rangle}$. Therefore, in this case, the ratio of the wealth shares of m and n changes from $f_m^{\langle t \rangle}/f_m^{\langle t \rangle}$ to $b_m f_m^{\langle t \rangle}/b_n f_m^{\langle t \rangle}$. In other words, the ratio is multiplied by b_m/b_n.

So we see that whenever horse A wins a race, the ratio of wealth shares of bettors m and n changes by a factor of a_m/a_n, while whenever horse B wins a race, the ratio of their wealth shares changes by a factor of b_m/b_n.

Suppose we apply these changes, starting from wealth shares f_m and f_n, over a sequence of races S in which A wins k times and B wins ℓ times. Then we end up with a ratio of wealth shares that's equal to

$$\frac{f_m a_m^k b_m^\ell}{f_n a_n^k b_n^\ell}. \tag{22.20}$$

The point is that this is exactly the same as Equation (22.15), describing the ratio of posterior probabilities that a Bayesian learner puts on the hypotheses (a_m, b_m) and (a_n, b_n), starting from prior probabilities of f_m and f_n. So the analogy is perfect: the wealth shares of the bettors evolve exactly like the posterior probabilities on hypotheses under Bayes' Rule. That is, the market treats each bettor as a hypothesis about the two horses, and in response to the outcome of the race it adjusts that bettor's wealth share in exactly the same way that a Bayesian learner would adjust the probability on the hypothesis.

We can draw two main conclusions from this.

- First, the inverse odds maintained by the market are computed from the wealth-share-weighted average of the bettors' beliefs, using Equation (22.9) from Section 22.3. This equation is parallel to the Bayesian learner's Equation (22.14) by which he determines the predicted probability on horse A. Hence, the market's inverse odds follow the results of Bayesian learning as well.

- Since the ratios of wealth shares evolve just as the posterior probabilities of hypotheses, we can conclude that if there is a unique bettor whose beliefs are closest in relative entropy to the correct probabilities (a, b), then in the limit the wealth share of this bettor will converge to 1. So the market is selecting for bettors with more accurate beliefs, where "accuracy" here refers to the bettor's distance from the truth in relative entropy. Combined with our previous observation about the inverse odds, we see that in the limit, the assets (i.e. the possible bets) are priced according to the most accurate information held by any of the market participants.

It is also important to note that in the special case when one of the bettors has correct beliefs, this bettor will acquire a wealth share of 1 in the limit, and the market will come to reflect the bettor's (correct) beliefs.

Extensions and Interpretations. We have kept the model very simple so as to make the calculations clear. But it is possible to extend the model to incorporate a number of further considerations.

First, we are assuming that the bettors have fixed beliefs and do not learn from observing the outcomes of races. This makes it easy to isolate the effect of wealth dynamics in the market, distinguishing it from the learning dynamics of individual participants. But while it is a bit messy, it is not particularly difficult to combine this analysis of wealth dynamics with Bayesian learning by the bettors. Second, our analysis assumes that each bettor reinvests his entire wealth in the market in each time step. However, this too can be extended to a model in which bettors must decide both how much to reinvest in the market, as well as how to allocate this investment across the different options [64].

Our overall conclusion, that the market selects for the trader with the most accurate beliefs, and asymptotically prices assets according to these beliefs, applies equally well in other settings such as prediction markets. Notice that the argument here about the performance of the market is not based on the benefits of averaging, as in our previous discussion of the "wisdom of crowds." Rather, in the analysis here, the crowd is exactly as smart as its smartest participant in the limit, since in the limit it is only this participant whose beliefs affect the market's predictions. As noted earlier, this idea draws on a long history of economic arguments for market efficiency based on natural selection [11, 157, 172], in which smarter traders come to hold an increasingly large fraction of the wealth in the market, and thereby exert an increasingly large influence on the market. The model here puts this intuition on more precise footing [64], and subsequent research has expanded on it in important ways [65, 362].

While these expanded models are too complex to describe in detail here, they relate in interesting ways to some of the issues from earlier in the chapter. First, and rather surprisingly, the more complex models show that the assumption of logarithmic utility, on which the model here is based, is in fact not important for the general conclusion about market selection. A more general and abstract analysis shows that we only need the assumption that traders are risk-averse — that is, their utility gain from increased wealth decreases as their wealth grows. The recent line of research also shows that these results apply to more complex markets, provided that there is a rich enough set of assets being traded. Intuitively, if there aren't enough traded assets, then there may not be enough ways for traders with better beliefs to take advantage of traders with worse beliefs, and thus the traders with worse beliefs may not get driven out of the market. The richness condition that is needed for stock markets is exactly the condition discussed in Section 22.4. The conclusion of this analysis is that if there is a rich enough set of assets traded in the stock market, then in the long run the market prices assets as correctly as possible given the collection of traders' beliefs that are made available to the market.

22.11 Exercises

1. Consider a betting market with two horses A and B and two bettors 1 and 2, as in Section 22.3. Let's suppose that each bettor has wealth w. Bettor 1 believes there is a probability of $1/2$ that horse A will win, and a probability of $1/2$ that horse B will win. Bettor 2 believes there is a probability of $1/4$ that horse A will win, and a probability of $3/4$ that horse B will win. Both bettors have logarithmic utility for wealth, and they each choose bets to maximize expected utility of wealth given their beliefs.

   (a) How much money should bettors 1 and 2 each bet on horses A and B respectively?

   (b) Find the equilibrium inverse odds on horse A and on horse B.

   (c) How much money will bettor 1 have after the race if horse A wins? How about if horse B wins?

2. Consider a betting market with two horses A and B and two bettors 1 and 2 as in Section 22.3. Let's suppose that each bettor has wealth w. Bettor 1's beliefs are (a_1, b_1) where the first number in the pair is bettor 1's probability of horse A winning the race. Both bettors have logarithmic utility for wealth. Bettor 1 chooses his bets to maximize

his expected utility of wealth using his beliefs, as in the chapter. Bettor 2, however, behaves differently; he believes that the inverse odds are the correct probabilities and he maximizes his expected utility using these inverse odds.

(a) Bettor 1's optimal bet on horse A is some function of his wealth and his beliefs. Let's call this function $f_1(w, a_1)$. Determine this function.

(b) Suppose that bettor 2 knows the equilibrium inverse odds on horse A, which we will call ρ_A. Bettor 2's optimal bet on horse A is some function of his wealth and the equilibrium inverse odds on horse A. Let's call this function $f_2(w, \rho_A)$. Determine this function.

(c) If we take Equation (22.8) from Section 22.3, applied to the betting rules in question, then we see that the equilibrium inverse odds on horse A must solve the equation

$$\frac{f_1(w, a_1)}{2w} + \frac{f_2(w, \rho_A)}{2w} = \rho_A.$$

Using this observation, find the equilibrium inverse odds on horse A.

(d) Now let's generalize this idea to many bettors. Suppose that most bettors are like better 2; they trust that the inverse odds are in some sense correct and they use them in deciding how to bet. Only a few bettors are like bettor 1; they have beliefs and they bet according to their beliefs. Would you expect the "Wisdom of Crowds" idea to be more or less likely to be true in this market than in a market in which each bettor has beliefs and bets according to their own beliefs? Does your answer depend on which bettors bet using inverse odds as their beliefs and which ones use their own beliefs? (Think about which bettors are more likely to have correct beliefs.)

3. Consider the model of the market for lemons. Suppose that there are three types of used cars: good ones, medium ones and lemons, and that sellers know which type of car they have. Buyers do not know which type of car a seller has. The fraction of used cars of each type is $\frac{1}{3}$ and buyers know this. Let's suppose that a seller who has a good car values it at \$8,000, a seller with a medium car values it at \$5,000 and a seller with a lemon values the lemon at \$1,000. A seller is willing to sell his car for any price greater than or equal to his value for the car; the seller is not willing to sell the car at a price below the value of the car. Buyers values for good cars, medium cars and lemons are, \$9,000, \$8,000 and \$4,000, respectively. As in Chapter 22 we will assume that buyers are risk-neutral; that is, they are willing to pay their expected value of a car.

(a) Is there an equilibrium in the used-car market in which all types of cars are sold? Explain briefly.

(b) Is there an equilibrium in the used-car market in which only medium quality cars and lemons are sold? Explain briefly.

(c) Is there an equilibrium in the used-car market in which only lemons are sold? Explain briefly.

4. Consider the model of the market for lemons from Chapter 22. Suppose that there are two types of used cars — good ones and lemons — and that sellers know which type of car they have. Buyers do not know which type of car a seller has. The fraction of used cars of each type is $\frac{1}{2}$ and buyers know this. Let's suppose that a seller who has a good car values it at \$10,000 and a seller with a lemon values the lemon at \$5,000. A seller is willing to sell his car for any price greater than or equal to his value for the car; the seller is not willing to sell the car at a price below the value of the car. Buyers' values for good cars and lemons are \$14,000 and \$8,000, respectively. As in Chapter 22 we will assume that buyers are risk-neutral; that is, they are willing to pay their expected value of a car.

   (a) Is there an equilibrium in the used-car market in which all types of cars are sold? Briefly explain.

   (b) Is there an equilibrium in the used-car market in which only lemons are sold? Briefly explain.

5. Consider the model of the market for lemons. Suppose that there are three types of used cars: good ones, medium ones, and lemons, and that sellers know which type of car they have. Buyers do not know which type of car a seller has. The fraction of used cars of each type is $\frac{1}{3}$ and buyers know this. Let's suppose that a seller who has a good car values it at \$4,000, a seller with a medium car values it at \$3,000 and a seller with a lemon values it at \$0. A seller is willing to sell his car for any price greater than or equal to his value for the car; the seller is not willing to sell the car at a price below the value of the car. Buyers' values for good cars, medium cars, and lemons are \$10,000, \$4,000 and \$1,000, respectively. As in Chapter 22 we will assume that buyers are willing to pay their expected value of a car. We will also assume that there are at least as many buyers as used cars.

   (a) Is there an equilibrium in the used car market in which all types of used cars are sold? If so, find some equilibrium price for used cars such that all used cars are sold, together with a brief explanation of why all cars are sold. If not, explain why not.

   (b) Now suppose that someone develops a way for sellers of good used cars to certify that their cars are good cars. All sellers of good used cars do this and they are no longer part of the general market for uncertified used cars, which now consists only of medium used cars and lemons in equal numbers. Is there an equilibrium in the market for these remaining, uncertified used cars in which both medium used cars and lemons are sold? If so, find some equilibrium price for used cars such that medium used cars and lemons are sold, together with a brief explanation. If not, explain why not.

6. Consider the model of the market for lemons. Suppose that there are two types of used cars, good ones and lemons, and that sellers know which type of car they have. Buyers do not know which type of car a seller has. The fraction of used cars that are good cars is g and buyers know this fraction. Let's suppose that a seller who has a good car values it at \$10,000 and that a seller with a lemon values the lemon at \$4,000. A seller is willing to sell his car for any price greater than or equal to his value for the car; the seller is not willing to sell the car at a price below the value of the car. Buyers values for good cars and lemons are, \$12,000 and \$5,000, respectively. As in Chapter 22 we will assume that buyers are risk neutral; that is, they are willing to pay their expected value of a car.

   (a) Suppose that you observe that used cars sell for a price of \$10,000. What can you say about the fraction of used cars that are lemons?

   (b) Suppose, instead that the fraction of used cars that are lemons is $g = 0.5$. What is the maximum selling price for used cars?

7. In this question we are going to examine how a tax on the purchase of used cars might affect the price and quantity of used cars traded. Suppose that there are two types of used cars: good ones and bad ones. Sellers of used cars know the type of car that they own. Buyers do not know which type of car any particular seller has. Buyers do know that there are good and bad used cars, and they know that of the 100 people who own used cars and are interested in selling their car, 50 have good cars and 50 have bad cars. Let's suppose that there are 200 possible buyers of used cars. [As in Chapter 22 we want to assume that there are more buyers than sellers to make the analysis straightforward.] A seller who has a good used car values it at \$8,000 and a seller who has a bad used car values it at \$3,000. A seller is willing to sell his car for any price greater than or equal to his value for the car; no seller is willing to sell his car for a price less than his value for the car. Buyers values for good and bad used cars are \$10,000 and \$6,000, respectively. As in Chapter 22 we will assume that buyers each want at most one used car and they are willing to pay their expected value for a used car.

   (a) Find all of the equilibria in the market for used cars. For each equilibrium provide the price of used cars and the number of used cars traded.

(b) Now suppose that the government places a tax of \$100 on the purchase of used cars. That is, anyone who buys a used car must pay a tax of \$100 on the purchase of the car. This effectively lowers the values that buyers place on any type of used car by \$100. Find all of the equilibria in the market for used cars.

(c) Now let's change the setup of the problem a bit so that there are three types of used cars: good ones, bad ones and lemons. There are 50 sellers with good cars, 50 with bad cars and 50 with lemons. Buyers and sellers values for good and bad used cars are the same as before. Everyone (both buyers and sellers) values a lemon at 0. There are still 200 buyers. (i) There is no tax on the purchase of used cars. Find all of the equilibria in the market for used cars. (ii) Now the government imposes a tax of \$100 on the purchase of used cars. Find all of the equilibria in the market for used cars.

8. A group of researchers have been investigating the quality and seaworthiness of five-year-old boats in the U.S. They are using a classification of the boats into five possible categories: excellent, good, medium, poor, and dangerous. They have concluded that there are no excellent five-year-old boats, and that most of these boats are of medium and/or lower quality. To conduct their study these researchers pretended to be potential buyers of five-year-old boats. They examined a very large number of five-year-old boats offered for sale both by private sellers (individuals) and by boat dealers. Based on the results of their study this group of researchers has concluded that there should be an investigation by the U. S. Coast Guard into the quality of these older boats. What concerns do you have about the methodology the researchers used in their study? Can you suggest an alternative approach that they might have used in order to draw a more careful conclusion about the actual quality distribution of five-year-old boats?