CHAPTER 9

AUCTIONS AND MECHANISM DESIGN

In most real-world markets, sellers do not have perfect knowledge of market demand. Instead, sellers typically have only statistical information about market demand. Only the buyers themselves know precisely how much of the good they are willing to buy at a particular price. In this chapter, we will revisit the monopoly problem under this more typical circumstance.

Perhaps the simplest situation in which the above elements are present occurs when a single object is put up for auction. There, the seller is typically unaware of the buyers' values but may nevertheless have some information about the distribution of values across buyers. In such a setting, there are a number of standard auction forms that the seller might use to sell the good – first-price, second-price, Dutch, English. Do each of these standard auctions raise the same revenue for the seller? If not, which is best? Is there a non-standard yet even better selling mechanism for the seller? To answer these and other questions, we will introduce and employ some of the tools from the theory of *mechanism design*.

Mechanism design is a general theory about how and when the design of appropriate institutions can achieve particular goals. This theory is especially germane when the designer requires information possessed only by others to achieve his goal. The subtlety in designing a successful mechanism lies in ensuring that the mechanism gives those who possess the needed information the incentive to reveal it to the designer. This chapter provides an introduction to the theory of mechanism design. We shall begin by considering the problem of designing a revenue-maximising selling mechanism. We then move on to the problem of efficient resource allocation. In both cases, the design problem will be subject to informational constraints – the agents possessing private information will have to be incentivised to report their information truthfully.

9.1 THE FOUR STANDARD AUCTIONS

Consider a seller with a single object for sale who wishes to sell the object to one of N buyers for the highest possible price. How should the seller go about achieving this goal? One possible answer is to hold an auction. Many distinct auctions have been put to use at one time or another, but we will focus on the following four standard auctions.[1]

[1] We shall assume throughout and unless otherwise noted that in all auctions ties in bids are broken at random: each tied bidder is equally likely to be deemed the winner.

- **First-Price, Sealed-Bid:** Each bidder submits a sealed bid to the seller. The highest bidder wins and pays his bid for the good.
- **Second-Price, Sealed-Bid:** Each bidder submits a sealed bid to the seller. The highest bidder wins and pays the second-highest bid for the good.
- **Dutch Auction:** The seller begins with a very high price and begins to reduce it. The first bidder to raise his hand wins the object at the current price.
- **English Auction:** The seller begins with very low price (perhaps zero) and begins to increase it. Each bidder signals when he wishes to drop out of the auction. Once a bidder has dropped out, he cannot resume bidding later. When only one bidder remains, he is the winner and pays the current price.

Can we decide even among these four which is best for the seller? To get a handle on this problem, we must begin with a model.

9.2 THE INDEPENDENT PRIVATE VALUES MODEL

A single risk-neutral seller wishes to sell an indivisible object to one of N risk-neutral buyers. The seller values the object at zero euros.[2] Buyer i's value for the object, v_i, is drawn from the interval $[0, 1]$ according to the distribution function $F_i(v_i)$ with density function $f_i(v_i)$.[3] We shall assume that the buyers' values are mutually independent. Each buyer knows his own value but not the values of the other buyers. However, the density functions, $f_1, \ldots, f_N$, are public information and so known by the seller and all buyers. In particular, while the seller is unaware of the buyers' exact values, he knows the distribution from which each value is drawn. If buyer i's value is v_i, then if he wins the object and pays p, his payoff (i.e., von Neumann-Morgenstern utility) is $v_i - p$, whereas his payoff is $-p$ if he must pay p but does not win the object.[4]

This is known as the 'independent, private values' model. **Independent** refers to the fact that each buyer's *private information* (in this case, each buyer's value) is independent of every other buyer's private information. **Private value** refers to the fact that once a buyer employs his own private information to assess the value of the object, this assessment would be unaffected were he subsequently to learn any other buyer's private information, i.e., each buyer's *private* information is sufficient for determining his *value*.[5]

Throughout this chapter, we will assume that the setting in which our monopolist finds himself is well-represented by the independent private values model. We can now

---

[2]This amounts to assuming that the object has already been produced and that the seller's use value for it is zero.

[3]Recall that $F_i(v_i)$ denotes the probability that i's value is less than or equal to v_i, and that $f_i(v_i) = F_i'(v_i)$. The latter relation can be equivalently expressed as $F_i(v_i) = \int_0^{v_i} f_i(x)\,dx$. Consequently, we will sometimes refer to f_i and sometimes refer to F_i since each one determines the other.

[4]Although such an outcome is not possible in any one of the four auctions above, there are other auctions (i.e., all-pay auctions) in which payments must be made whether or not one wins the object.

[5]There are more general models in which buyers with private information would potentially obtain yet *additional* information about the value of the object were they to learn *another buyer's* private information, but we shall not consider such models here.

begin to think about how the seller's profits vary with different auction formats. Note that with the production decision behind him and his own value equal to zero, profit-maximisation is equivalent to revenue-maximisation.

Before we can determine the seller's revenues in each of the four standard auctions, we must understand the bidding behaviour of the buyers across the different auction formats. Let us start with the first-price auction.

9.2.1 BIDDING BEHAVIOUR IN A FIRST-PRICE, SEALED-BID AUCTION

To understand bidding behaviour in a first-price auction, we shall, for simplicity, assume that the buyers are ex ante symmetric. That is, we shall suppose that for all buyers $i = 1, \ldots, N$, $f_i(v) = f(v)$ for all $v \in [0, 1]$.

Clearly, the main difficulty in determining the seller's revenue is in determining how the buyers, let us agree to call them *bidders* now, will bid. But note that if you are one of the bidders, then because you would prefer to win the good at a lower price rather than a higher one, you will want to bid low when the others are bidding low and you will want to bid higher when the others bid higher. Of course, you do not know the bids that the others submit because of the sealed-bid rule. Yet, *your optimal bid will depend on how the others bid.* Thus, the bidders are in a strategic setting in which the optimal action (bid) of each bidder depends on the actions of others. Consequently, to determine the behaviour of the bidders, we shall employ the game theoretic tools developed in Chapter 7.

Let us consider the problem of how to bid from the point of view of bidder i. Suppose that bidder i's value is v_i. Given this value, bidder i must submit a sealed bid, b_i. Because b_i will in general depend on i's value, let us write $b_i(v_i)$ to denote bidder i's bid when his value is v_i. Now, because bidder i must be prepared to submit a bid $b_i(v_i)$ for each of his potential values $v_i \in [0, 1]$, we may view bidder i's **strategy** as a *bidding function* $b_i\colon [0, 1] \to \mathbb{R}_+$, mapping each of his values into a (possibly different) non-negative bid.

Before we discuss payoffs, it will be helpful to focus our attention on a natural class of bidding strategies. It seems very natural to expect that bidders with higher values will place higher bids. So, let us restrict attention to *strictly increasing* bidding functions. Next, because the bidders are ex ante symmetric, it is also natural to suppose that bidders with the same value will submit the same bid. With this in mind, we shall focus on finding a strictly increasing bidding function, $\hat{b}\colon [0, 1] \to \mathbb{R}_+$, that is optimal for each bidder to employ, given that all other bidders employ this bidding function as well. That is, we wish to find a symmetric Nash equilibrium in strictly increasing bidding functions.

Now, let us suppose that we find a symmetric Nash equilibrium given by the strictly increasing bidding function $\hat{b}(\cdot)$. By definition it must be payoff-maximising for a bidder, say i, with value v to bid $\hat{b}(v)$ given that the other bidders employ the same bidding function $\hat{b}(\cdot)$. Because of this, we can usefully employ what may at first appear to be a rather mysterious exercise.

The mysterious but useful exercise is this: imagine that bidder i cannot attend the auction and that he sends a friend to bid for him. The friend knows the equilibrium bidding function $\hat{b}(\cdot)$, but he does not know bidder i's value. Now, if bidder i's value is v, bidder i would like his friend to submit the bid $\hat{b}(v)$ on his behalf. His friend can do this for him once bidder i calls him and tells him his value. Clearly, bidder i has no incentive to lie to his friend about his value. That is, among all the values $r \in [0, 1]$ that bidder i with value v can report to his friend, his payoff is maximised by reporting his true value, v, to his friend. This is because reporting the value r results in his friend submitting the bid $\hat{b}(r)$ on his behalf. But if bidder i were there himself he would submit the bid $\hat{b}(v)$.

Let us calculate bidder i's expected payoff from reporting an arbitrary value, r, to his friend when his value is v, given that all other bidders employ the bidding function $\hat{b}(\cdot)$. To calculate this expected payoff, it is necessary to notice just two things. First, bidder i will win only when the bid submitted for him is highest. That is, when $\hat{b}(r) > \hat{b}(v_j)$ for all bidders $j \neq i$. Because $\hat{b}(\cdot)$ is strictly increasing this occurs precisely when r exceeds the values of all $N-1$ other bidders. Letting F denote the distribution function associated with f, the probability that this occurs is $(F(r))^{N-1}$ which we will denote $F^{N-1}(r)$. Second, bidder i pays only when he wins and he then pays his bid, $\hat{b}(r)$. Consequently, bidder i's expected payoff from reporting the value r to his friend when his value is v, given that all other bidders employ the bidding function $\hat{b}(\cdot)$, can be written

$$u(r, v) = F^{N-1}(r)(v - \hat{b}(r)). \tag{9.1}$$

Now, as we have already remarked, because $\hat{b}(\cdot)$ is an equilibrium, bidder i's expected payoff-maximising bid when his value is v must be $\hat{b}(v)$. Consequently, (9.1) must be maximised when $r = v$, i.e., when bidder i reports his true value, v, to his friend. So, if we differentiate the right-hand side with respect to r, the resulting derivative must be zero when $r = v$. Differentiating yields

$$\frac{dF^{N-1}(r)(v - \hat{b}(r))}{dr} = (N-1)F^{N-2}(r)f(r)(v - \hat{b}(r)) - F^{N-1}(r)\hat{b}'(r). \tag{9.2}$$

Evaluating the right-hand side at $r = v$, where it is equal to zero, and rearranging yields,

$$(N-1)F^{N-2}(v)f(v)\hat{b}(v) + F^{N-1}(v)\hat{b}'(v) = (N-1)vf(v)F^{N-2}(v). \tag{9.3}$$

Looking closely at the left-hand side of (9.3), we see that it is just the derivative of the product $F^{N-1}(v)\hat{b}(v)$ with respect to v. With this observation, we can rewrite (9.3) as

$$\frac{dF^{N-1}(v)\hat{b}(v)}{dv} = (N-1)vf(v)F^{N-2}(v). \tag{9.4}$$

Now, because (9.4) must hold for every v, it must be the case that

$$F^{N-1}(v)\hat{b}(v) = (N-1)\int_0^v xf(x)F^{N-2}(x)\,dx + \text{constant}.$$

Noting that a bidder with value zero must bid zero, we conclude that the constant above must be zero. Hence, it must be the case that

$$\hat{b}(v) = \frac{N-1}{F^{N-1}(v)} \int_0^v xf(x)F^{N-2}(x)\,dx,$$

which can be written more succinctly as

$$\hat{b}(v) = \frac{1}{F^{N-1}(v)} \int_0^v x dF^{N-1}(x). \tag{9.5}$$

There are two things to notice about the bidding function in (9.5). First, as we had assumed, it is strictly increasing in v (see Exercise 9.1). Second, it has been uniquely determined. Hence, in conclusion, we have proven the following.

THEOREM 9.1 ***First-Price Auction Symmetric Equilibrium***

If N bidders have independent private values drawn from the common distribution, F, then bidding

$$\hat{b}(v) = \frac{1}{F^{N-1}(v)} \int_0^v x dF^{N-1}(x)$$

whenever one's value is v constitutes a symmetric Nash equilibrium of a first-price, sealed-bid auction. Moreover, this is the only symmetric Nash equilibrium.[6]

EXAMPLE 9.1 Suppose that each bidder's value is uniformly distributed on $[0, 1]$. Then $F(v) = v$ and $f(v) = 1$. Consequently, if there are N bidders, then each employs the bidding function

$$\begin{aligned}
\hat{b}(v) &= \frac{1}{v^{N-1}} \int_0^v x dx^{N-1} \\
&= \frac{1}{v^{N-1}} \int_0^v x(N-1)x^{N-2}\,dx \\
&= \frac{N-1}{v^{N-1}} \int_0^v x^{N-1}\,dx \\
&= \frac{N-1}{v^{N-1}} \frac{1}{N} v^N \\
&= v - \frac{v}{N}.
\end{aligned}$$

[6]Strictly speaking, we have not shown that this is an equilibrium. We have shown that *if* a symmetric equilibrium exists, then this must be it. You are asked to show that this is indeed an equilibrium in an exercise. You might also wonder about the existence of asymmetric equilibria. It can be shown that there are none, although we shall not do so here.

So, each bidder *shades* his bid, by bidding less than his value. Note that as the number of bidders increases, the bidders bid more aggressively. □

Because $F^{N-1}(\cdot)$ is the distribution function of the highest value among a bidder's $N-1$ competitors, the bidding strategy displayed in Theorem 9.1 says that each bidder bids the expectation of the second highest bidder's value conditional on his own value being highest. But, because the bidders use the *same strictly increasing* bidding function, having the highest value is equivalent to having the highest bid and so equivalent to winning the auction. So, we may say:

In the unique symmetric equilibrium of a first-price, sealed-bid auction, each bidder bids the expectation of the second-highest bidder's value conditional on winning the auction.

The idea that one ought to bid *conditional on winning* is very intuitive in a first-price auction because of the feature that one's bid matters only when one wins the auction. Because this feature is present in other auctions as well, this idea should be considered one of the basic insights of our strategic analysis.

Having analysed the first-price auction, it is an easy matter to describe behaviour in a Dutch auction.

9.2.2 BIDDING BEHAVIOUR IN A DUTCH AUCTION

In a Dutch auction, each bidder has a single decision to make, namely, 'At what price should I raise my hand to signal that I am willing to buy the good at that price?' Moreover, the bidder who chooses the highest price wins the auction and pays this price. Consequently, by replacing the word 'price' by 'bid' in the previous sentence we see that this auction is equivalent to a first-price auction! So, we can immediately conclude the following.

THEOREM 9.2 ***Dutch Auction Symmetric Equilibrium***

If N bidders have independent private values drawn from the common distribution, F, then raising one's hand when the price reaches

$$\frac{1}{F^{N-1}(v)}\int_0^v x dF^{N-1}(x)$$

whenever one's value is v constitutes a symmetric Nash equilibrium of a Dutch auction. Moreover, this is the only symmetric Nash equilibrium.

Clearly then, the first-price and Dutch auctions raise exactly the same revenue for the seller, ex post (i.e., for every realisation of bidder values $v_1, \ldots, v_N$).

We now turn to the second-price, sealed-bid auction.

9.2.3 BIDDING BEHAVIOUR IN A SECOND-PRICE, SEALED-BID AUCTION

One might wonder why we would bother considering a second-price auction at all. Is it not obvious that a first-price auction must yield higher revenue for the seller? After all, in a first-price auction the seller receives the *highest* bid, whereas in a second-price auction he receives only the *second-highest* bid.

While this might sound convincing, it neglects a crucial point: *The bidders will bid differently in the two auctions.* In a first-price auction, a bidder has an incentive to raise his bid to increase his chances of winning the auction, yet he has an incentive to reduce his bid to lower the price he pays when he does win. In a second-price auction, the second effect is absent because when a bidder wins, the amount he pays is independent of his bid. So, we should expect bidders to bid *more aggressively* in a second-price auction than they would in a first-price auction. Therefore, there is a chance that a second-price auction will generate higher expected revenues for the seller than will a first-price auction. When we recognise that bidding behaviour changes with the change in the auction format, the question of which auction raises more revenue is not quite so obvious, is it?

Happily, analysing bidding behaviour in a second-price, sealed-bid auction is remarkably straightforward. Unlike our analysis of the first-price auction, we need not restrict attention to the case involving symmetric bidders. That is, we shall allow the density functions $f_1, \ldots, f_N$, from which the bidders' values are independently drawn, to differ.[7]

Consider bidder i with value v_i, and let B denote the highest bid submitted by the other bidders. Of course, B is unknown to bidder i because the bids are sealed. Now, if bidder i were to win the auction, his bid would be highest and B would then be the second-highest bid. Consequently, bidder i would have to pay B for the object. In effect, then, the price that bidder i must pay for the object is the highest bid, B, submitted by the other bidders.

Now, because bidder i's value is v_i, he would strictly want to win the auction when his value exceeds the price he would have to pay, i.e., when $v_i > B$; and he would strictly want to lose when $v_i < B$. When $v_i = B$ he is indifferent between winning and losing. Can bidder i bid in a manner that guarantees that he will win when $v_i > B$ and that he will lose when $v_i < B$, even though he does not know B? The answer is yes. He can guarantee precisely this simply by bidding his value, v_i!

By bidding v_i, bidder i is the high bidder, and so wins, when $v_i > B$, and he is not the high bidder, and so loses, when $v_i < B$. Consequently, bidding his value is a payoff-maximising bid for bidder i *regardless of the bids submitted by the other bidders* (recall that B was the highest bid among any arbitrary bids submitted by the others). Moreover, because bidding below one's value runs the risk of losing the auction when one would have strictly preferred winning it, and bidding above one's value runs the risk of winning the auction for a price above one's value, bidding one's value is a weakly dominant bidding strategy. So, we can state the following.

[7] In fact, even the independence assumption can be dropped. (See Exercise 9.5.)

THEOREM 9.3 ***Second-Price Auction Equilibrium***

If N bidders have independent private values, then bidding one's value is the unique weakly dominant bidding strategy for each bidder in a second-price, sealed-bid auction.

This brings us to the English auction.

9.2.4 BIDDING BEHAVIOUR IN AN ENGLISH AUCTION

In contrast to the auctions we have considered so far, in an English auction there are potentially many decisions a bidder has to make. For example, when the price is very low, he must decide at which price he would drop out when no one has yet dropped out. But, if some other bidder drops out first, he must then decide at which price to drop out *given* the remaining active bidders, and so on. Despite this, there is a close connection between the English and second-price auctions.

In an English auction, as in a second-price auction, it turns out to be a dominant strategy for a bidder to drop out when the price reaches his value, regardless of which bidders remain active. The reason is rather straightforward. A bidder i with value v_i who, given the history of play and the current price $p < v_i$, considers dropping out can do no worse by planning to remain active a little longer and until the price reaches his value, v_i. By doing so, the worst that can happen is that he ends up dropping out when the price does indeed reach his value. His payoff would then be zero, just as it would be if he were to drop out now at price p. However, it might happen, were he to remain active, that all other bidders would drop out before the price reaches v_i. In this case, bidder i would be strictly better off by having remained active since he then wins the object at a price strictly less than his value v_i, obtaining a positive payoff. So, we have the following.

THEOREM 9.4 ***English Auction Equilibrium***

If N bidders have independent private values, then dropping out when the price reaches one's value is the unique weakly dominant bidding strategy for each bidder in an English auction.[8]

Given this result, it is easy to see that the bidder with the highest value will win in an English auction. But what price will he pay for the object? That, of course, depends on the price at which his *last remaining competitor* drops out of the auction. But his last remaining competitor will be the bidder with the *second-highest value*, and he will, like all bidders, drop out when the price reaches his value. Consequently, the bidder with highest value wins and pays a price equal to the second-highest value. Hence, we see that the outcome of the English auction is identical to that of the second-price auction. In particular, the English and second-price auctions earn exactly the same revenue for the seller, ex post.

[8]As in the second-price auction case, this weak dominance result does not rely on the independence of the bidder's values. It holds even if the values are correlated. However, it is important that the values are *private*.

9.2.5 REVENUE COMPARISONS

Because the first-price and Dutch auctions raise the same ex post revenue and the second-price and English auctions raise the same ex post revenue, it remains only to compare the revenues generated by the first- and second-price auctions. Clearly, these auctions need not raise the same revenue ex post. For example, when the highest value is quite high and the second-highest is quite low, running a first-price auction will yield more revenue than a second-price auction. On the other hand, when the first- and second-highest values are close together, a second-price auction will yield higher revenues than will a first-price auction.

Of course, when the seller must decide which of the two auction forms to employ, he does not know the bidders' values. However, knowing how the bidders bid as functions of their values, and knowing the distribution of bidder values, the seller can calculate the *expected revenue* associated with each auction. Thus, the question is, which auction yields the highest expected revenue, a first- or a second-price auction? Because our analysis of the first-price auction involved symmetric bidders, we must assume symmetry here to compare the expected revenue generated by a first-price versus a second-price auction. So, in what follows, $f(\cdot)$ will denote the common density of each bidder's value and $F(\cdot)$ will denote the associated distribution function.

Let us begin by considering the expected revenue, R_{FPA}, generated by a first-price auction (FPA). Because the highest bid wins a first-price auction and because the bidder with the highest value submits the highest bid, if v is the highest value among the N bidder values, then the seller's revenue is $\hat{b}(v)$. So, if the highest value is distributed according to the density $g(v)$, the seller's expected revenue can be written

$$R_{FPA} = \int_0^1 \hat{b}(v)g(v)\,dv.$$

Because the density, g, of the maximum of N independent random variables with common density f and distribution F is NfF^{N-1},[9] we have

$$R_{FPA} = N\int_0^1 \hat{b}(v)f(v)F^{N-1}(v)\,dv. \tag{9.6}$$

We have seen that in a second-price auction, because each bidder bids his value, the seller receives as price the second-highest value among the N bidder values. So, if $h(v)$ is the density of the second-highest value, the seller's expected revenue, R_{SPA}, in a second-price auction can be written

$$R_{SPA} = \int_0^1 vh(v)\,dv.$$

[9]To see this, note that the highest value is less than or equal to v if and only if all N values are, and that this occurs with probability $F^N(v)$. Hence, the distribution function of the highest value is F^N. Because the density function is the derivative of the distribution function the result follows.

Because the density, h, of the second-highest of N independent random variables with common density f and distribution function F is $N(N-1)F^{N-2}f(1-F)$,[10] we have

$$R_{SPA} = N(N-1)\int_0^1 vF^{N-2}(v)f(v)(1-F(v))\,dv. \tag{9.7}$$

We shall now compare the two. From (9.6) and (9.5) we have

$$\begin{aligned}
R_{FPA} &= N\int_0^1 \left[\frac{1}{F^{N-1}(v)}\int_0^v x\,dF^{N-1}(x)\right] f(v)F^{N-1}(v)\,dv \\
&= N(N-1)\int_0^1 \left[\int_0^v xF^{N-2}(x)f(x)\,dx\right] f(v)\,dv \\
&= N(N-1)\int_0^1 \int_0^v [xF^{N-2}(x)f(x)f(v)]\,dx\,dv \\
&= N(N-1)\int_0^1 \int_x^1 [xF^{N-2}(x)f(x)f(v)]\,dv\,dx \\
&= N(N-1)\int_0^1 xF^{N-2}(x)f(x)(1-F(x))\,dx \\
&= R_{SPA},
\end{aligned}$$

where the fourth equality follows from interchanging the order of integration (i.e., from $dxdv$ to $dvdx$), and the final equality follows from (9.7).

EXAMPLE 9.2 Consider the case in which each bidder's value is uniform on $[0, 1]$ so that $F(v) = v$ and $f(v) = 1$. The expected revenue generated in a first-price auction is

$$\begin{aligned}
R_{FPA} &= N\int_0^1 \hat{b}(v)f(v)F^{N-1}(v)\,dv \\
&= N\int_0^1 \left[v - \frac{v}{N}\right] v^{N-1}\,dv \\
&= (N-1)\int_0^1 v^N\,dv \\
&= \frac{N-1}{N+1}.
\end{aligned}$$

[10] One way to see this is to treat probability density like probability. Then the probability (density) that some particular bidder's value is v is $f(v)$ and the probability that exactly one of the remaining $N-1$ other bidders' values is above this is $(N-1)F^{N-2}(v)(1-F(v))$. Consequently, the probability that this particular bidder's value is v and it is second-highest is $(N-1)f(v)F^{N-2}(v)(1-F(v))$. Because there are N bidders, the probability (i.e., density) that the second-highest value is v is then $N(N-1)f(v)F^{N-2}(v)(1-F(v))$.

On the other hand, the expected revenue generated in a second-price auction is

$$\begin{aligned}
R_{SPA} &= N(N-1)\int_0^1 vF^{N-2}(v)f(v)(1-F(v))\,dv \\
&= N(N-1)\int_0^1 v^{N-1}(1-v)\,dv \\
&= N(N-1)\left[\frac{1}{N}-\frac{1}{N+1}\right] \\
&= \frac{N-1}{N+1}.
\end{aligned}$$

□

Remarkably, the first- and second-price auctions raise the *same* expected revenue, regardless of the common distribution of bidder values! So, we may state the following:

If N bidders have independent private values drawn from the common distribution, F, then all four standard auction forms (first-price, second-price, Dutch, and English) raise the same expected revenue for the seller.

This *revenue equivalence* result may go some way towards explaining why we see all four auction forms in practice. Were it the case that one of them raised more revenue than the others on average, then we would expect that one to be used rather than any of the others. But what is it that accounts for the coincidence of expected revenue in these auctions? Our next objective is to gain some insight into why this is so.

9.3 The Revenue Equivalence Theorem

To explain the equivalence of revenue in the four standard auction forms, we must first find a way to fit all of these auctions into a single framework. With this in mind, we now define the notion of a *direct selling mechanism.*[11]

DEFINITION 9.1 ***Direct Selling Mechanism***

A direct selling mechanism is a collection of N probability assignment functions,

$$p_1(v_1,\ldots,v_N),\ldots,p_N(v_1,\ldots,v_N),$$

and N cost functions

$$c_1(v_1,\ldots,v_N),\ldots,c_N(v_1,\ldots,v_N).$$

For every vector of values $(v_1,\ldots,v_N)$ *reported by the N bidders,* $p_i(v_1,\ldots,v_N)\in[0,1]$ *denotes the probability that bidder i receives the object and* $c_i(v_1,\ldots,v_N)\in\mathbb{R}$

[11] Our presentation is based upon Myerson (1981).

denotes the payment that bidder i must make to the seller. The sum of the probabilities, $p_1(v_1, \ldots, v_N) + \cdots + p_N(v_1, \ldots, v_N)$ *is always no greater than unity.*

A direct selling mechanism works as follows. Because the seller does not know the bidders' values, he asks them to report them to him simultaneously. He then takes those reports, $v_1, \ldots, v_N$, which need not be truthful, and assigns the object to one of the bidders according to the probabilities $p_i(v_1, \ldots, v_N)$, $i = 1, \ldots, N$, keeping the object with the residual probability, and secures the payment $c_i(v_1, \ldots, v_N)$ from each bidder $i = 1, \ldots, N$. It is assumed that the entire direct selling mechanism – the probability assignment functions and the cost functions – are public information, and that the seller must carry out the terms of the mechanism given the vector of reported values.

Several points are worthy of note. First, although the sum of probabilities $p_1 + \cdots + p_N$ can never exceed unity, we allow this sum to fall short of unity because we want to allow the seller to keep the object.[12] Second, a bidder's cost may be negative. Third, a bidder's cost may be positive even when that bidder does not receive the object (i.e., when that bidder's probability of receiving the object is zero).

Clearly, the seller's revenue will depend on the reports submitted by the bidders. Will they be induced to report truthfully? If not, how will they behave? These are very good questions, but let us put them aside for the time being. Instead, we introduce what will turn out to be an extremely important special kind of direct selling mechanism, namely, those in which the bidders find it in their interest to report truthfully. These mechanisms are called *incentive-compatible*. Before introducing the formal definition, we introduce a little notation.

Consider a direct selling mechanism $(p_i(\cdot), c_i(\cdot))_{i=1}^N$. Suppose that bidder i's value is v_i and he considers reporting that his value is r_i. If all other bidders always report their values truthfully, then bidder i's expected payoff is

$$u_i(r_i, v_i) = \int_0^1 \cdots \int_0^1 (p_i(r_i, v_{-i})v_i - c_i(r_i, v_{-i}))f_{-i}(v_{-i})dv_{-i},$$

where $f_{-i}(v_{-i}) = f(v_1) \cdots f(v_{i-1})f(v_{i+1}) \cdots f(v_N)$ and $dv_{-i} = dv_1 \cdots dv_{i-1}dv_{i+1} \cdots dv_N$.

For every $r_i \in [0, 1]$, let

$$\bar{p}_i(r_i) = \int_0^1 \cdots \int_0^1 p_i(r_i, v_{-i})f_{-i}(v_{-i})dv_{-i}$$

and

$$\bar{c}_i(r_i) = \int_0^1 \cdots \int_0^1 c_i(r_i, v_{-i})f_{-i}(v_{-i})dv_{-i}.$$

[12]This is more generality than we need at the moment because the seller never keeps the object in any of the four standard auctions. However, this will be helpful a little later.

Therefore, $\bar{p}_i(r_i)$ is the probability that i receives the object when he reports r_i and $\bar{c}_i(r_i)$ is i's expected payment when he reports r_i, with both of these being conditional on all others always reporting truthfully. Consequently, bidder i's expected payoff when his value is v_i and he reports r_i can be written as

$$u_i(r_i, v_i) = \bar{p}_i(r_i)v_i - \bar{c}_i(r_i), \tag{9.8}$$

when all other bidders report their values truthfully.

We can now state the definition of an **incentive-compatible direct selling mechanism**.

DEFINITION 9.2 ***Incentive-Compatible Direct Selling Mechanisms***

A direct selling mechanism is incentive-compatible if when the other bidders always report their values truthfully, each bidder i's expected payoff is maximised by always reporting his value truthfully – i.e., the mechanism is incentive-compatible if for each bidder i, and for each of his values $v_i \in [0, 1]$, $u_i(r_i, v_i)$ as defined in (9.8) is maximised in $r_i \in [0, 1]$ when $r_i = v_i$. We shall then say that it is a Bayesian-Nash equilibrium for each bidder to always report his value truthfully.[13]

Note very carefully what the definition does *not* say. It does not say that reporting truthfully is best for a bidder regardless of the others' reports. It *only* says that a bidder can do no better than to report truthfully so long as all other bidders report truthfully. Thus, although truthful reporting is a Bayesian-Nash equilibrium in an incentive-compatible mechanism, it need not be a dominant strategy for any player.

You might wonder how all of this is related to the four standard auctions. We will now argue that each of the four standard auctions can be equivalently viewed as an incentive-compatible direct selling mechanism. In fact, understanding incentive-compatible direct selling mechanisms will not only be the key to understanding the connection between the four standard auctions, but it will be central to our understanding revenue-maximising auctions as well.

Consider a first-price auction with symmetric bidders. We would like to construct an 'equivalent' direct selling mechanism in which truth-telling is an equilibrium. To do this, we shall employ the first-price auction equilibrium bidding function $\hat{b}(\cdot)$. The idea behind our construction is simple. Instead of the bidders submitting bids computed by plugging their values into the equilibrium bidding function, the bidders will be asked to submit their values and the seller will then compute their equilibrium bids for them. Recall that because $\hat{b}(\cdot)$ is strictly increasing, a bidder wins the object in a first-price auction if and only if he has the highest value.

[13] This would in fact be a consequence of our Chapter 7 definition of Bayesian-Nash equilibrium but for the fact that we restricted attention to finite type spaces there.

Consider, then, the following direct selling mechanism, where $\hat{b}(\cdot)$ is the equilibrium bidding function for the first-price auction given in (9.5):

$$p_i(v_1, \ldots, v_N) = \begin{cases} 1, & \text{if } v_i > v_j \text{ for all } j \neq i \\ 0, & \text{otherwise,} \end{cases}$$

$$\text{and} \tag{9.9}$$

$$c_i(v_1, \ldots, v_N) = \begin{cases} \hat{b}(v_i), & \text{if } v_i > v_j \text{ for all } j \neq i \\ 0, & \text{otherwise.} \end{cases}$$

Look closely at this mechanism. Note that the bidder with the highest reported value, v, receives the object and he pays $\hat{b}(v)$ for it, just as he would have in a first-price auction equilibrium. So, if the bidders report their values truthfully, then the bidder with the highest value, v, wins the object and makes the payment $\hat{b}(v)$ to the seller. Consequently, if this mechanism is incentive-compatible, the seller will earn exactly the same ex post revenue as he would with a first-price auction.

To demonstrate that this mechanism is incentive-compatible we need to show that truth-telling is a Nash equilibrium. So, let us suppose that all other bidders report their values truthfully and that the remaining bidder has value v. We must show that this bidder can do no better than to report his value truthfully to the seller. So, suppose that this bidder considers reporting value r. He then wins the object and makes a payment of $\hat{b}(r)$ if and only if $r > v_j$ for all other bidders j. Because the other $N-1$ bidders' values are independently distributed according to F, this event occurs with probability $F^{N-1}(r)$. Consequently, this bidder's expected payoff from reporting value r when his true value is v is

$$F^{N-1}(r)(v - \hat{b}(r)).$$

But this is exactly the payoff in (9.1), which we already know is maximised when $r = v$. Hence, the direct selling mechanism (9.9) is indeed incentive-compatible.

Let us reconsider what we have accomplished here. Beginning with the equilibrium of a first-price auction, we have constructed an incentive-compatible direct selling mechanism whose truth-telling equilibrium results in the same ex post assignment of the object to bidders and the same ex post payments by them. In particular, it results in the same ex post revenue for the seller. Moreover, this method of constructing a direct mechanism is quite general. Indeed, beginning with the equilibrium of any of the four standard auctions, we can similarly construct an incentive-compatible direct selling mechanism that yields the same ex post assignment of the object to bidders and the same ex post payments by them. (You are asked to do this in an exercise.)

In effect, we have shown that each of the four standard auctions is equivalent to some incentive-compatible direct selling mechanism. Because of this, we can now gain insight into the former by studying the latter.

9.3.1 INCENTIVE-COMPATIBLE DIRECT SELLING MECHANISMS: A CHARACTERISATION

Because incentive-compatible mechanisms are so important, it is very helpful to know how to identify them. The following result provides a complete characterisation. It states that a direct mechanism is incentive-compatible if two conditions are met. First, it must be the case that reporting a higher value leads a bidder to expect that he will receive the object with higher probability. Second, the cost a bidder expects to pay must be related in a very particular way to the probability with which he expects to receive the object.

THEOREM 9.5 ***Incentive-Compatible Direct Selling Mechanisms***

A direct selling mechanism $(p_i(\cdot), c_i(\cdot))_{i=1}^N$ is incentive-compatible if and only if for every bidder i

(i) $\bar{p}_i(v_i)$ *is non-decreasing in* v_i *and,*

(ii) $\bar{c}_i(v_i) = \bar{c}_i(0) + \bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)dx$, *for every* $v_i \in [0, 1]$.

Proof: Suppose the mechanism is incentive-compatible. We must show that (i) and (ii) hold.

To see that (i) holds, note that by incentive compatibility, for all $r_i, v_i \in [0, 1]$,

$$\bar{p}_i(r_i)v_i - \bar{c}_i(r_i) = u_i(r_i, v_i) \le u_i(v_i, v_i) = \bar{p}_i(v_i)v_i - \bar{c}_i(v_i).$$

Adding and subtracting $\bar{p}_i(v_i)r_i$ to the right-hand side, this implies

$$\bar{p}_i(r_i)v_i - \bar{c}_i(r_i) \le [\bar{p}_i(v_i)r_i - \bar{c}_i(v_i)] + \bar{p}_i(v_i)(v_i - r_i).$$

But a careful look at the term in square brackets reveals that it is $u_i(v_i, r_i)$, bidder i's expected payoff from reporting v_i when his true value is r_i. By incentive compatibility, this must be no greater than $u_i(r_i, r_i)$, his payoff when he reports his true value, r_i. Consequently,

$$\begin{aligned}
\bar{p}_i(r_i)v_i - \bar{c}_i(r_i) &\le [\bar{p}_i(v_i)r_i - \bar{c}_i(v_i)] + \bar{p}_i(v_i)(v_i - r_i) \\
&\le u_i(r_i, r_i) + \bar{p}_i(v_i)(v_i - r_i) \\
&= [\bar{p}_i(r_i)r_i - \bar{c}_i(r_i)] + \bar{p}_i(v_i)(v_i - r_i).
\end{aligned}$$

That is,

$$\bar{p}_i(r_i)v_i - \bar{c}_i(r_i) \le [\bar{p}_i(r_i)r_i - \bar{c}_i(r_i)] + \bar{p}_i(v_i)(v_i - r_i),$$

which, when rewritten, becomes

$$(\bar{p}_i(v_i) - \bar{p}_i(r_i))(v_i - r_i) \ge 0.$$

So, when $v_i > r_i$, it must be the case that $\bar{p}_i(v_i) \ge \bar{p}_i(r_i)$. We conclude that $\bar{p}_i(\cdot)$ is non-decreasing. Hence, (i) holds. (See also Exercise 9.7.)

To see that (ii) holds, note that because bidder i's expected payoff must be maximised when he reports truthfully, the derivative of $u_i(r_i, v_i)$ with respect to r_i must be zero when

$r_i = v_i$.[14] Computing this derivative yields

$$\frac{\partial u_i(r_i, v_i)}{\partial r_i} = \bar{p}_i'(r_i)v_i - \bar{c}_i'(r_i),$$

and setting this to zero when $r_i = v_i$ yields

$$\bar{c}_i'(v_i) = \bar{p}_i'(v_i)v_i. \tag{P.1}$$

Because v_i was arbitrary, (P.1) must hold for every $v_i \in [0, 1]$. Consequently,

$$\begin{aligned}\bar{c}_i(v_i) - \bar{c}_i(0) &= \int_0^{v_i} \bar{c}_i'(x)\,dx \\ &= \int_0^{v_i} \bar{p}_i'(x)x\,dx \\ &= \bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)\,dx,\end{aligned}$$

where the first equality follows from the fundamental theorem of calculus, the second from (P.1), and the third from integration by parts. Consequently, for every bidder i and every $v_i \in [0, 1]$,

$$\bar{c}_i(v_i) = \bar{c}_i(0) + \bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)\,dx, \tag{P.2}$$

proving (ii).

We must now show the converse. So, suppose that (i) and (ii) hold. We must show that $u_i(r_i, v_i)$ is maximised in r_i when $r_i = v_i$. To see this, note that substituting (ii) into (9.8) yields

$$u_i(r_i, v_i) = \bar{p}_i(r_i)v_i - \left[\bar{c}_i(0) + \bar{p}_i(r_i)r_i - \int_0^{r_i} \bar{p}_i(x)\,dx\right]. \tag{P.3}$$

This can be rewritten as

$$u_i(r_i, v_i) = -\bar{c}_i(0) + \int_0^{v_i} \bar{p}_i(x)\,dx - \left\{\int_{r_i}^{v_i} (\bar{p}_i(x) - \bar{p}_i(r_i))\,dx\right\},$$

where this expression is valid whether $r_i \le v_i$ or $r_i \ge v_i$.[15] Because by (i) $\bar{p}_i(\cdot)$ is non-decreasing, the integral in curly brackets is non-negative for all r_i and v_i. Consequently,

$$u_i(r_i, v_i) \le -\bar{c}_i(0) + \int_0^{v_i} \bar{p}_i(x)\,dx. \tag{P.4}$$

---

[14] We are ignoring two points here. The first is whether $u_i(r_i, v_i)$ is in fact differentiable in r_i. Although it need not be everywhere differentiable, incentive compatibility implies that it must be differentiable almost everywhere and that the analysis we shall conduct can be made perfectly rigorous. We will not pursue these details here. The second point we ignore is the first-order condition at the two non-interior values $v_i = 0$ or 1. Strictly speaking, the derivatives at these boundary points need not be zero. But there is no harm in this because these two values each occur with probability zero.

[15] Recall the convention in mathematics that when $a < b$, $\int_b^a f(x)\,dx = -\int_a^b f(x)\,dx$.

But, by (P.3), the right-hand side of (P.4) is equal to $u_i(v_i, v_i)$. Consequently,

$$u_i(r_i, v_i) \leq u_i(v_i, v_i),$$

so that $u_i(r_i, v_i)$ is indeed maximised in r_i when $r_i = v_i$. ∎

Part (ii) of Theorem 9.5 says that if a direct mechanism is incentive-compatible there must be a connection between the probability assignment functions and the cost functions. In particular, it says that once the probability assignment function has been chosen and once a bidder's expected cost conditional on having value zero is chosen, the remainder of the expected cost function is chosen as well. To put it differently, under incentive compatibility a bidder's expected payment conditional on his value is completely determined by his expected payment when his value is zero and his probability assignment function. This observation is essential for understanding the following result.

THEOREM 9.6 ***Revenue Equivalence***

If two incentive-compatible direct selling mechanisms have the same probability assignment functions and every bidder with value zero is indifferent between the two mechanisms, then the two mechanisms generate the same expected revenue for the seller.

Proof: The seller's expected revenue is

$$\begin{aligned}
R &= \int_0^1 \cdots \int_0^1 \sum_{i=1}^N c_i(v_1, \ldots, v_N) f(v_1) \ldots f(v_N)\, dv_1 \ldots dv_N \\
&= \sum_{i=1}^N \int_0^1 \cdots \int_0^1 c_i(v_1, \ldots, v_N) f(v_1) \ldots f(v_N)\, dv_1 \ldots dv_N \\
&= \sum_{i=1}^N \int_0^1 \left[\int_0^1 \cdots \int_0^1 c_i(v_i, v_{-i}) f_{-i}(v_{-i})\, dv_{-i} \right] f_i(v_i)\, dv_i \\
&= \sum_{i=1}^N \int_0^1 \bar{c}_i(v_i) f_i(v_i)\, dv_i \\
&= \sum_{i=1}^N \int_0^1 \left[\bar{c}_i(0) + \bar{p}_i(v_i) v_i - \int_0^{v_i} \bar{p}_i(x)\, dx \right] f_i(v_i)\, dv_i \\
&= \sum_{i=1}^N \int_0^1 \left[\bar{p}_i(v_i) v_i - \int_0^{v_i} \bar{p}_i(x)\, dx \right] f_i(v_i)\, dv_i + \sum_{i=1}^N \bar{c}_i(0),
\end{aligned}$$

where the fourth equality follows from the definition of $\bar{c}_i(v_i)$ and the fifth equality follows from (ii) of Theorem 9.5.

Consequently, the seller's expected revenue depends only on the probability assignment functions and the amount bidders expect to pay when their values are zero. Because a bidder's expected payoff when his value is zero is completely determined by his expected payment when his value is zero, the desired result follows. ∎

The revenue equivalence theorem provides an explanation for the apparently coincidental equality of expected revenue among the four standard auctions. We now see that this follows because, with symmetric bidders, each of the four standard auctions has the same probability assignment function (i.e., the object is assigned to the bidder with the highest value), and in each of the four standard auctions a bidder with value zero receives expected utility equal to zero.

The revenue equivalence theorem is very general and allows us to add additional auctions to the list of those yielding the same expected revenue as the four standard ones. For example, a first-price, all-pay auction, in which the highest among all sealed bids wins but *every* bidder pays an amount equal to his bid, also yields the same expected revenue under bidder symmetry as the four standard auctions. You are asked to explore this and other auctions in the exercises.

9.3.2 EFFICIENCY

Before closing this section, we briefly turn our attention to the allocative properties of the four standard auctions. As we have already noted several times, each of these auctions allocates the object to the bidder who values it most. That is, each of these auctions is efficient. In the case of the Dutch and the first-price auctions, this result relies on bidder symmetry. Without symmetry, different bidders in a first-price auction, say, will employ different strictly increasing bidding functions. Consequently, if one bidder employs a lower bidding function than another, then the one may have a higher value yet be outbid by the other.

9.4 DESIGNING A REVENUE MAXIMISING MECHANISM

By now we understand very well the four standard auctions, their equilibria, their expected revenue, and the relation among them. But do these auctions, each generating the same expected revenue (under bidder symmetry), maximise the seller's expected revenue? Or is there a better selling mechanism for the seller? If there is a better selling mechanism what form does it take? Do the bidders submit sealed bids? Do they bid sequentially? What about a combination of the two? Is an auction the best selling mechanism?

9.4.1 THE REVELATION PRINCIPLE

Apparently, finding a revenue-maximising selling mechanism is likely to be a difficult task. Given the freedom to choose any selling procedure, where do we start? The key observation is to recall how we were able to construct an incentive-compatible direct selling mechanism from the equilibrium of a first-price auction, and how the outcome of the first-price auction was exactly replicated in the direct mechanism's truth-telling equilibrium. As it turns out, the same type of construction can be applied to any selling procedure. That is, given an arbitrary selling procedure and a Nash equilibrium in which each bidder employs a strategy mapping his value into payoff-maximising behaviour under that selling procedure, we can construct an equivalent incentive-compatible direct selling mechanism. The requisite probability assignment and cost functions map each vector of values to the probabilities and costs that each bidder would experience according to the equilibrium strategies in the original selling procedure. So constructed, this direct selling mechanism is incentive-compatible and yields the same (probabilistic) assignment of the object and the same expected costs to each bidder as well as the same expected revenue to the seller.

Consequently, if some selling procedure yields the seller expected revenue equal to R, then so too does some incentive-compatible direct selling mechanism. But this means that *no selling mechanism among all conceivable selling mechanisms yields more revenue for the seller than the revenue-maximising, incentive-compatible direct selling mechanism.* We can, therefore, restrict our search for a revenue-maximising selling procedure to the (manageable) set of incentive-compatible direct selling mechanisms. In this way, we have simplified our problem considerably while losing nothing.

This simple but extremely important technique for reducing the set of mechanisms to the set of incentive-compatible direct mechanisms is an instance of what is called the **revelation principle**. This principle is used again and again in the theory of mechanism design and we will see it in action again in Section 9.5 when we consider the problem of achieving efficient outcomes in a private information setting.

9.4.2 INDIVIDUAL RATIONALITY

There is one additional restriction we must now consider. Because participation by the bidders is entirely voluntary, no bidder's expected payoff can be negative given his value. Otherwise, whenever he has that value, he will simply not participate in the selling mechanism. Thus, we must restrict attention to incentive-compatible direct selling mechanisms that are **individually rational**, i.e., that yield each bidder, regardless of his value, a non-negative expected payoff in the truth-telling equilibrium.

Now, in an incentive-compatible mechanism bidder i with value v_i will receive expected payoff $u_i(v_i, v_i)$ in the truth-telling equilibrium. So, an incentive-compatible direct selling mechanism is individually rational if this payoff is always non-negative, i.e., if

$$u_i(v_i, v_i) = \bar{p}_i(v_i)v_i - \bar{c}_i(v_i) \geq 0 \text{ for all } v_i \in [0, 1].$$

However, by incentive compatibility, (ii) of Theorem 9.5 tells us that

$$\bar{c}_i(v_i) = \bar{c}_i(0) + \bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)dx, \text{ for every } v_i \in [0, 1].$$

Consequently, an incentive-compatible direct selling mechanism is individually rational if and only if

$$u_i(v_i, v_i) = \bar{p}_i(v_i)v_i - \bar{c}_i(v_i) = -\bar{c}_i(0) + \int_0^{v_i} \bar{p}_i(x)\,dx \geq 0 \text{ for every } v_i \in [0, 1],$$

which clearly holds if and only if

$$\bar{c}_i(0) \leq 0. \tag{9.10}$$

Consequently, an incentive-compatible direct selling mechanism is individually rational if and only if each bidder's expected cost when his value is zero is non-positive.

9.4.3 AN OPTIMAL SELLING MECHANISM

We have now reduced the task of finding the optimal selling mechanism to maximising the seller's expected revenue among all individually rational, incentive-compatible direct selling mechanisms, $p_i(\cdot)$ and $c_i(\cdot)$, $i = 1, \ldots, N$. Because Theorem 9.5 characterises all incentive-compatible selling mechanisms, and because an incentive-compatible direct selling mechanism is individually rational if and only if $\bar{c}_i(0) \leq 0$, our task has been reduced to solving the following problem: choose a direct selling mechanism $p_i(\cdot)$, $c_i(\cdot)$, $i = 1, \ldots, N$, to maximise

$$R = \sum_{i=1}^{N} \int_0^1 \left[\bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)\,dx \right] f_i(v_i)\,dv_i + \sum_{i=1}^{N} \bar{c}_i(0)$$

subject to

(i) $\bar{p}_i(v_i)$ is non-decreasing in v_i,

(ii) $\bar{c}_i(v_i) = \bar{c}_i(0) + \bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)\,dx$, for every $v_i \in [0, 1]$,

(iii) $\bar{c}_i(0) \leq 0$,

where the expression for the seller's expected revenue follows from incentive compatibility precisely as in the proof of Theorem 9.6.

It will be helpful to rearrange the expression for the seller's expected revenue.

$$\begin{aligned} R &= \sum_{i=1}^{N} \int_0^1 \left[\bar{p}_i(v_i)v_i - \int_0^{v_i} \bar{p}_i(x)\,dx \right] f_i(v_i)\,dv_i + \sum_{i=1}^{N} \bar{c}_i(0) \\ &= \sum_{i=1}^{N} \left[\int_0^1 \bar{p}_i(v_i)v_i f_i(v_i)\,dv_i - \int_0^1 \int_0^{v_i} \bar{p}_i(x) f_i(v_i)\,dx\,dv_i \right] + \sum_{i=1}^{N} \bar{c}_i(0). \end{aligned}$$

By interchanging the order of integration in the iterated integral (i.e., from $dxdv_i$ to dv_idx), we obtain

$$R = \sum_{i=1}^{N} \left[\int_0^1 \bar{p}_i(v_i) v_i f_i(v_i) dv_i - \int_0^1 \int_x^1 \bar{p}_i(x) f_i(v_i) dv_i dx \right] + \sum_{i=1}^{N} \bar{c}_i(0)$$

$$= \sum_{i=1}^{N} \left[\int_0^1 \bar{p}_i(v_i) v_i f_i(v_i) dv_i - \int_0^1 \bar{p}_i(x)(1 - F_i(x)) dx \right] + \sum_{i=1}^{N} \bar{c}_i(0).$$

By replacing the dummy variable of integration, x, by v_i, this can be written equivalently as

$$R = \sum_{i=1}^{N} \left[\int_0^1 \bar{p}_i(v_i) v_i f_i(v_i) dv_i - \int_0^1 \bar{p}_i(v_i)(1 - F_i(v_i)) dv_i \right] + \sum_{i=1}^{N} \bar{c}_i(0)$$

$$= \sum_{i=1}^{N} \int_0^1 \bar{p}_i(v_i) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] f_i(v_i) dv_i + \sum_{i=1}^{N} \bar{c}_i(0).$$

Finally, recalling that

$$\bar{p}_i(r_i) = \int_0^1 \cdots \int_0^1 p_i(r_i, v_{-i}) f_{-i}(v_{-i}) dv_{-i}$$

we may write

$$R = \sum_{i=1}^{N} \int_0^1 \cdots \int_0^1 p_i(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N$$

$$+ \sum_{i=1}^{N} \bar{c}_i(0),$$

or

$$R = \int_0^1 \cdots \int_0^1 \left\{ \sum_{i=1}^{N} p_i(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \right\} f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N$$

$$+ \sum_{i=1}^{N} \bar{c}_i(0). \tag{9.11}$$

So, our problem is to maximise (9.11) subject to the constraints (i)–(iii) above. For the moment, let us concentrate on the first term in (9.11), namely

$$\int_0^1 \cdots \int_0^1 \left\{ \sum_{i=1}^N p_i(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \right\} f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N. \quad (9.12)$$

Clearly, (9.12) would be maximised if the term in curly brackets were maximised for each vector of values $v_1, \ldots, v_N$. Now, because the $p_i(v_1, \ldots, v_N)$ are non-negative and sum to one or less, the $N+1$ numbers $p_1(v_1, \ldots, v_N), \ldots, p_N(v_1, \ldots, v_N)$, $1 - \sum_{i=1}^N p_i(v_1, \ldots, v_N)$ are non-negative and sum to one. So, the sum above in curly brackets, which can be rewritten as

$$\sum_{i=1}^N p_i(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] + \left(1 - \sum_{i=1}^N p_i(v_1, \ldots, v_N) \right) \cdot 0,$$

is just a weighted average of the $N+1$ numbers

$$\left[v_1 - \frac{1 - F_1(v_1)}{f_1(v_1)} \right], \ldots, \left[v_N - \frac{1 - F_N(v_N)}{f_N(v_N)} \right], 0.$$

But then the sum in curly brackets can be no larger than the largest of these bracketed terms if one of them is positive, and no larger than zero if all of them are negative. Suppose now that no two of the bracketed terms are equal to one another. Then, if we define

$$p_i^*(v_1, \ldots, v_N) = \begin{cases} 1, \text{ if } v_i - \frac{1-F_i(v_i)}{f_i(v_i)} > \max\left(0, v_j - \frac{1-F_j(v_j)}{f_j(v_j)}\right) \text{ for all } j \neq i, \\ 0, \text{ otherwise,} \end{cases} \quad (9.13)$$

it must be the case that

$$\sum_{i=1}^N p_i(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \leq \sum_{i=1}^N p_i^*(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right].$$

Therefore, if the bracketed terms are distinct with probability one, we will have

$$R = \int_0^1 \cdots \int_0^1 \left\{ \sum_{i=1}^N p_i(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \right\} f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N$$
$$+ \sum_{i=1}^N \bar{c}_i(0)$$

$$\leq \int_0^1 \cdots \int_0^1 \left\{ \sum_{i=1}^{N} p_i^*(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \right\} f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N$$
$$+ \sum_{i=1}^{N} \bar{c}_i(0),$$

for all incentive-compatible direct selling mechanisms $p_i(\cdot)$, $c_i(\cdot)$. For the moment, then, let us assume that the bracketed terms are distinct with probability one. We will introduce an assumption on the bidders' distributions that guarantees this shortly.[16]

Because constraint (iii) implies that each $\bar{c}_i(0) \leq 0$, we can also say that for all incentive-compatible direct selling mechanisms $p_i(\cdot)$, $c_i(\cdot)$, the seller's revenue can be no larger than the following upper bound:

$$R \leq \int_0^1 \cdots \int_0^1 \left\{ \sum_{i=1}^{N} p_i^*(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \right\} f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N. \tag{9.14}$$

We will now construct an incentive-compatible direct selling mechanism that *achieves* this upper bound. Consequently, this mechanism will maximise the seller's revenue, and so will be optimal for the seller.

To construct this optimal mechanism, let the probability assignment functions be the $p_i^*(v_1, \ldots, v_N)$, $i = 1, \ldots, N$, in (9.13). To complete the mechanism, we must define cost functions $c_i^*(v_1, \ldots, v_N)$, $i = 1, \ldots, N$. But constraint (ii) requires that for each v_i, bidder i's expected cost and probability of receiving the object, $\bar{c}_i^*(v_i)$ and $\bar{p}_i^*(v_i)$, be related as follows

$$\bar{c}_i^*(v_i) = \bar{c}_i^*(0) + \bar{p}_i^*(v_i)v_i - \int_0^{v_i} \bar{p}_i^*(x)dx.$$

Now, because the $\bar{c}_i^*$ and $\bar{p}_i^*$ are averages of the c_i^* and p_i^*, this required relationship between *averages* will hold if it holds for *each and every* vector of values $v_1, \ldots, v_N$. That is, (ii) is guaranteed to hold if we define the c_i^* as follows: for every $v_1, \ldots, v_N$,

$$c_i^*(v_1, \ldots, v_N) = c_i^*(0, v_{-i}) + p_i^*(v_1, \ldots, v_N)v_i - \int_0^{v_i} p_i^*(x, v_{-i})dx. \tag{9.15}$$

To complete the definition of the cost functions and to satisfy constraint (iii), we shall set $c_i^*(0, v_{-i}, \ldots, v_{-i}) = 0$ for all i and all $v_2, \ldots, v_n$. So, our candidate for a revenue-maximising, incentive-compatible direct selling mechanism is as follows: for

[16] The assumption is given in (9.18).

every $i = 1, \ldots, N$ and every $v_1, \ldots, v_N$

$$p_i^*(v_1, \ldots, v_N) = \begin{cases} 1, \text{ if } v_i - \frac{1-F_i(v_i)}{f_i(v_i)} > \max\left(0,\, v_j - \frac{1-F_j(v_j)}{f_j(v_j)}\right) \text{ for all } j \neq i, \\ 0, \text{ otherwise}; \end{cases} \tag{9.16}$$

and

$$c_i^*(v_1, \ldots, v_N) = p_i^*(v_1, \ldots, v_N)v_i - \int_0^{v_i} p_i^*(x, v_{-i})dx. \tag{9.17}$$

By construction, this mechanism satisfies constraints (ii) and (iii), and it achieves the upper bound for revenues in (9.14). To see this, simply substitute the p_i^* into (9.11) and recall that by construction $\bar{c}_i^*(0) = 0$ for every i. The result is that the seller's revenues are

$$R = \int_0^1 \cdots \int_0^1 \left\{ \sum_{i=1}^N p_i^*(v_1, \ldots, v_N) \left[v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \right] \right\} f_1(v_1) \ldots f_N(v_N) dv_1 \ldots dv_N,$$

their maximum possible value.

So, if we can show that our mechanism's probability assignment functions defined in (9.16) satisfy constraint (i), then this mechanism will indeed be the solution we are seeking.

Unfortunately, the p_i^* as defined in (9.16) need not satisfy (i). To ensure that they do, we need to restrict the distributions of the bidders' values. Consider, then, the following assumption: For every $i = 1, \ldots, N$

$$v_i - \frac{1 - F_i(v_i)}{f_i(v_i)} \quad \text{is } \textit{strictly increasing} \text{ in } v_i. \tag{9.18}$$

This assumption is satisfied for a number of distributions, including the uniform distribution. Moreover, you are asked to show in an exercise that it holds whenever each F_i is any convex function, not merely that of the uniform distribution.[17] Note that in addition to ensuring that (i) holds, this assumption also guarantees that the numbers $v_1 - (1 - F_1(v_1))/f_1(v_1), \ldots, v_N - (1 - F_N(v_N))/f_N(v_N)$ are distinct with probability one, a requirement that we earlier employed but had left unjustified until now.

Let us now see why (9.18) implies that (i) is satisfied. Consider some bidder i and some fixed vector of values, v_{-i}, for the other bidders. Now, suppose that $\bar{v}_i > \underline{v}_i$ and that $p_i^*(\underline{v}_i, v_{-i}) = 1$. Then, by the definition of p_i^*, it must be the case that $\underline{v}_i - (1 - F_i(\underline{v}_i))/f_i(\underline{v}_i)$ is positive and strictly greater than $v_j - (1 - F_j(v_j))/f_j(v_j)$ for all $j \neq i$. Consequently, because $v_i - (1 - F_i(v_i))/f_i(v_i)$ is strictly increasing it must

[17]When this assumption fails, the mechanism we have constructed here is not optimal. One can nevertheless construct the optimal mechanism, but we shall not do so here. Thus, the additional assumption we are making here is only for simplicity's sake.

also be the case that $\bar{v}_i - (1 - F_i(\bar{v}_i))/f_i(\bar{v}_i)$ is both positive and strictly greater than $v_j - (1 - Fj(v_j))/f_j(v_j)$ for all $j \neq i$, which means that $p_i^*(\bar{v}_i, v_{-i}) = 1$. Thus, we have shown that if $p_i^*(v_i, v_{-i}) = 1$, then $p_i^*(v_i', v_{-i}) = 1$ for all $v_i' > v_i$. But because p_i^* takes on either the value 0 or 1, $p_i^*(v_i, v_{-i})$ is non-decreasing in v_i for every v_{-i}. This in turn implies that $\bar{p}_i^*(v_i)$ is non-decreasing in v_i, so that constraint (i) is indeed satisfied.

In the end then, our hard work has paid off handsomely. We can now state the following.

THEOREM 9.7 ***An Optimal Selling Mechanism***

If N bidders have independent private values with bidder i's value drawn from the continuous positive density f_i satisfying (9.18), then the direct selling mechanism defined in (9.16) and (9.17) yields the seller the largest possible expected revenue.

9.4.4 A CLOSER LOOK AT THE OPTIMAL SELLING MECHANISM

Let us see if we can simplify the description of the optimal selling mechanism by studying its details. There are two parts to the mechanism, the manner in which it allocates the object – the p_i^* – and the manner in which it determines payments – the c_i^*.

The allocation portion of the optimal mechanism is straightforward. Given the reported values $v_1, \ldots, v_N$, the object is given to the bidder i whose $v_i - (1 - F_i(v_i))/f_i(v_i)$ is strictly highest and positive. Otherwise, the seller keeps the object. But it is worth a little effort to try to interpret this allocation scheme.

What we shall argue is that $v_i - (1 - F_i(v_i))/f_i(v_i)$ represents the marginal revenue, $MR_i(v_i)$, that the seller obtains from increasing the probability that the object is assigned to bidder i when his value is v_i. To see this without too much notation we shall provide an intuitive argument. Consider the effect of increasing the probability that the object is awarded bidder i when his value is v_i. This enables the seller to increase the cost to v_i so as to leave his utility unchanged. Because the density of v_i is $f_i(v_i)$, the seller's revenue increases at the rate $v_i f_i(v_i)$ as a result of this change. On the other hand, incentive compatibility forces a connection between the probability that the good is assigned to bidder i with value v_i and the cost assessed to all higher values $v_i' > v_i$. Indeed, according to constraint (ii), increasing the probability that lower values receive the object reduces one-for-one the cost that all higher values can be assessed. Because there is a mass of $1 - F_i(v_i)$ values above v_i, this total reduction in revenue is $1 - F_i(v_i)$. So, altogether the seller's revenues increase by $v_i f_i(v_i) - (1 - F_i(v_i))$. But this is the total effect due to the density $f_i(v_i)$ of values equal to v_i. Consequently, the marginal revenue associated with each v_i is $MR_i(v_i) = v_i - (1 - F_i(v_i))/f_i(v_i)$.

The allocation rule now makes perfect sense. If $MR_i(v_i) > MR_j(v_j)$, the seller can increase revenue by reducing the probability that the object is assigned to bidder j and increasing the probability that it is assigned to bidder i. Clearly then, the seller maximises his revenue by assigning all probability (i.e., probability one) to the bidder with the highest $MR_i(v_i)$, so long as it is positive. If all the marginal revenues are negative, the seller does best by reducing all of the bidders' probabilities to zero, i.e., the seller keeps the object.

The payment portion of the mechanism is a little less transparent. To get a clearer picture of what is going on, suppose that when the (truthfully) reported values are $v_1, \ldots, v_N$, bidder i does not receive the object, i.e., that $p_i^*(v_i, v_{-i}) = 0$. What must bidder i pay according to the mechanism? The answer, according to (9.17), is

$$\begin{aligned} c_i^*(v_i, v_{-i}) &= p_i^*(v_i, v_{-i})v_i - \int_0^{v_i} p_i^*(x, v_{-i})\,dx \\ &= 0 \cdot v_i - \int_0^{v_i} p_i^*(x, v_{-i})\,dx. \end{aligned}$$

But recall that, by virtue of assumption (9.18), $p_i^*(\cdot, v_{-i})$ is non-decreasing. Consequently, because $p_i^*(v_i, v_{-i}) = 0$, it must be the case that $p_i^*(x, v_{-i}) = 0$ for every $x \le v_i$. Hence the integral above must be zero so that

$$c_i^*(v_i, v_{-i}) = 0.$$

So, we have shown that according to the optimal mechanism, if bidder i does not receive the object, he pays nothing.

Suppose now that bidder i does receive the object, i.e., that $p_i^*(v_i, v_{-i}) = 1$. According to (9.17), he then pays

$$\begin{aligned} c_i^*(v_i, v_{-i}) &= p_i^*(v_i, v_{-i})v_i - \int_0^{v_i} p_i^*(x, v_{-i})\,dx \\ &= v_i - \int_0^{v_i} p_i^*(x, v_{-i})\,dx. \end{aligned}$$

Now, because p_i^* takes on the value 0 or 1, is non-decreasing and continuous from the left in i's value, and $p_i^*(v_i, v_{-i}) = 1$, there must be a largest value for bidder i, $r_i^* < v_i$, such that $p_i^*(r_i^*, v_{-i}) = 0$. Note that r_i^* will generally depend on v_{-i} so it would be more explicit to write $r_i^*(v_{-i})$. Note then that by the very definition of $r_i^*(v_{-i})$, $p_i^*(x, v_{-i})$ is equal to 1 for every $x > r_i^*(v_{-i})$, and is equal to 0 for every $x \le r_i^*(v_{-i})$. But this means that

$$\begin{aligned} c_i^*(v_i, v_{-i}) &= v_i - \int_{r_i^*(v_{-i})}^{v_i} 1\,dx \\ &= v_i - (v_i - r_i^*(v_{-i})) \\ &= r_i^*(v_{-i}). \end{aligned}$$

So, when bidder i wins the object, *he pays a price,* $r_i^*(v_{-i})$, *that is independent of his own reported value.* Moreover, the price he pays is the maximum value he could have reported, given the others' reported values, without receiving the object.

Putting all of this together, we may rephrase the revenue-maximising selling mechanism defined by (9.16) and (9.17) in the following manner.

THEOREM 9.8 ***The Optimal Selling Mechanism Simplified***

If N bidders have independent private values with bidder i's value drawn from the continuous positive density f_i and each $v_i - (1 - F_i(v_i))/f_i(v_i)$ is strictly increasing, then the following direct selling mechanism yields the seller the largest possible expected revenue:

For each reported vector of values, $v_1, \ldots, v_N$, the seller assigns the object to the bidder i whose $v_i - (1 - F_i(v_i))/f_i(v_i)$ is strictly largest and positive. If there is no such bidder, the seller keeps the object and no payments are made. If there is such a bidder i, then only this bidder makes a payment to the seller in the amount r_i^*, where $r_i^* - (1 - F_i(r_i^*))/f_i(r_i^*) = 0$ or $\max_{j \neq i} v_j - (1 - F_j(v_j))/f_j(v_j)$, whichever is largest. Bidder i's payment, r_i^*, is, therefore, the largest value he could have reported, given the others' reported values, without receiving the object.

As we know, this mechanism is incentive-compatible. That is, truth-telling is a Nash equilibrium. But, in fact, the incentive to tell the truth in this mechanism is much stronger than this. In this mechanism it is, in fact, a *dominant strategy* for each bidder to report his value truthfully to the seller; even if the other bidders do not report their values truthfully, bidder i can do no better than to report his value truthfully to the seller. You are asked to show this in one of the exercises.

One drawback of this mechanism is that to implement it, the seller must know the distributions, F_i, from which the bidders' values are drawn. This is in contrast to the standard auctions that the seller can implement without any bidder information whatsoever. Yet there is a connection between this optimal mechanism and the four standard auctions that we now explore.

9.4.5 EFFICIENCY, SYMMETRY, AND COMPARISON TO THE FOUR STANDARD AUCTIONS

In the optimal selling mechanism, the object is not always allocated efficiently. Sometimes the bidder with the highest value does not receive the object. In fact, there are *two ways* that inefficiency can occur in the optimal selling mechanism. First, the outcome can be inefficient because the seller sometimes keeps the object, even though his value for it is zero and all bidders have positive values. This occurs when every bidder i's value v_i is such that $v_i - (1 - F_i(v_i))/f_i(v_i) \leq 0$. Second, even when the seller does assign the object to one of the bidders, it might not be assigned to the bidder with the highest value. To see this, consider the case of two bidders, 1 and 2. If the bidders are asymmetric, then for some $v \in [0, 1]$, $v - (1 - F_1(v))/f_1(v) \neq v - (1 - F_2(v))/f_2(v)$. Indeed, let us suppose that for this particular value, v, $v - (1 - F_1(v))/f_1(v) > v - (1 - F_2(v))/f_2(v) > 0$. Consequently, when both bidders' values are v, bidder 1 will receive the object. But, by continuity, even if bidder 1's value falls slightly to $v' < v$, so long as v' is close enough to v, the inequality $v' - (1 - F_1(v'))/f_1(v') > v - (1 - F_2(v))/f_2(v) > 0$ will continue to hold. Hence, bidder 1 will receive the object even though his value is strictly below that of bidder 2.

The presence of inefficiencies is not surprising. After all, the seller is a monopolist seeking maximal profits. In Chapter 4, we saw that a monopolist will restrict output below the efficient level so as to command a higher price. The same effect is present here. But, because there is only one unit of an indivisible object for sale, the seller here restricts supply by sometimes keeping the object, depending on the vector of reports. But this accounts for only the first kind of inefficiency. The second kind of inefficiency that arises here did not occur in our brief look at monopoly in Chapter 4. The reason is that there we assumed that the monopolist was unable to distinguish one consumer from another. Consequently, the monopolist had to charge all consumers the same price. Here, however, we are assuming that the seller *can* distinguish bidder i from bidder j and that the seller knows that i's distribution of values is F_i and that j's is F_j. This additional knowledge allows the monopolist to discriminate between the bidders, which leads to higher profits.

Let us now eliminate this second source of inefficiency by supposing that bidders are symmetric. Because the four standard auctions all yield the same expected revenue for the seller under symmetry, this will also allow us to compare the standard auctions with the optimal selling mechanism.

How does symmetry affect the optimal selling mechanism? If the bidders are symmetric, then $f_i = f$ and $F_i = F$ for every bidder i. Consequently, the optimal selling mechanism is as follows: if the vector of reported values is $(v_1, \ldots, v_N)$, the bidder i with the highest positive $v_i - (1 - F(v_i))/f(v_i)$ receives the object and pays the seller r_i^*, the largest value he could have reported, given the other bidder's reported values, without winning the object. If there is no such bidder i, the seller keeps the object and no payments are made.

But let us think about this for a moment. Because we are assuming that $v - (1 - F(v))/f(v)$ is strictly increasing in v, the object is actually awarded to the bidder i with the strictly highest value v_i, so long as $v_i - (1 - F_i(v_i))/f_i(v_i) > 0$ – that is, so long as $v_i > \rho^* \in [0, 1]$, where

$$\rho^* - \frac{1 - F(\rho^*)}{f(\rho^*)} = 0. \tag{9.19}$$

(You are asked to show in an exercise that a unique such ρ^* is guaranteed to exist.)

Now, how large can bidder i's reported value be before he is awarded the object? Well, he does not get the object unless his reported value is strictly highest and strictly above ρ^*. So, the largest his report can be without receiving the object is the largest of the other bidders' values or ρ^*, whichever is larger. Consequently, when bidder i does receive the object he pays either ρ^* or the largest value reported by the other bidders, whichever is larger.

Altogether then, the optimal selling mechanism is as follows: the bidder whose reported value is strictly highest and strictly above ρ^* receives the object and pays the larger of ρ^* and the largest reported value of the other bidders.

Remarkably, this optimal direct selling mechanism can be mimicked by running a second-price auction with reserve price ρ^*. That is, an auction in which the bidder with the highest bid strictly above the reserve price wins and pays the second-highest bid or the reserve price, whichever is larger. If no bids are above the reserve price, the seller keeps the object and no payments are made. This is optimal because, just as in a standard second-price auction, it is a dominant strategy to bid one's value in a second-price auction with a reserve price.

This is worth highlighting.

THEOREM 9.9 ***An Optimal Auction Under Symmetry***

If N bidders have independent private values, each drawn from the same continuous positive density f, where $v - (1 - F(v))/f(v)$ is strictly increasing, then a second price auction with reserve price ρ^* satisfying $\rho^* - (1 - F(\rho^*))/f(\rho^*) = 0$, maximises the seller's expected revenue.

You might wonder about the other three standard auctions. Will adding an appropriate reserve price render these auctions optimal for the seller too? The answer is yes, and this is left for you to explore in the exercises.

So, we have now come full circle. The four standard auctions – first-price, second-price, Dutch, and English – all yield the same revenue under symmetry. Moreover, by supplementing each by an appropriate reserve price, the seller maximises his expected revenue. Is it any wonder then that these auctions are in such widespread use?

9.5 Designing Allocatively Efficient Mechanisms

We now turn our attention away from profit maximisation and towards allocative efficiency. The basic question is how to achieve a Pareto-efficient outcome when critical pieces of information are held privately by individuals in society. Such information might include, for example, individual preferences, production costs, income, etc.

As in Chapter 6, we allow for a broad collection of circumstances by letting X denote the set of social states. To keep matters simple, we assume that X is finite. Once again, the members of X might be allocations in an exchange or production economy, candidates running for office, etc. We also introduce a distinguished good called 'money', whose role will be apparent shortly. Individuals will care about the social state $x \in X$ as well as about how much money they have. Thus, the social state does not completely describe all that is utility-relevant for individuals. For any fixed social state, an individual can use his money to purchase desirable commodities that are independent of, and have no effect upon, the social state.[18]

There are N individuals in society. To capture the idea that they might have critical pieces of private information, we introduce a set of possible 'types' for each individual. Let T_i denote the finite set of types of individual i. As in our Chapter 7 analysis of Bayesian games, we introduce probabilities over the players' types. In particular, we assume here that there is a common prior, q, where $q(t) > 0$ is the probability that the vector of

[18]It is also possible to interpret 'money' instead as a separate commodity that individuals directly desire. But we will stick with the monetary interpretation.

types of the N individuals is $t = (t_1, \ldots, t_N) \in T = \times_{i=1}^{N} T_i$. Moreover, we assume that the types are independent, so that $q(t) = q_1(t_1) \cdots q_N(t_N)$. Consequently, no individual's type provides any information about other individuals' types.[19]

9.5.1 QUASI-LINEAR UTILITY AND PRIVATE VALUES

For the remainder of this chapter, we shall restrict the domain of individual preferences to those that can be represented by *quasi-linear utility functions*. That is, if individual i has m dollars and $x \in X$ is the social state, his von Neumann-Morgenstern utility is

$$v_i(x, t_i) + m$$

when his type is $t_i \in T_i$.

Because we are interpreting m as money, $v_i(x, t_i)$ is correctly interpreted as the value, in dollars, that individual i places on the social state x when his type is t_i. Also, note that individual i's value for the social state $x \in X$ depends only on his type t_i and not on the types of the other individuals. Therefore, each individual has private values. Consequently, just as in Section 9.2, this is an *independent private values* model.[20] Let us consider an example.

EXAMPLE 9.3 Consider a small town with N individuals. The town has been selected by the state to receive either a new swimming pool (S) or a new bridge (B) and must decide which it wants. Thus, there are two social states, S and B, and therefore the set of social states is $X = \{S, B\}$. Each individual i in the town has quasi-linear preferences and has private information t_i regarding the value he places on the pool and on the bridge. Specifically, the values individual i places on the swimming pool (S) and on the bridge (B) are given by,

$$v_i(x, t_i) = \begin{cases} t_i + 5, & \text{if } x = S \\ 2t_i, & \text{if } x = B \end{cases},$$

where his type t_i is equally likely to take on any of the values $1, 2, \ldots, 9$ and where the types are independent across individuals.

Each individual is therefore as likely to strictly prefer the swimming pool over the bridge (i.e., $t_i \in \{1, 2, 3, 4\}$) as he is to strictly prefer the bridge over the swimming pool (i.e., $t_i \in \{6, 7, 8, 9\}$. Only the individual himself knows which of these is the case and by how much he prefers one social state over the other. And the more extreme an individual's type, the more he prefers one of the social states over the other. □

Quasi-linearity is a strong assumption. It implies that there is a common rate at which utility can be substituted across individuals regardless of the social state and regardless of

[19] None of our analysis here depends on finite type spaces. In particular, all of our formulae and conclusions are valid when, for example, each T_i is a Euclidean cube and each $q_i(t_i)$ is the probability density that i's type is t_i. In that case, summations over T, T_{-i} and T_i become integrals.

[20] In fact the single object model of Section 9.2 is itself a special case of a quasi-linear independent private values model.

individual types. Its advantage is that it yields a convenient characterisation of efficient social states. To see this most clearly, suppose that individuals have no private information, i.e., suppose there are no types. Individual i's utility function is then simply $v_i(x) + m$. Even though individuals care about both the social state and the amount of money they have, it turns out that the amount of money they end up with is more or less irrelevant to determining which social states are compatible with Pareto efficiency. Indeed, we claim the following:

With quasi-linear preferences, a social state $\hat{x} \in X$ is Pareto efficient if and only if it maximises the sum of the non-monetary parts of individual utilities, i.e., if and only if it solves,

$$\max_{x \in X} \sum_{i=1}^{N} v_i(x). \tag{9.20}$$

Let us see why this claim is true. Suppose, for example, that the social state happens to be $x \in X$ but that $y \in X$ satisfies,

$$\sum_{i=1}^{N} v_i(y) > \sum_{i=1}^{N} v_i(x). \tag{9.21}$$

We would like to show that a Pareto improvement is available. In fact, we shall show that a Pareto improvement can be obtained by switching the social state from x to y.

Now, even though (9.21) holds, merely switching the social state from x to y need not result in a Pareto improvement because some individual utilities may well fall in the move from x to y. The key idea is to compensate those individuals whose utilities fall by transferring to them income from individuals whose utilities rise. It is here where the common rate at which income translates into utility across individuals is absolutely central.

For each individual i, define the income transfer, τ_i, as follows:

$$\tau_i = v_i(x) - v_i(y) + \frac{1}{N}\sum_{i=1}^{N}(v_i(y) - v_i(x)).$$

If $\tau_i > 0$, then individual i receives τ_i dollars while if $\tau_i < 0$ individual i is taxed τ_i dollars. By construction, the τ_i sum to zero and so these are indeed income transfers among the N individuals.

After changing the state from x to y and carrying out the income transfers, the change in individual i's utility is,

$$v_i(y) + \tau_i - v_i(x),$$

which, by (9.21), is strictly positive. Hence, each individual is strictly better off after the change. This proves that the social state x is not Pareto efficient and establishes the 'only if' part of the claim in (9.20).[21] You are asked to establish the 'if' part of the claim in Exercise 9.26.

9.5.2 EX POST PARETO EFFICIENCY

There are several stages at which economists typically think about Pareto efficiency, the *ex ante stage* prior to individuals finding out their types, the *interim stage*, where each knows only his own type, and the *ex post stage* when all types are known by all individuals.

In general, the more uncertainty there is, the greater is the scope for mutually beneficial insurance. Hence, we expect ex ante Pareto efficiency to imply interim Pareto efficiency to imply ex post Pareto efficiency. We will focus here only on the latter, ex post Pareto efficiency.

Because individual preferences over social states are a function of individual types, $t_1, \ldots, t_N$, achieving ex post Pareto-efficient outcomes will typically require the social state to depend upon individual types. With this in mind, call a function $x\colon T \to X$ an *allocation function.* Thus, an allocation function specifies, for each vector of individual types, a social state in X.[22]

Analogous to the claim established in (9.20), ex post Pareto efficiency of an allocation function is characterised by maximisation of the sum of individual utilities. This leads us to the following definition.

DEFINITION 9.3 ***Ex Post Pareto Efficiency***

An allocation function $\hat{x}\colon T \to X$ is ex post Pareto-efficient if for each $t \in T$, $\hat{x}(t) \in X$ solves,

$$\max_{x \in X} \sum_{i=1}^{N} v_i(x, t_i),$$

where the maximisation is over all social states. We then also say that $\hat{x}(t)$ is an ex post efficient social state given $t \in T$.

Thus, $\hat{x}\colon T \to X$ is ex post Pareto efficient if for each type vector $t \in T$, the social state, $\hat{x}(t)$, maximises the sum of individual ex post utilities given t.

9.5.3 DIRECT MECHANISMS, INCENTIVE COMPATIBILITY AND THE REVELATION PRINCIPLE

Our question of interest is whether it is possible to always achieve an ex post efficient allocation, despite the fact that individual utilities are private information. How might we go about achieving this goal? The possibilities are in fact rather daunting. For example, we

[21] We implicitly assume that individuals whose transfers are negative, i.e., who are taxed, have sufficient income to pay the tax.

[22] We do not instead call this a social choice function, as in Section 6.5 of Chapter 6, because we do not require the range of $x(\cdot)$ to be all of X.

might ask individuals one at a time to publicly announce their type (of course they might lie). We might then ask whether anyone believes that someone lied about their type, suitably punishing (via taxes) those whose announcements are doubted by sufficiently many others – the hope being that this might encourage honest reports. On the other hand, we might not ask individuals to report their types at all. Rather, we might ask them to vote directly for the social state they would like implemented. But what voting system ought we employ? Plurality rule? Pairwise majority with ties broken randomly? Should the votes be by secret ballot? Or public and sequential? As you can sense, we could go on and on. There are endless possibilities for designing a system, or mechanism, in the pursuit of achieving our goal.

Fortunately, just as in the single-good revenue-maximisation setting, the **revelation principle** applies here and it allows us to limit our search to the set of incentive-compatible direct mechanisms. Before we discuss this second application of the revelation principle any further, it is useful to have on record two definitions. They are the extensions of Definitions 9.1 and 9.2 to the present more general setup.

DEFINITION 9.4 ***Direct Mechanisms***

A direct mechanism consists of a collection of probability assignment functions,

$$p^x(t_1, \ldots, t_N), \text{ one for each } x \in X,$$

and N cost functions,

$$c_1(t_1, \ldots, t_N), \ldots, c_N(t_1, \ldots, t_N).$$

For every vector of types $(t_1, \ldots, t_N) \in T$ reported by the N individuals, $p^x(t_1, \ldots, t_N) \in [0, 1]$ denotes the probability that the social state is $x \in X$ and $c_i(t_1, \ldots, t_N) \in \mathbb{R}$ denotes individual i's cost, i.e., the amount he must pay. Because some social state must be chosen, we require $\sum_{x \in X} p^x(t_1, \ldots, t_N) = 1$ for every $(t_1, \ldots, t_N) \in T$.

Because of the similarity between Definitions 9.1 and 9.4, there is little need for further discussion except to say that Definition 9.4 becomes equivalent to Definition 9.1 when (i) there is a single object available, (ii) there are $N+1$ individuals consisting of N bidders and one seller, and (iii) the social states are the $N+1$ allocations in which, either, one of the bidders ends up with the good or the seller ends up with the good. (See Exercise 9.28.)

Given a direct mechanism, $p, c_1, \ldots, c_N$, it is useful to define, as in Section 9.3, $u_i(r_i, t_i)$ to be individual i's expected utility from reporting that his type is $r_i \in T_i$ when his true type is $t_i \in T_i$ and given that all other individuals always report their types truthfully. That is,

$$u_i(r_i, t_i) = \sum_{t_{-i} \in T_{-i}} q_{-i}(t_{-i}) \left(\sum_{x \in X} p^x(r_i, t_{-i}) v_i(x, t_i) - c_i(r_i, t_{-i}) \right),$$

where $q_{-i}(t_{-i}) = \Pi_{j\neq i} q_j(t_j)$. As before, we can simplify this formula by defining

$$\bar{p}_i^x(r_i) = \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i})p^x(r_i, t_{-i}),$$

and

$$\bar{c}_i(r_i) = \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i})c_i(r_i, t_{-i}). \tag{9.22}$$

Then,

$$u_i(r_i, t_i) = \sum_{x\in X} \bar{p}_i^x(r_i)v_i(x, t_i) - \bar{c}_i(r_i) \tag{9.23}$$

DEFINITION 9.5 ***Incentive-Compatible Direct Mechanisms***

A direct mechanism is incentive-compatible if when the other individuals always report their types truthfully, each individual's expected utility is maximised by always reporting his type truthfully, i.e., the mechanism is incentive-compatible if for each individual i and for each of his types $t_i \in T_i$, $u_i(r_i, t_i)$ as defined in (9.23) is maximised in $r_i \in T_i$ when $r_i = t_i$. Or put yet another way, the mechanism is incentive-compatible if it is a Bayesian-Nash equilibrium for each individual to always report his type truthfully.[23]

With these definitions in hand, it is worthwhile to informally discuss how the revelation principle allows us to reduce our search to the set of incentive-compatible direct mechanisms. So, suppose that we manage to design some, possibly quite complex, extensive form game for individuals in society to play, where the payoffs to the individuals at the endpoints are defined by the utility they receive from some social state and income distribution at that endpoint. Because the strategies they choose to adopt may depend upon their types, any 'equilibrium' they play (i.e., Nash, subgame-perfect, sequential) will be a Bayesian-Nash equilibrium of the game's strategic form. Suppose that in some such Bayesian-Nash equilibrium, an ex post efficient social state is always certain to occur. We would then say that the given extensive form game (mechanism) successfully implements an ex post efficient outcome. According to the revelation principle, a direct incentive-compatible mechanism can do precisely the same thing. Here's how. Instead of having the individuals play their strategies themselves, design a new (direct) mechanism that simply plays their strategies for them after they report their types. Consequently, if the other individuals always report honestly, then, from your perspective, it is as if you are participating in the original extensive form game against them. But in that game, it was optimal for you to carry out the actions specified by your strategy conditional on your actual type. Consequently, it is optimal for you to report your type truthfully in the new direct mechanism so that those same actions are carried out on your behalf. Hence, the new direct

[23] Because the type spaces T_i are finite here, our Chapter 7 definition of a Bayesian-Nash equilibrium applies. If the type spaces are infinite we would simply define the truth-telling equilibrium to be a Bayesian-Nash equilibrium.

mechanism is incentive-compatible and always yields the same ex post efficient social state and income distribution as would the old. That's all there is to it!

As a matter of terminology, we call an incentive-compatible direct mechanism *ex post efficient* if it assigns probability one to a set of ex post efficient social states given any vector of reported types $t \in T$, i.e., if for every $t \in T$, $p^x(t) > 0$ implies $x \in X$ is ex post efficient when the vector of types is t.

9.5.4 THE VICKREY-CLARKE-GROVES MECHANISM

We now introduce one of the most important direct mechanisms in the theory of mechanism design, the Vickrey-Clarke-Groves (VCG) mechanism. As we shall see, it plays a central role in the theory we shall develop here. In particular, it will solve the ex post efficient allocation problem that we have so far set for ourselves.

An interesting feature of the VCG mechanism is that it can be thought of as a generalisation of a second-price auction. Recall that in a second-price auction for a single good, the highest bidder wins and pays the second-highest bid. As we know, it is therefore a dominant strategy for each bidder to bid his value, and so the bidder with highest value wins and pays the second highest value. This auction is sometimes described as one in which the winner 'pays his externality'. The reason is that if the winner were not present, the bidder with second highest value would have won. Thus, the winning bidder, by virtue of his presence, precludes the second-highest value from being realised – he imposes an externality. Of course, he pays for the good precisely the amount of the externality he imposes, and the end result is efficient.

The 'paying one's externality' idea generalises nicely to our current situation as follows. Let $\hat{x}\colon T \to X$ be an ex post efficient allocation function. That is, for each $t \in T$, let $\hat{x}(t)$ be a solution to,

$$\max_{x \in X} \sum_{i=1}^{N} v_i(x, t_i).$$

Such a solution always exists because X is finite. If there are multiple solutions choose any one of them. The ex post efficient allocation function $\hat{x}(\cdot)$ is therefore well-defined and it will remain fixed for the rest of this chapter.

Let us think about the externality imposed by each individual i on the remaining individuals under the assumption that ex post efficiency can be achieved. The trick to computing individual i's externality is to think about the difference his presence makes to the total utility of the others.

When individual i is present and the vector of types is $t \in T$, the social state is $\hat{x}(t)$ and the total utility of the others is,[24]

$$\sum_{j \neq i} v_j(\hat{x}(t), t_j).$$

[24]We can safely ignore the income individuals may have because we will ultimately be interested only in utility differences between different outcomes and therefore individual incomes will always cancel. In other words, it is harmless to compute utilities as if initial individual incomes are zero.

That was simple enough. But what is the total utility of the others when individual i is *not* present? This too is straightforward if we assume that in the absence of individual i – i.e., if society consists only of the $N-1$ individuals $j \neq i$ – the social state is chosen in an ex post efficient manner *relative to those who remain*.

For each $t_{-i} \in T_{-i}$, let $\tilde{x}^i(t_{-i}) \in X$ solve,

$$\max_{x \in X} \sum_{j \neq i} v_j(x, t_j).$$

That is, $\tilde{x}^i \colon T_{-i} \to X$ is an ex post efficient allocation function in the society without individual i.

It is now a simple matter to compute the difference that i's presence makes to the total utility of the others. Evidently, when the type vector is $t \in T$, the difference in the utility of the others when i is not present as compared to when he is present is,

$$\sum_{j \neq i} v_j(\tilde{x}^i(t_{-i}), t_j) - \sum_{j \neq i} v_j(\hat{x}(t), t_j).$$

Call this difference the *externality imposed by individual i.*

Note that one's externality is always non-negative and is typically positive because, by definition, $\tilde{x}^i(t_{-i}) \in X$ maximises the sum of utilities of individuals $j \neq i$ when their vector of types is t_{-i}. You should convince yourself that, in the case of a single good, each individual's externality is zero except for the individual with highest value, whose externality is the second highest value – just as it should be.

Consider now the following important mechanism, called the *VCG mechanism* after Vickrey, Clarke, and Groves, who independently provided important contributions leading to its development.[25]

DEFINITION 9.6 ***The Vickrey-Clarke-Groves Mechanism***

Each individual simultaneously reports his type to the designer. If the reported vector of types is $t \in T$, the social state $\hat{x}(t)$ is chosen. In addition, each individual i is assessed a monetary cost equal to,

$$c_i^{VCG}(t) = \sum_{j \neq i} v_j(\tilde{x}^i(t_{-i}), t_j) - \sum_{j \neq i} v_j(\hat{x}(t), t_j).$$

That is, each individual pays his externality based on the reported types. The c_i^{VCG} are called the VCG cost functions.

The key idea behind the VCG mechanism is to define individual costs so that each individual internalises the externality that, through his report, he imposes on the rest of society. Let us return to Example 9.3 to see what the VCG mechanism looks like there.

[25] See Vickrey (1961), Clarke (1971) and Groves (1973).

EXAMPLE 9.4 Consider the situation in Example 9.3. If the vector of reported types is $t \in T$, then it is efficient for the town to build the bridge if $\sum_i v_i(B, t_i) > \sum_i v_i(S, t_i)$.[26] Given the definition of the v_i, this leads to the following ex post efficient allocation function. For each $t \in T$,

$$\hat{x}(t) = \begin{cases} B, & \text{if } \sum_{i=1}^{N}(t_i - 5) > 0 \\ S, & \text{otherwise.} \end{cases}$$

According to the VCG mechanism, if the reported vector of types is $t \in T$, then the social state is $\hat{x}(t)$. It remains to describe the cost, $c_i^{VCG}(t)$, individual i must pay. Let us think about the externality that individual i imposes on the others. Suppose, for example, that the others report very high types, e.g., $t_j = 9$ for all $j \neq i$. Then, if there are at least two other individuals, the bridge will be built regardless of i's report. Indeed, the bridge will be built whether or not individual i is present. Hence, individual i's externality, and so also his cost, is zero in this case. Similarly, i's externality and cost will be zero whenever his presence does not change the outcome. With this in mind, let us say that individual i is pivotal for the social state $x \in \{S, B\}$ at the type vector $t \in T$ when, given reports t, his presence changes the social state from x' to x. For example, individual i is pivotal for B at $t \in T$ if $\sum_{j=1}^{N}(t_j - 5) > 0$ and $\sum_{j \neq i}(t_j - 5) \leq 0$, because the first (strict) inequality implies that the social state is B when he is present and the second (weak) inequality implies it is S when he is absent. In this circumstance, i's externality and cost is $c_i^{VCG}(t) = \sum_{j \neq i}(t_j + 5) - \sum_{j \neq i} 2t_j$, i.e., the difference between the others' total utility when he is absent and their total utility when he is present. Altogether then, $c_i^{VCG}(t)$ is as follows,

$$c_i^{VCG}(t) = \begin{cases} \sum_{j \neq i}(5 - t_j), & \text{if } i \text{ is pivotal for } B \text{ at } t \in T, \\ \sum_{j \neq i}(t_j - 5), & \text{if } i \text{ is pivotal for } S \text{ at } t \in T, \\ 0, & \text{otherwise.} \end{cases}$$

□

So far so good, but will the VCG mechanism actually succeed in implementing an ex post efficient outcome? By construction, the mechanism chooses an outcome that is ex post efficient based on the reported vector of types. However, individuals are free to lie about their types, and, if they do, the outcome will typically *not* be ex post efficient with respect to the *actual* vector of types. Hence, for this mechanism to work, it must induce individuals to report their types truthfully. Our next result establishes that the VCG mechanism does indeed do so.

THEOREM 9.10 ***Truth-Telling is Dominant in the VCG Mechanism***

In the VCG mechanism it is a weakly dominant strategy for each individual to report his type truthfully. Hence, the VCG mechanism is incentive-compatible and ex post efficient.

[26] We assume that the swimming pool is built if the two sums are equal.

Proof: We must show that truthful reporting is a weakly dominant strategy for an arbitrary individual i. Suppose then that the others report $t_{-i} \in T_{-i}$, which need not be truthful. Suppose also that individual i's type is $t_i \in T_i$ and that he reports $r_i \in T_i$. His utility would then be,[27]

$$v_i(\hat{x}(r_i, t_{-i}), t_i) - c_i^{VCG}(r_i, t_{-i}). \tag{P.1}$$

Note that $\hat{x}(\cdot)$ and $c^{VCG}(\cdot)$ are evaluated at i's reported type, r_i, while $v_i(x, \cdot)$ is evaluated at i's actual type, t_i. We must show that (P.1) is maximised when individual i reports truthfully, i.e., when $r_i = t_i$.

Substituting the definition of $c_i^{VCG}(r_i, t_{-i})$ into (P.1), i's utility can be written as,

$$\begin{aligned} v_i(\hat{x}(r_i, t_{-i}), t_i) - c_i^{VCG}(r_i, t_{-i}) &= v_i(\hat{x}(r_i, t_{-i}), t_i) - \Bigg(\sum_{j \neq i} v_j(\tilde{x}^i(t_{-i}), t_j) \\ &\qquad - \sum_{j \neq i} v_j(\hat{x}(r_i, t_{-i}), t_j)\Bigg) \\ &= \sum_{j=1}^{N} v_j(\hat{x}(r_i, t_{-i}), t_j) - \sum_{j \neq i} v_j(\tilde{x}^i(t_{-i}), t_j). \end{aligned} \tag{P.2}$$

Hence, we must show that setting $r_i = t_i$ maximises the right-hand side of the second equality (P.2). To see why this is so, note that r_i appears only in the first summation there and so it suffices to show that,

$$\sum_{j=1}^{N} v_j(\hat{x}(t_i, t_{-i}), t_j) \geq \sum_{j=1}^{N} v_j(\hat{x}(r_i, t_{-i}), t_j), \text{ for all } t_{-i} \in T_{-i}. \tag{P.3}$$

But by the definition of $\hat{x}(t_i, t_{-i})$,

$$\sum_{j=1}^{N} v_j(\hat{x}(t_i, t_{-i}), t_j) \geq \sum_{j=1}^{N} v_j(x, t_j), \text{ for all } x \in X.$$

Hence, in particular, (P.3) is satisfied because $\hat{x}(r_i, t_{-i}) \in X$ for all $r_i \in T_i$. ∎

To test your understanding of this proof and also of the VCG mechanism, you should try to show, with and without the aid of the proof, that it is a dominant strategy to tell the truth in the VCG mechanism that is explicitly defined in Example 9.4.

Several remarks are in order. First, because each individual's cost, $c_i^{VCG}(t)$, is always non-negative, the mechanism never runs a deficit and typically runs a surplus.

[27] We can safely ignore i's initial level of income since it simply adds a constant to all of our utility calculations.

Second, one might therefore wonder whether any individual would prefer to avoid paying his cost by not participating in the mechanism. To properly address this question we must specify what would happen if an individual were to choose not to participate. An obvious specification is to suppose that the VCG mechanism would be applied as usual, but only to those who do participate. With this in mind, we can show that it is an equilibrium for all N individuals to participate.

If all individuals participate and report truthfully (a dominant strategy), then individual i's payoff when the vector of types is t is,

$$v_i(\hat{x}(t), t_i) - c_i^{VCG}(t) = \sum_{j=1}^{N} v_j(\hat{x}(t), t_j) - \sum_{j \neq i} v_j(\tilde{x}^i(t_{-i}), t_j). \tag{9.24}$$

On the other hand, if individual i chooses not to participate, he avoids paying the cost $c_i^{VCG}(t)$, but the social state becomes instead $\tilde{x}^i(t_{-i})$, i.e., an ex post efficient social state for the $N-1$ participating individuals who report their types. Consequently, if individual i chooses not to participate his utility will be

$$v_i(\tilde{x}^i(t_{-i}), t_i). \tag{9.25}$$

By the definition of $\hat{x}(t)$,

$$\sum_{j=1}^{N} v_j(\hat{x}(t), t_j) \geq \sum_{j=1}^{N} v_j(\tilde{x}^i(t_{-i}), t_j),$$

because $\tilde{x}^i(t_{-i}) \in X$. Rearranging this, we obtain,

$$\sum_{j=1}^{N} v_j(\hat{x}(t), t_j) - \sum_{j \neq i} v_j(\tilde{x}^i(t_{-i}), t_j) \geq v_i(\tilde{x}^i(t_{-i}), t_i).$$

Hence, by (9.24), i's utility from participating, exceeds (9.25), his utility from not participating. Thus, it is an equilibrium for all individuals to voluntarily participate in the VCG mechanism.

Third, the dominance of truth-telling in the VCG mechanism might appear to contradict Theorem 6.4 (the Gibbard-Satterthwaite Theorem) of Chapter 6. Indeed, the function $\hat{x}(\cdot)$ maps vectors of types (which index individual utility functions) into social choices in such a way that no individual can ever gain by reporting untruthfully. That is, $\hat{x}(\cdot)$ is strategy-proof. Moreover, because we have assumed nothing about the range of $\hat{x}(\cdot)$, the range might very well be all of X (if not, simply remove those elements of X that are absent from the range). In that case, $\hat{x}(\cdot)$ is a strategy-proof social choice function. But it is certainly not dictatorial! (Consider the one-good case, for example.) But, rest assured, there is no contradiction because, in contrast to the situation in Chapter 6, we have restricted the domain of preferences here to those that are quasi-linear. This restriction permits us to avoid the negative conclusion of the Gibbard-Satterthwaite theorem.

9.5.5 ACHIEVING A BALANCED BUDGET: EXPECTED EXTERNALITY MECHANISMS

As already noted, the VCG mechanism runs a surplus because each individual's cost is non-negative, and sometimes positive, regardless of the reported vector of types. But what happens to the revenue that is generated? Does it matter? In fact, it does.

Suppose, for example, that there are no other individuals in society but the N individuals participating in the VCG mechanism. Then any revenue that is generated must either be redistributed or destroyed.

If any amount of the revenue is destroyed, the overall outcome, which consists of the social state plus the amount of money each individual possesses is clearly *not* ex post Pareto efficient. Therefore, destroying any portion of the revenue is simply not an option. Hence, the only option consistent with our goal is to redistribute the revenue among the N individuals. But this causes problems as well.

If the revenue is redistributed to the N individuals, then the costs, $c_i^{VCG}(t)$, are no longer the correct costs. They instead *overstate* actual costs because they do not take into account the redistributed revenue. Consequently, it is not at all clear that, once individuals take into account their share of the revenue that is generated, it remains a dominant strategy to report their types truthfully. And of course, if individuals lie about their types, the social state chosen will typically not be ex post efficient. This is a potentially serious problem. Fortunately, because individual utilities are quasi-linear and individual types are independent, this problem can be solved so long as revenue is redistributed in a sufficiently careful manner.

Before getting to the solution, let us note that if the revenue generated is redistributed among the N individuals, then the sum of the actual (net) payments made by all of them must be zero. For example, if there are just two individuals, and one of them ultimately ends up paying a dollar, then the other individual must receive that dollar because there is simply nowhere else for it to go. Thus, what we are really looking for are mechanisms in which the costs always add up to zero. Such mechanisms are called **budget-balanced**.

DEFINITION 9.7 ***Budget-Balanced Cost Functions***

The cost functions, $c_1, \ldots, c_N$, are budget-balanced if they sum to zero regardless of the reported vector of types, i.e., if

$$\sum_{i=1}^{N} c_i(t) = 0, \quad \textit{for every } t \in T.$$

If a direct mechanism's cost functions are budget-balanced then we say that the mechanism is budget-balanced as well.

Thus, a budget-balanced mechanism not only wastes no money, it is completely self-sufficient, requiring no money from the outside. We will now adjust the VCG costs so that they result in a budget-balanced mechanism.

When the vector of reported types is $t \in T$ individual i's VCG cost, $c_i^{VCG}(t)$, is his externality. Thus, according to the formula in (9.22), the quantity

$$\begin{aligned}\bar{c}_i^{VCG}(t_i) &= \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i})c_i^{VCG}(t_i, t_{-i})\\ &= \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i})\left(\sum_{j\neq i} v_j(\tilde{x}^i(t_{-i}), t_j) - \sum_{j\neq i} v_j(\hat{x}(t), t_j)\right),\end{aligned} \tag{9.26}$$

is i's *expected externality when his type is t_i*. It turns out that these expected externalities can be used to define costs in a way that delivers ex post efficiency and a balanced budget.

THEOREM 9.11 ***The Budget-Balanced Expected Externality Mechanism***

Consider the mechanism in which, when the vector of reported types is $t \in T$, the ex post efficient social state $\hat{x}(t)$ is chosen and individual i's cost is,

$$\bar{c}_i^{VCG}(t_i) - \bar{c}_{i+1}^{VCG}(t_{i+1}),$$

where the $\bar{c}_i^{VCG}(t_i)$ are defined by (9.26), and $i+1=1$ when $i=N$. Then, this mechanism, which we call the budget-balanced expected externality mechanism, is incentive-compatible, ex post efficient, and budget-balanced. Furthermore, in the truth-telling equilibrium, every individual is voluntarily willing to participate regardless of his type.

The cost functions in the budget-balanced expected externality mechanism can be described as follows. Arrange the N individuals clockwise around a circular table. The mechanism requires each individual i to pay the person on his right his expected externality, $\bar{c}_i^{VCG}(t_i)$, given his reported type, t_i.[28] Mechanisms like this are sometimes called *expected externality mechanisms*.[29] Two points are worth emphasising. First, $\bar{c}_i^{VCG}(t_i) - \bar{c}_{i+1}^{VCG}(t_{i+1})$ is individual i's actual cost when the vector of reported types is $t \in T$. It is *not* his expected cost. Second, because individual i pays his expected externality to one person and receives another's expected externality, his actual cost is less than his expected externality. Hence his expected cost in the new mechanism is lower than in the original VCG mechanism.

Proof: The mechanism is clearly budget-balanced. (Do you see why?) Furthermore, if every individual always reports his type truthfully, then the ex post efficient social state $\hat{x}(t)$ will be chosen when the vector of types is $t \in T$. Hence it suffices to show that truthful reporting is a Bayesian-Nash equilibrium and that each individual is willing to participate.

[28] Paying one's externality to a single other individual keeps the formula for the new cost functions simple. But paying any number of the others one's expected externality would do just as well. See Exercise 9.29.
[29] See Arrow (1979) and d'Aspremont and Gérard-Varet (1979).

Let $u_i^{VCG}(r_i, t_i)$ denote individual i's expected utility in the VCG mechanism when his type is t_i and he reports that it is r_i, and when other individuals always report their types truthfully. Then,

$$u_i^{VCG}(r_i, t_i) = \sum_{t_{-i}\in T_{-i}} q(t_{-i})v_i(\hat{x}(r_i, t_{-i}), t_i) - \bar{c}_i^{VCG}(r_i),$$

because the first term (the summation) is his expected utility from the social state, and the second, with the negative sign, is his expected cost given his report. We already know that truth-telling is a Bayesian-Nash equilibrium in the VCG mechanism (indeed it is a dominant strategy). Hence, $u_i^{VCG}(r_i, t_i)$ is maximised in r_i when $r_i = t_i$.

In the new mechanism, i's expected costs when he reports r_i and the others (in particular individual $i+1$) report truthfully are,

$$\bar{c}_i^{VCG}(r_i) - \sum_{t_{i+1}\in T_{i+1}} q_{i+1}(t_{i+1})\bar{c}_{i+1}^{VCG}(t_{i+1}).$$

Hence, his expected utility when he reports r_i in the new mechanism and when all others report truthfully is,

$$\sum_{t_{-i}\in T_{-i}} q(t_{-i})v_i(\hat{x}(r_i, t_{-i}), t_i) - \bar{c}_i^{VCG}(r_i) + \bar{c}_{i+1},$$

where $\bar{c}_{i+1} = \sum_{t_{i+1}\in T_{i+1}} q_{i+1}(t_{i+1})\bar{c}_{i+1}^{VCG}(t_{i+1})$ is a constant. But this last expression is equal to,

$$u_i^{VCG}(r_i, t_i) + \bar{c}_{i+1},$$

and so it too is maximised in r_i when $r_i = t_i$. Hence, truth-telling is a Bayesian-Nash equilibrium in the new mechanism.

Furthermore, because $\bar{c}_{i+1}$ is always non-negative (because it is equal to individual $i+1$'s ex ante expected VCG cost), individual i is at least as well off in the new mechanism since he expects his costs to be weakly lower, regardless of his type. Hence, because individuals are willing to participate in the VCG mechanism regardless of their types, the same is true in the new mechanism. ∎

Note carefully that Theorem 9.11 does not say that truth-telling is a weakly dominant strategy in the new budget-balanced mechanism. It says only that truth-telling is a Bayesian-Nash equilibrium. Consequently, although we gain a balanced budget (and hence full efficiency) when we adjust the cost functions of the VCG mechanism, we lose the otherwise very nice property of dominant strategy equilibrium.[30]

[30] In fact, there are theorems stating that it is impossible to achieve both in a wide variety of circumstances. See Green and Laffont (1977) and Holmstrom (1979b).

EXAMPLE 9.5 Continuing with Examples 9.3 and 9.4, suppose that there are just two individuals, i.e., $N = 2$. As you are asked to show in Exercise 9.30, the cost formula given in Theorem 9.11 yields, for the two individuals here, budget-balanced cost functions that can be equivalently described by the following table.

If your reported type is:	1	2	3	4	5	6	7	8	9
You pay the other individual:	$\frac{10}{9}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{9}$	0	0	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{2}{3}$

Let us understand why the entries in the table are as they are. The 'circular seating' description following Theorem 9.11 implies that the entries in the second row of the table are simply the expected VCG costs, i.e., the $\bar{c}_i^{VCG}(t_i)$. In particular, the fourth entry in the second row is $\bar{c}_1^{VCG}(4)$, individual 1's expected VCG cost when he reports that his type is $t_1 = 4$. By reporting $t_1 = 4 < 5$, he can be pivotal only for the swimming pool, and even then he is pivotal only when individual 2 reports $t_2 = 6$, in which case his VCG cost, i.e., his externality, is $c_1^{VCG}(4, 6) = 6 - 5$ (see Example 9.4). Because individual 2 reports truthfully and the probability that $t_2 = 6$ is $1/9$, individual 1's expected externality is therefore $\bar{c}_1^{VCG}(4) = \frac{1}{9}(6 - 5) = \frac{1}{9}$, as in the table.

Note that one's payment to the other individual is higher the more extreme is one's report. This is in keeping with the idea that, for correct incentives, individuals should pay their externality (but keep in mind that the amount paid according to the table is *not* one's cost, because each individual also receives a payment from the other individual). Indeed, the more extreme an individual's report, the more likely it is that he gets his way, or, equivalently, the less likely it is that the other individual gets their way. Requiring individuals to pay more when their reports are more extreme keeps them honest.

Thus, when $N = 2$, the budget-balanced expected externality mechanism for the town is as follows. The two individuals are asked to report their types and make payments to one another according to the table above. The bridge is built if the sum of the reports exceeds 10 and the swimming pool is built otherwise. This mechanism is incentive-compatible, ex post efficient, budget-balanced, and leads to voluntary participation. □

Theorem 9.11 provides an affirmative answer to the question of whether one can design a mechanism that ensures an ex post efficient outcome in a quasi-linear utility, independent private values setting. Thus, we have come quite a long way. But there are important situations that our analysis so far does not cover and it is now time to get to them.

9.5.6 PROPERTY RIGHTS, OUTSIDE OPTIONS, AND INDIVIDUAL RATIONALITY CONSTRAINTS

Up to now we have implicitly assumed that, on the one hand, individuals cannot be forced to give up their income and, on the other hand, they have no property rights over social states. These implicit assumptions show up in our analysis when we consider whether individuals are willing to participate in the mechanism.[31] Indeed, we presumed that when an individual chooses not to participate, two things are true. First, his income is unchanged, implying that he cannot be forced to give it up. Second, the set of social states available to the remaining individuals is also unchanged, implying that the individual himself has no control – i.e., no property rights – over them.

The 'no property rights over social states' assumption sometimes makes perfect sense. For example, when the mechanism is an auction and the N participating individuals are bidders, it is natural to suppose that no bidder has any effect on the availability of the good should he decide not to participate. But what if we include the seller as one of the individuals participating in the mechanism? It typically will not make sense to assume that the good will remain available to the bidders if the seller chooses not to participate.[32] Or, consider a situation in which a firm-owner has the technology to produce a good (at some cost) that a consumer might value. In this case too, the set of social states is not the same for the consumer alone as it is with the consumer and firm-owner together. Or, suppose one is interested in dissolving a partnership (e.g., a law firm, a marriage, etc.) efficiently, where each partner has rights to the property that is jointly owned. In order to cover these and other important situations we must generalise our model.

The key to accommodating property rights over social states is to be more flexible about individual participation decisions. To get us moving in the right direction, consider a situation involving a seller who owns an object and potential buyer. The seller's value for the object is some $v_s \in [0, 1]$ known only to the seller, and the buyer's value for the object is some $v_b \in [0, 1]$ known only to the buyer. If we wish to give the seller property rights over the object, then we cannot force him to trade it away. Consequently, the seller will participate in a mechanism only if he expects to receive utility at least v_s from doing so, because he can achieve this utility by not participating and keeping the object for himself. The notable feature of this example is that the value to the seller of not participating depends non-trivially on his private type, v_s. We will now incorporate this idea into our general model.

For each individual i, and for each $t_i \in T_i$, let $IR_i(t_i)$ denote i's expected utility when he does not participate in the mechanism and his type is t_i. Thus, in the example of the previous paragraph, letting individual 1 be the seller, we have $IR_1(v_s) = v_s$ for each $v_s \in [0, 1]$, and letting individual 2 be the buyer, we have $IR_2(v_b) = 0$ for each $v_b \in [0, 1]$.

DEFINITION 9.8 ***Individual Rationality***

An incentive-compatible direct mechanism is individually rational if for each individual i and for each $t_i \in T_i$, i's expected utility from participating in the mechanism and reporting truthfully when his type is t_i is at least $IR_i(t_i)$ when the others participate and always report their types truthfully.

---

[31] Participation in the mechanism, as always, implies a commitment to abide by its outcome.

[32] Note that in our treatment of auctions, the seller is always better off participating in the auction than not. Hence, this issue also arises there but is well taken care of.

Equivalently, an incentive-compatible direct mechanism, $p, c_1, \ldots, c_N$, is individually rational if the following individual rationality constraints are satisfied:

$$\sum_{x\in X} \bar{p}_i^x(t_i)v_i(x, t_i) - \bar{c}_i(t_i) \geq IR_i(t_i), \text{ for every } i \text{ and every } t_i \in T_i.$$

Thus, if an incentive-compatible direct mechanism is individually rational, it is optimal for each individual to voluntarily participate in the mechanism regardless of his type because his expected utility is at least as high when he participates as when he does not.

The individual rationality constraints appearing in Definition 9.8 are additional constraints, above and beyond the constraints imposed by incentive-compatibility and any other constraints of interest, such as ex post efficiency. The higher are the values $IR_i(t_i)$, the more difficult it will be to construct an incentive-compatible ex post efficient mechanism. Because property rights over social states often increase the $IR_i(t_i)$, their presence can create difficulties.

Note that we can always return to the model of no property rights over social states by defining $IR_i(t_i) = \min_{x\in X} v_i(x, t_i)$ since, given our harmless convention of zero initial income for each individual, this is the least utility one can expect when one's type is t_i.

Finally, note that introducing the functions, $IR_i(t_i)$, has an added benefit. They permit us to model the possibility that individuals have 'outside options', even if they have no property rights over the social states *per se*. For example, suppose that an individual has the opportunity to participate in one of several mechanisms and has no property rights over social states in any of them – e.g. think of a bidder considering which one of several auctions to attend. If you are designing one of the mechanisms, and $U_i^k(t_i)$ is individual i's expected utility from participating in any one of the other mechanisms $k = 1, \ldots, K$ when his type is t_i, then i will voluntarily participate in your mechanism only if his expected utility from doing so is at least $\max_k U_i^k(t_i)$. Consequently, to correctly assess i's participation decision in your mechanism we would define $IR_i(t_i) = \max_k U_i^k(t_i)$.

Let us take another look at Example 9.3 now that we can include property rights.

EXAMPLE 9.6 Reconsider Example 9.3 but suppose that the town itself must finance the building of either the bridge or the swimming pool and that building neither (i.e., 'Don't Build' (D)) is a third social state that is available. The types are as before as are the utilities for the bridge and pool. But we must specify utilities for building nothing. Suppose that individual 1 is the only engineer in town and that he would be the one to build the bridge or the pool. His utility for the social state D is

$$v_1(D, t_1) = 10,$$

while for every other individual $i > 1$,

$$v_i(D, t_i) = 0.$$

You may think of $v_1(D, t_1) = 10$ as the engineer's (opportunity) cost of building either the bridge or the pool. So, if the engineer cannot be forced to build (i.e., if he has property rights over the social state D), then the mechanism must give him at least an expected utility of 10 because he can ensure this utility simply by not building anything. Hence, for every $t \in T$, we have $IR_1(t_1) = 10$ and $IR_i(t_i) = 0$ for $i > 1$. As we now show, the expected externality mechanism that worked so beautifully without property rights no longer works.

Note that it is always efficient to build something, because total utility is equal to 10 if nothing is built, while it is strictly greater than 10 (assuming the engineer is not the only individual) if the swimming pool is built. Suppose there are just two individuals, the engineer and one other. The expected externality mechanism described in Example 9.5 fails to work because the engineer will sometimes refuse to build. For instance, if the engineer's type is $t_1 < 4$, then whatever are the reports, the mechanism will indicate that either the bridge or the pool be built and individual 2's payment to the engineer will be no more than $10/9$. Consequently, even ignoring the payment the engineer must make to individual 2, the engineer's expected utility if he builds is strictly less than 10 because $\max(t_1 + 5, 2t_1) + 10/9 < 10$ when $t_1 < 4$. The engineer is therefore strictly better off exercising his right not to build. So, under the expected externality mechanism, the outcome is inefficient whenever $t_1 < 4$ because the engineer's individual rationality constraint is violated. □

Can the type of difficulty encountered in Example 9.6 be remedied? That is, is it always possible to design an incentive-compatible, ex post efficient, budget-balanced direct mechanism that is also individually rational? In general, the answer is 'No' (we will return to the specific case of Example 9.6 a little later). However, we can come to an essentially complete understanding of when it is possible and when it is not. Let us begin by providing conditions under which it is possible.[33]

this is why we study IR-VCG mechanisms

9.5.7 THE IR-VCG MECHANISM: SUFFICIENCY OF EXPECTED SURPLUS

For each individual i, let $U_i^{VCG}(t_i)$ be his expected utility when his type is t_i in the truth-telling (dominant-strategy) equilibrium of the VCG mechanism. Hence,

$$U_i^{VCG}(t_i) = \sum_{t_{-i}\in T_{-i}} \sum_{x\in X} q_{-i}(t_{-i})v_i(\hat{x}(t_i, t_{-i}), t_i) - \bar{c}_i^{VCG}(t_i).$$

As we now know, this mechanism need be neither budget-balanced nor individually rational.

Let us first try and achieve individual rationality in the simplest possible way, namely by giving a fixed amount of money to each individual so that they are willing to participate in the VCG mechanism no matter what their type. Let ψ_i denote the *participation subsidy* given to individual i. When will it be large enough so that he always chooses to participate in the VCG mechanism? The answer, of course, is that it must be such that,

$$U_i^{VCG}(t_i) + \psi_i \geq IR_i(t_i), \text{ for every } t_i \in T_i,$$

[33]The remainder of this chapter draws heavily from Krishna and Perry (1998). Another very nice treatment can be found in Williams (1999).

or equivalently, such that

$$\psi_i \geq IR_i(t_i) - U_i^{VCG}(t_i), \text{ for every } t_i \in T_i,$$

or finally, such that

$$\psi_i \geq \max_{t_i \in T_i}(IR_i(t_i) - U_i^{VCG}(t_i)).$$

Consequently, the *minimum* participation subsidy we must give to individual i (and it may be negative) so that he is willing to participate in the VCG mechanism regardless of his type, is

$$\psi_i^* = \max_{t_i \in T_i}(IR_i(t_i) - U_i^{VCG}(t_i)). \tag{9.27}$$

Using these minimum participation subsidies, we now define a new mechanism, called the *individually rational VCG mechanism*, or simply the IR-VCG mechanism. In the IR-VCG mechanism, each individual i reports his type and is given ψ_i^* dollars no matter what type he reports. If the vector of reports is t, the social state is $\hat{x}(t)$ and individual i must in addition pay his Vickrey cost, $c_i^{VCG}(t)$. Consequently, in total, individual i's cost is $c_i^{VCG}(t) - \psi_i^*$.

Because the participation subsidies, ψ_i^*, are handed out regardless of the reports, they have no effect on one's incentives to lie. Hence it remains a dominant strategy to report one's type truthfully. Moreover, by construction, the IR-VCG mechanism is individually rational and ex post efficient. Hence, the IR-VCG mechanism is incentive-compatible, ex post efficient, and individually rational. The only problem is that it might not be budget-balanced. To balance the budget, we might try the same trick used in Theorem 9.11, namely to seat the individuals around a circular table and have each one pay the person on their right their expected cost given their report. But there is a problem. Because the VCG costs have been reduced by the participation subsidies, it might be that an individual's expected cost is now negative. He would then not be *paying* the individual on his right. Instead, he would be *taking money away from* the individual on his right (who is also paying his expected cost to the individual on his right). This additional expense for him, the individual on his right, might lead to a violation of his individual rationality constraint. If so, our 'circular-table' trick balances the budget, but it results in a mechanism that is no longer individually rational. Thus, balancing the budget when expected costs are negative for some individuals requires a more sophisticated method than that described in Theorem 9.11, if indeed balancing the budget is possible at all.

Say that an incentive-compatible direct mechanism with cost functions $c_1, \ldots, c_N$, *runs an expected surplus* if, in the truth-telling equilibrium, ex ante expected revenue is non-negative, i.e., if

$$\sum_{t \in T} q(t) \left(\sum_{i=1}^{N} c_i(t) \right) \geq 0.$$

Note that the VCG mechanism runs an expected surplus because $c_i^{VCG}(t) \geq 0$ for every i and every t. On the other hand, the IR-VCG may or may not run an expected surplus because it reduces the expected surplus of the VCG mechanism by the amount of the participation subsidies. We can now state the following result, which holds for *any* incentive-compatible mechanism, not merely for the particular mechanisms we have considered so far.

THEOREM 9.12 ***Achieving a Balanced Budget***

Suppose that an incentive-compatible direct mechanism with cost functions $c_1, \ldots, c_N$ runs an expected surplus. For each individual i and every $t \in T$ replace the cost function c_i by the cost function,

$$c_i^B(t) = \bar{c}_i(t_i) - \bar{c}_{i+1}(t_{i+1}) + \bar{c}_{i+1} - \frac{1}{N}\sum_{j=1}^{N}\bar{c}_j,$$

where $\bar{c}_i(t_i)$ is defined by (9.22), $\bar{c}_i = \sum_{t\in T} q(t)c_i(t)$, and $i+1=1$ when $i=N$. Then the resulting mechanism – with the same probability assignment function – is budget-balanced and remains incentive-compatible. Moreover, the resulting mechanism is weakly preferred by every type of every individual to the original mechanism. Therefore, if the original mechanism was individually rational, so is the new budget-balanced mechanism.

Analogous to the budget-balanced expected externality mechanism, the cost functions, c_i^B, defined in Theorem 9.12 can be described rather simply. Seat the N individuals in order, from 1 to N, clockwise around a circular table. If individual i reports t_i, he pays every other individual the fixed amount $\bar{c}_{i+1}/N$ and pays the individual on his right the additional amount $\bar{c}_i(t_i)$. That's it!

If you carefully trace through who pays whom how much when the vector of reports is t, you will find that, in the end, individual i's net cost is $c_i^B(t)$. Once again, the nice thing about this way of looking at the cost functions c_i^B is that it is 'obvious' that they balance the budget. Why? Because the N individuals are simply making payments among themselves. Hence, no money leaves the system (so no revenue is generated) and no money is pumped into the system (so no losses are generated). Therefore, the budget must balance. Let us now prove Theorem 9.12.

Proof: As we have already noted, the cost functions, $c_1^B, \ldots, c_N^B$, are budget-balanced. (For a more direct proof, add them up!)

Second, given the cost function c_i^B, individual i's expected cost when he reports r_i and the others report truthfully is,

$$\bar{c}_i^B(r_i) = \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i})c_i^B(r_i, t_{-i}).$$

Substituting for the definition of $c_i^B(r_i, t_{-i})$ gives,

$$
\begin{aligned}
\bar{c}_i^B(r_i) &= \sum_{t_{-i}\in T_{-i}} q_{-i}(t_{-i}) \left(\bar{c}_i(t_i) - \bar{c}_{i+1}(t_{i+1}) + \bar{c}_{i+1} - \frac{1}{N}\sum_{j=1}^{N} \bar{c}_j \right) \\
&= \bar{c}_i(r_i) - \left(\sum_{t_{i+1}\in T_{i+1}} q_{i+1}(t_{i+1})\bar{c}_{i+1}(t_{i+1}) \right) + \bar{c}_{i+1} - \frac{1}{N}\sum_{j=1}^{N} \bar{c}_j \\
&= \bar{c}_i(r_i) - \bar{c}_{i+1} + \bar{c}_{i+1} - \frac{1}{N}\sum_{j=1}^{N} \bar{c}_j \\
&= \bar{c}_i(r_i) - \frac{1}{N}\sum_{j=1}^{N} \bar{c}_j. \qquad \text{(P.1)}
\end{aligned}
$$

Hence, individual i's expected cost when he reports r_i differs from the original by a fixed constant.

Given the original cost functions, c_i, let $u_i(r_i, t_i)$ denote individual i's expected utility from reporting r_i when his type is t_i and when all others report truthfully, and let $u_i^B(r_i, t_i)$ denote the analogous quantity with the new cost functions, c_i^B. Because the probability assignment function has not changed, the result from (P.1) and the formula in (9.23) imply,

$$
u_i^B(r_i, t_i) = u_i(r_i, t_i) + \frac{1}{N}\sum_{j=1}^{N} \bar{c}_j. \qquad \text{(P.2)}
$$

Therefore, because $u_i(r_i, t_i)$ is maximised in r_i when $r_i = t_i$, the same is true of $u_i^B(r_i, t_i)$ and we conclude that the new mechanism is incentive-compatible.

Finally, the assumption that the original mechanism runs an expected surplus means precisely that,

$$
\sum_{j=1}^{N} \bar{c}_j \geq 0.
$$

Consequently, evaluating (P.2) at $r_i = t_i$, the expected utility of every type of each individual is at least as high in the truth-telling equilibrium with the new cost functions as with the old. ∎

Let us note a few things about Theorem 9.12. First it provides explicit budget-balanced cost functions derived from the original cost functions that maintain incentive-compatibility. Second, not only do we achieve a balanced budget, we do so while ensuring that every individual, regardless of his type, is at least as well off in the truth-telling equilibrium of the new mechanism as he was in the old. Thus, if individuals were willing to participate in the old mechanism they are willing to participate in the new mechanism as well, regardless of their type.[34] Consequently, an immediate implication of Theorem 9.12 is the following.

THEOREM 9.13 ***IR-VCG Expected Surplus: Sufficiency***

Suppose that the IR-VCG mechanism runs an expected surplus, i.e., suppose that,

$$\sum_{t \in T} \sum_{i=1}^{N} q(t)(c_i^{VCG}(t) - \psi_i^*) \geq 0.$$

Then, the following direct mechanism is incentive-compatible, ex post efficient, budget-balanced, and individually rational: Each individual reports his type. If the reported vector of types is $t \in T$, then the social state is $\hat{x}(t)$, and individual i pays the cost,

$$\bar{c}_i^{VCG}(t_i) - \psi_i^* - \bar{c}_{i+1}^{VCG}(t_{i+1}) + \bar{c}_{i+1}^{VCG} - \frac{1}{N}\sum_{j=1}^{N}(\bar{c}_j^{VCG} - \psi_j^*),$$

where $\bar{c}_j^{VCG}(t_j)$ is defined by (9.26) and $\bar{c}_j^{VCG} = \sum_{t_j \in T_j} q_j(t_j)\bar{c}_j^{VCG}(t_j)$ is individual j's ex ante expected VCG cost.

The proof of Theorem 9.13 really is immediate because the IR-VCG mechanism is incentive-compatible, ex post efficient, and individually rational. So, if it runs an expected surplus, adjusting its cost functions, $c_i^{VCG}(t) - \psi_i^*$, according to Theorem 9.12, results in an incentive-compatible, ex post efficient, budget-balanced, and individually rational mechanism. You now need only convince yourself that the resulting mechanism is precisely that which is defined in Theorem 9.13. (Do convince yourself.)

Theorem 9.13 identifies expected surplus in the IR-VCG mechanism as a *sufficient* condition for the existence of a mechanism that satisfies all of our demands, i.e., incentive compatibility, ex post efficiency, budget-balancedness, and individual rationality. Moreover, Theorem 9.13 explicitly constructs such a mechanism.

EXAMPLE 9.7 In the light of Theorem 9.13, let us reconsider Example 9.6 when there are just two individuals, one of whom is the engineer. We know from Example 9.6 that the budget-balanced expected externality mechanism is not individually rational. In particular, when the engineer's type is low, he is better off not participating. Thus, the engineer's

[34]Theorem 9.12 remains true, and the proof is identical to that given here, even when the private value assumption fails – i.e., even when $v_i(x, t)$ depends on t_{-i} as well as on t_i. On the other hand, the proof given here depends crucially on the assumption that the types are independent across individuals.

participation subsidy ψ_1^* must be strictly positive. Let us check whether the IR-VCG mechanism runs an expected surplus here. According to (9.27),

$$\psi_1^* = \max_{t_1 \in T_1}(IR_1(t_1) - U_1^{VCG}(t_1)).$$

Because $IR_1(t_1) = 10$ for all $t_1 \in T_1$, we have,

$$\begin{aligned}\psi_1^* &= \max_{t_1 \in T_1}(10 - U_1^{VCG}(t_1)) \\ &= 10 - \min_{t_1 \in T_1} U_1^{VCG}(t_1).\end{aligned}$$

Thus, we must compute the minimum expected VCG utility over the engineer's types. It is not difficult to argue that the higher is the engineer's type, the better off he must be in the VCG mechanism (see Exercise 9.32). Hence, the minimum occurs when $t_1 = 1$, and his expected utility in the VCG mechanism is,

$$U_1^{VCG}(1) = (1 + 5) - \frac{10}{9},$$

because the pool will be built regardless of 2's report, and his expected *VCG* cost when his type is $t_1 = 1$ is, $\bar{c}_1^{VCG}(1) = 10/9$ (see Exercise 9.33). Hence,

$$\begin{aligned}\psi_1^* &= 10 - U_1^{VCG}(1) \\ &= 10 - 6 + \frac{10}{9} \\ &= \frac{46}{9}.\end{aligned}$$

Similarly, because $IR_2(t_2) = 0$ for all $t_2 \in T_2$ and because $\bar{c}_2^{VCG}(1) = 20/9$ (see Exercise 9.33),

$$\begin{aligned}\psi_2^* &= 0 - U_2^{VCG}(1) \\ &= 0 - ((1 + 5) - \frac{20}{9}) \\ &= -\frac{34}{9},\end{aligned}$$

so that,

$$\psi_1^* + \psi_2^* = \frac{4}{3}.$$

As you are asked to show in Exercise 9.33, the VCG's ex ante expected revenue here is $50/27 > 4/3$. Consequently, the IR-VCG mechanism runs an expected surplus and therefore it is possible to ensure that the outcome is ex post efficient even while respecting the engineer's individual rationality constraint. Exercise 9.33 asks you to explicitly provide a mechanism that does the job. □

9.5.8 THE NECESSITY OF IR-VCG EXPECTED SURPLUS

Up to this point in our analysis of the quasi-linear model we have assumed that the type set of each individual is finite. This was for simplicity only. Everything we have done up to now goes through with essentially no changes even when the type sets are infinite, e.g., intervals of real numbers ($T_i = [0, 1]$) or products of intervals of real numbers ($T_i = [0, 1]^K$).

But we now wish to show that an expected surplus in the IR-VCG mechanism is not only a sufficient condition but that it is also a *necessary condition* for the existence of an incentive-compatible, ex post efficient, budget-balanced, individually rational mechanism. And in order to do this, we must abandon finite type spaces. So, we will make the simplest possible assumption that suits our purposes. We will assume that each individuals' type space, T_i, is the unit interval $[0, 1]$.[35] We will continue to assume that the set of social states, X, is finite.

Our objective is modest. We wish to give the reader a good sense of why an expected surplus in the IR-VCG mechanism is also a necessary condition for ex post efficient mechanism design when type sets are intervals without always taking care of the finer technical details that arise with infinite type spaces.[36] Fortuitously, the notation largely remains as it has been except that sums over types are now integrals, and the probabilities, $q_i(t_i)$, are now probability density functions. So, for example, given a direct mechanism, $p, c_1, \ldots, c_N$, we now have for each individual i and type $t_i \in [0, 1]$,

$$\bar{p}_i^x(t_i) = \int_{T_{-i}} p_i^x(t_i, t_{-i})q_{-i}(t_{-i})dt_{-i}, \text{ for each } x \in X,$$

and

$$\bar{c}_i(t_i) = \int_{T_{-i}} c_i(t_i, t_{-i})q_{-i}(t_{-i})dt_{-i},$$

rather than their finite summation counterparts in (9.22). Consequently, individual i's expected utility from reporting r_i when his type is t_i and the others report truthfully is once again,

$$u_i(r_i, t_i) = \sum_{x \in X} \bar{p}_i^x(r_i)v_i(x, t_i) - \bar{c}_i(r_i),$$

exactly as in (9.23) where type sets are finite.

Now, suppose that the direct mechanism, $p, c_1, \ldots, c_N$, is incentive compatible. This implies that $u_i(r_i, t_i)$ is maximised in r_i at $r_i = t_i$. Assuming differentiability wherever we

[35] One can just as well allow type spaces to be Euclidean cubes. But we shall not do so here.

[36] E.g., measurability, continuity, or differentiability of probability assignment functions or cost functions.

need it, this yields the following first-order condition for every individual i and every $t_i \in (0, 1)$.

$$\left.\frac{\partial u_i(r_i, t_i)}{\partial r_i}\right|_{r_i=t_i} = \sum_{x\in X} \bar{p}_i^{x\prime}(t_i)\, v_i(x, t_i) - \bar{c}_i'(t_i) = 0,$$

so that

$$\bar{c}_i'(t_i) = \sum_{x\in X} \bar{p}_i^{x\prime}(t_i)\, v_i(x, t_i). \tag{9.28}$$

Consequently, if two mechanisms, $p, c_{A1}, \ldots, c_{AN}$, and $p, c_{B1}, \ldots, c_{BN}$ have the same probability assignment function, then the derivative of the expected costs in the A mechanism, $\bar{c}_{Ai}'(t_i)$, must satisfy (9.28) as must the derivative of the expected costs in the B mechanism, $\bar{c}_{iB}'(t_i)$. Hence, for all i and all $t_i \in (0, 1)$,

$$\bar{c}_{Ai}'(t_i) = \sum_{x\in X} \bar{p}_i^{x\prime}(t_i)\, v_i(x, t_i) = \bar{c}_{Bi}'(t_i).$$

That is, the derivatives of the expected cost functions must be identical. But then, so long as the fundamental theorem of calculus can be applied, the expected cost functions themselves must differ by a constant because,

$$\bar{c}_{Ai}(t_i) - \bar{c}_{Ai}(0) = \int_0^{t_i} \bar{c}_{Ai}'(s)\,ds = \int_0^{t_i} \bar{c}_{Bi}'(s)\,ds = \bar{c}_{Bi}(t_i) - \bar{c}_{Bi}(0).$$

In obtaining this conclusion we assumed differentiability of the mechanism and also that the derivative of expected cost is sufficiently well-behaved so that the fundamental theorem of calculus applies. These assumptions are in fact unnecessary for the result so long as, for example, $\partial v_i(x, t_i)/\partial t_i$ exists and is continuous in $t_i \in [0, 1]$ for each $x \in X$. Moreover, note that because (9.28) depends only on the *expected probabilities*, $\bar{p}_i^x$, it is enough that the two mechanisms have the same expected probability assignment functions. We state the following result without proof.

THEOREM 9.14 ***Costs Differ by a Constant***

Suppose that for each individual i, $\partial v_i(x, t_i)/\partial t_i$ exists and is continuous in $t_i \in [0, 1]$ for each $x \in X$. If two incentive-compatible mechanisms have the same expected probability assignment functions, $\bar{p}_i^x$, then for each i, individual i's expected cost functions in the two mechanisms differ by a constant (which may depend upon i).

From this we immediately obtain a general revenue equivalence result that is worthwhile stating in passing. It generalises Theorem 9.6.

THEOREM 9.15 ***A General Revenue Equivalence Theorem***

Suppose that for each individual i, $\partial v_i(x, t_i)/\partial t_i$ exists and is continuous in $t_i \in [0, 1]$ for each $x \in X$. If two incentive-compatible mechanisms have the same expected probability *assignment functions, $\bar{p}_i^x$, and every individual is indifferent between the two mechanisms when his type is zero, then the two mechanisms generate the same expected revenue.*

We leave the proof of Theorem 9.15 to you as an exercise (see Exercise 9.35). Another immediate consequence of Theorem 9.14 is the following. Suppose that for each $t \in T$ there is a unique ex post efficient social state. Then any two ex post efficient incentive-compatible mechanisms have the same probability assignment functions. Hence, by Theorem 9.14, because the VCG mechanism is incentive-compatible and ex post efficient, any other incentive compatible ex post efficient mechanism must have expected cost functions that differ from the VCG expected costs (i.e., the expected externalities) by a constant. Indeed, if you look back at all of the ex post efficient mechanisms we constructed, expected costs differ by a constant from the VCG expected costs. This fact is the basis of our next result.

The assumption that for each $t \in T$ there is a unique ex post efficient social state is very strong when there are finitely many social states.[37] Fortunately there are much weaker assumptions that have the same effect.

THEOREM 9.16 ***Maximal Revenue Subject to Efficiency and Individual Rationality***

Suppose that for each individual i, $\partial v_i(x, t_i)/\partial t_i$ exists and is continuous in $t_i \in [0, 1]$ for each $x \in X$. In addition, suppose that for each individual i and for each $t_{-i} \in T_{-i}$, there is, for all but perhaps finitely many $t_i \in T_i$, a unique ex post efficient social state given (t_i, t_{-i}). Then the IR-VCG mechanism generates the maximum ex ante expected revenue among all incentive-compatible, ex post efficient, individually rational direct mechanisms.

Proof: Because, for each $t_{-j} \in T_{-j}$ there is, for all but finitely many $t_j \in T_j$, a unique ex post efficient social state given (t_j, t_{-j}), the expected probability assignment functions, $\bar{p}_i^x(t_i)$, are uniquely determined by ex post efficiency. Consider then some incentive-compatible, ex post efficient, individually rational direct mechanism with cost functions $c_1, \ldots, c_N$. According to the fact just stated, its expected probability assignment functions must coincide with those of the IR-VCG mechanism. So, by Theorem 9.14, its expected cost functions differ by a constant from the expected cost functions, $\bar{c}_i^{VCG}(t_i) - \psi_i^*$, of the IR-VCG mechanism.[38] Hence, for some constants, $k_1, \ldots, k_N$,

$$\bar{c}_i(t_i) = \bar{c}_i^{VCG}(t_i) - \psi_i^* - k_i, \tag{P.1}$$

for every individual i and every $t_i \in T_i$.

Now, because the expected probability assignment functions in the two mechanisms are the same and because the mechanism with cost functions, c_i, is individually rational,

[37] Indeed, if each $v_i(x, t_i)$ is continuous in t_i, then uniqueness implies that $\hat{x}(t)$ must be constant. That is, there must be a single social state that is ex post efficient regardless of the vector of types. But in that case there is no problem to begin with since there is no uncertainty about which social state is ex post efficient.

[38] Keep in mind that all of our formulae, including those defing the ψ_i^*, must be adjusted from sums over types to integrals over types. But otherwise they are the same.

(P.1) says that starting with the VCG mechanism and adjusting it by giving to each individual i the participation subsidy $\psi_i^* + k_i$ renders it individually rational in addition to ex post efficient. But because the participation subsidies ψ_i^* are, by definition, the *smallest* such subsidies, it must be the case that $k_i \geq 0$ for all i. Hence, by (P.1),

$$\bar{c}_i(t_i) \leq \bar{c}_i^{VCG}(t_i) - \psi_i^*,$$

so that each individual, regardless of his type, expects to pay a lower cost in the mechanism with cost functions, c_i, than in the IR-VCG mechanism. Hence, the IR-VCG mechanism generates at least as much expected revenue. ∎

We can now prove the result we have been heading towards.

THEOREM 9.17 ***IR-VCG Expected Surplus: Necessity***

Suppose that for each individual i, $\partial v_i(x, t_i)/\partial t_i$ exists and is continuous in $t_i \in [0, 1]$ for each $x \in X$. In addition, suppose that for each individual i and for each $t_{-i} \in T_{-i}$, there is, for all but perhaps finitely many $t_i \in T_i$, a unique ex post efficient social state given (t_i, t_{-i}). If there exists an incentive-compatible, ex post efficient, budget-balanced, individually rational direct mechanism, then the IR-VCG mechanism runs an expected surplus.

Proof: If such a mechanism exists, then, because it is budget-balanced, its expected revenues are zero. On the other hand, by Theorem 9.16, the IR-VCG mechanism raises at least as much ex ante expected revenue. Therefore, the ex ante expected revenue raised by the IR-VCG mechanism must be non-negative. ∎

We can apply Theorem 9.17 to obtain a special case of an important impossibility result due to Myerson and Satterthwaite (1983).

EXAMPLE 9.8 Consider a buyer and a seller. The seller owns an indivisible object and both the buyer and seller have quasi-linear preferences and private values for the object. There are two social states, B (the buyer receives the object) and S (the seller receives the object). It is convenient to index the two individuals by $i = b$ for the buyer and $i = s$ for the seller. The buyer's type, t_b, and the seller's type, t_s, are each drawn uniformly and independently from $[0, 1]$. For each individual i, t_i is his value for the object. So, for example, $v_b(B, t_b) = t_b$ and $v_b(S, t_b) = 0$, and similarly for the seller. The seller has property rights over the object and so $IR_s(t_s) = t_s$ because the seller can always choose not to participate in the mechanism and keep his object. On the other hand, $IR_b(t_b) = 0$ because non-participation leaves the buyer with zero utility.

We wish to know whether there exists a direct mechanism that is incentive-compatible, ex post efficient, budget-balanced, and individually rational. According to Theorems 9.13 and 9.17 the answer is 'yes' if and only if the IR-VCG mechanism runs an expected surplus. So, let's check.

First, we compute the VCG cost functions. If the buyer's type is t_b, then it is efficient for the object to go to the buyer when $t_b > t_s$. In this case, the buyer's externality is $t_s - 0$

because, without the buyer, the seller receives the object and obtains utility t_s from the social state, but with the buyer, the buyer receives the object and the seller obtains utility zero from the social state. On the other hand, if $t_b < t_s$, the buyer's externality is zero because with or without him the seller receives the object. Hence,

$$c_b^{VCG}(t_b, t_s) = \begin{cases} t_s, & \text{if } t_b > t_s \\ 0 & \text{if } t_b < t_s. \end{cases}$$

There is no need to specify who receives the object when $t_b = t_s$ because this event occurs with probability zero and will have no effect on expected costs. Indeed, from what we already know, the buyer's expected cost given his type is:

$$\bar{c}_b^{VCG}(t_b) = \int_0^{t_b} t_s dt_s = \frac{1}{2} t_b^2.$$

Similarly, because within the VCG mechanism the buyer and seller are symmetric,[39]

$$\bar{c}_s^{VCG}(t_s) = \frac{1}{2} t_s^2.$$

Consequently, if $U_i^{VCG}(t_i)$ is individual i's expected utility in the VCG mechanism when his type is t_i, then

$$\begin{aligned} U_b^{VCG}(t_b) &= \int_0^{t_b} t_b dt_s - \bar{c}_b^{VCG}(t_b) \\ &= t_b^2 - \frac{1}{2} t_b^2 \\ &= \frac{1}{2} t_b^2, \end{aligned}$$

where the integral in the first line is the utility the buyer expects from receiving the object when that is the efficient social state. Similarly, by symmetry,

$$U_s^{VCG}(t_s) = \frac{1}{2} t_s^2.$$

The IR-VCG mechanism runs an expected surplus when the expected revenue from the VCG mechanism exceeds the sum of the participation subsidies, $\psi_b^* + \psi_s^*$. The

[39]The fact that the seller owns the object plays no role in the VCG mechanism, which always operates as if there are no property rights over social states.

expected revenue from the VCG mechanism is,

$$\begin{aligned}\int_0^1 \bar{c}_b^{VCG}(t_b)\,dt_b + \int_0^1 \bar{c}_s^{VCG}(t_s)\,dt_s &= \int_0^1 \frac{1}{2}t_b^2 dt_b + \int_0^1 \frac{1}{2}t_s^2 dt_s \\ &= \frac{1}{6} + \frac{1}{6} \\ &= \frac{1}{3}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\psi_b^* &= \max_{t_b\in[0,1]} (IR_b(t_b) - U_b^{VCG}(t_b)) \\ &= \max_{t_b\in[0,1]} (0 - \frac{1}{2}t_b^2) \\ &= 0,\end{aligned}$$

and

$$\begin{aligned}\psi_s^* &= \max_{t_s\in[0,1]} (IR_s(t_b) - U_s^{VCG}(t_s)) \\ &= \max_{t_s\in[0,1]} (t_s - \frac{1}{2}t_s^2) \\ &= \frac{1}{2},\end{aligned}$$

so that

$$\psi_b^* + \psi_s^* = \frac{1}{2} > \frac{1}{3}.$$

Thus, we conclude that there *does not* exist an incentive-compatible, ex post efficient, budget-balanced, individually rational mechanism in this situation. □

There are several lessons to draw from Example 9.8. First, the example provides an explanation for the otherwise puzzling phenomenon of strikes and disagreements in bargaining situations. The puzzling thing about strikes is that one imagines that whatever agreement is eventually reached could have been reached without the strike, saving both sides time and resources. But the result in the example demonstrates that this 'intuition' is simply wrong. Sometimes there is no mechanism that can assure ex post efficiency – inefficiencies must occasionally appear. And one example of such an inefficiency is that associated with a strike.

Second, the example illustrates that property rights matter. A very famous result in law and economics is the 'Coase Theorem' which states, roughly, that if one's only interest is Pareto efficiency, property rights do not matter – e.g., whether a downstream fishery is given the legal right to clean water or an upstream steel mill is given the legal right to dump waste into the stream, the two parties will, through appropriate transfer payments to one another, reach a Pareto-efficient agreement. Our analysis reveals an important caveat, namely that the Coase Theorem can fail when the parties have private information about their preferences. If no individual has property rights over social states, we found that efficiency was always possible. However, when property rights are assigned (as in the buyer–seller example) an efficient agreement cannot always be guaranteed.

Third, the fact that property rights can get in the way of efficiency provides an important lesson for the privatisation of public assets (e.g. government sale of off-shore oil rights, or of radio spectrum for commercial communication (mobile phones, television, radio)). If the government's objective is efficiency, then it is important to design the privatisation mechanism so that it assigns the objects efficiently, if possible. This is because the assignment, by its nature, creates property rights. If the assignment is inefficient, and private information remains, the establishment of property rights may well lead to unavoidable, persistent, and potentially large efficiency losses.

Fourth, the example suggests that the lack of symmetry in ownership may play a role in the impossibility result. For example, a setting without property rights is one where property rights are symmetric and also one where it is possible to construct an ex post efficient budget-balanced mechanism with voluntary participation. In the exercises you are asked to explore this idea further (see also Cramton et al. (1987)).

An excellent question at this point is, 'What do we do when there does not exist an incentive-compatible, ex post efficient, budget-balanced, individually rational mechanism?' This is a terrific and important question, but one we will not pursue in this introduction to mechanism design. One answer, however, is that we do the next best thing. We instead search among all incentive-compatible mechanisms for those that cannot be Pareto improved upon either from the interim perspective (i.e., from the perspective of individuals once they know their type but no one else's), or from the ex ante perspective. An excellent example of this methodology can be found in Myerson and Satterthwaite (1983).

The theory of mechanism design is rich, powerful, and important, and, while we have only skimmed the surface here, we hope to have given you a sense of its usefulness in addressing the fundamental problem of resource allocation in the presence of private information.

9.6 EXERCISES

9.1 Show that the bidding strategy in (9.5) is strictly increasing.

9.2 Show in two ways that the symmetric equilibrium bidding strategy of a first-price auction with N symmetric bidders each with values distributed according to F, can be written as

$$\hat{b}(v) = v - \int_0^v \left(\frac{F(x)}{F(v)}\right)^{N-1} dx.$$

For the first way, use our solution from the text and apply integration by parts. For the second way, use the fact that $F^{N-1}(r)(v - \hat{b}(r))$ is maximised in r when $r = v$ and then apply the envelope theorem to conclude that $d(F^{N-1}(v)(v - \hat{b}(v))/dv = F^{N-1}(v)$; now integrate both sides from 0 to v.

9.3 This exercise will guide you through the proof that the bidding function in (9.5) is in fact a symmetric equilibrium of the first-price auction.

(a) Recall from (9.2) that

$$\frac{du(r, v)}{dr} = (N-1)F^{N-2}(r)f(r)(v - \hat{b}(r)) - F^{N-1}(r)\hat{b}'(r).$$

Using (9.3), show that

$$\begin{aligned}\frac{du(r, v)}{dr} &= (N-1)F^{N-2}(r)f(r)(v - \hat{b}(r)) - (N-1)F^{N-2}(r)f(r)(r - \hat{b}(r)) \\ &= (N-1)F^{N-2}(r)f(r)(v - r).\end{aligned}$$

(b) Use the result in part (a) to conclude that $du(r, v)/dr$ is positive when $r < v$ and negative when $r > v$, so that $u(r, v)$ is maximised when $r = v$.

9.4 Throughout this chapter we have assumed that both the seller and all bidders are risk neutral. In this question, we shall explore the consequences of risk aversion on the part of bidders.

There are N bidders participating in a first-price auction. Each bidder's value is independently drawn from [0,1] according to the distribution function F, having continuous and strictly positive density f. If a bidder's value is v and he wins the object with a bid of $b < v$, then his von Neumann-Morgenstern utility is $(v - b)^{\frac{1}{\alpha}}$, where $\alpha \geq 1$ is fixed and common to all bidders. Consequently, the bidders are risk averse when $\alpha > 1$ and risk neutral when $\alpha = 1$. (Do you see why?) Given the risk-aversion parameter α, let $\hat{b}_\alpha(v)$ denote the (symmetric) equilibrium bid of a bidder when his value is v. The following parts will guide you toward finding $\hat{b}_\alpha(v)$ and uncovering some of its implications.

(a) Let $u(r, v)$ denote a bidder's expected utility from bidding $\hat{b}_\alpha(r)$ when his value is v, given that all other bidders employ $\hat{b}_\alpha(\cdot)$. Show that

$$u(r, v) = F^{N-1}(r)(v - \hat{b}_\alpha(r))^{\frac{1}{\alpha}}.$$

Why must $u(r, v)$ be maximised in r when $r = v$?

(b) Use part (a) to argue that

$$[u(r, v)]^\alpha = [F^\alpha(r)]^{N-1}(v - \hat{b}_\alpha(r))$$

must be maximised in r when $r = v$.

(c) Use part (b) to argue that a first-price auction with the $N-1$ risk-averse bidders above whose values are each independently distributed according to $F(v)$, is equivalent to a first-price auction with $N-1$ *risk-neutral* bidders whose values are each independently distributed according to $F^\alpha(v)$. Use our solution for the risk-neutral case (see Exercise 9.2 above) to conclude that

$$\hat{b}_\alpha(v) = v - \int_0^v \left(\frac{F(x)}{F(v)}\right)^{\alpha(N-1)} dx.$$

(d) Prove that $\hat{b}_\alpha(v)$ is strictly increasing in $\alpha \geq 1$. Does this make sense? Conclude that as bidders become more risk averse, the seller's revenue from a first-price auction increases.

(e) Use part (d) and the revenue equivalence result for the standard auctions in the risk-neutral case to argue that when bidders are risk averse as above, a first-price auction raises *more* revenue for the seller than a second-price auction. Hence, these two standard auctions no longer generate the same revenue when bidders are risk averse.

(f) What happens to the seller's revenue as the bidders become infinitely risk averse (i.e., as $\alpha \to \infty$)?

DONE 9.5 In a private values model, argue that it is a weakly dominant strategy for a bidder to bid his value in a second-price auction even if the joint distribution of the bidders' values exhibits correlation.

9.6 Use the equilibria of the second-price, Dutch, and English auctions to construct incentive-compatible direct selling mechanisms for each of them in which the ex post assignment of the object to bidders as well as their ex post payments to the seller are unchanged.

9.7 Prove part (i) of Theorem 9.5 under the assumption that both $\bar{p}_i(v_i)$ and $\bar{c}_i(v_i)$ are differentiable at every $v_i \in [0, 1]$.

9.8 In a first-price, all-pay auction, the bidders simultaneously submit sealed bids. The highest bid wins the object and *every* bidder pays the seller the amount of his bid. Consider the independent private values model with symmetric bidders whose values are each distributed according to the distribution function F, with density f.

(a) Find the unique symmetric equilibrium bidding function. Interpret.

(b) Do bidders bid higher or lower than in a first-price auction?

(c) Find an expression for the seller's expected revenue.

(d) Both with and without using the revenue equivalence theorem, show that the seller's expected revenue is the same as in a first-price auction.

9.9 Suppose there are just two bidders. In a second-price, all-pay auction, the two bidders simultaneously submit sealed bids. The highest bid wins the object and both bidders pay the *second-highest* bid.

(a) Find the unique symmetric equilibrium bidding function. Interpret.

(b) Do bidders bid higher or lower than in a first-price, all-pay auction?

(c) Find an expression for the seller's expected revenue.

(d) Both with and without using the revenue equivalence theorem, show that the seller's expected revenue is the same as in a first-price auction.

9.10 Consider the following variant of a first-price auction. Sealed bids are collected. The highest bidder pays his bid but receives the object only if the outcome of the toss of a fair coin is heads. If the outcome is tails, the seller keeps the object and the high bidder's bid. Assume bidder symmetry.

(a) Find the unique symmetric equilibrium bidding function. Interpret.

(b) Do bidders bid higher or lower than in a first-price auction?

(c) Find an expression for the seller's expected revenue.

(d) Both with and without using the revenue equivalence theorem, show that the seller's expected revenue is exactly half that of a standard first-price auction.

9.11 Suppose all bidders' values are uniform on $[0, 1]$. Construct a revenue-maximising auction. What is the reserve price?

9.12 Consider again the case of uniformly distributed values on $[0, 1]$. Is a first-price auction with the same reserve price as in the preceding question optimal for the seller? Prove your claim using the revenue equivalence theorem.

9.13 Suppose the bidders' values are *i.i.d.*, each according to a uniform distribution on $[1, 2]$. Construct a revenue-maximising auction for the seller.

9.14 Suppose there are N bidders with independent private values where bidder i's value is uniform on $[a_i, b_i]$. Show that the following is a revenue-maximising, incentive-compatible direct selling mechanism. Each bidder reports his value. Given the reported values $v_1, \ldots, v_N$, bidder i wins the object if v_i is strictly larger than the $N-1$ numbers of the form $\max[a_i, b_i/2 + \max(0, v_j - b_j/2)]$ for $j \neq i$. Bidder i then pays the seller an amount equal to the largest of these $N-1$ numbers. All other bidders pay nothing.

9.15 A drawback of the direct mechanism approach is that the seller must know the distribution of the bidders' values to compute the optimal auction. The following exercise provides an optimal auction that is distribution-free for the case of two asymmetric bidders, 1 and 2, with independent private values. Bidder i's strictly positive and continuous density of values on $[0, 1]$ is f_i with distribution F_i. Assume throughout that $v_i - (1 - F_i(v_i))/f_i(v_i)$ is strictly increasing for $i = 1, 2$.

The auction is as follows. In the first stage, the bidders each simultaneously submit a sealed bid. Before the second stage begins, the bids are publicly revealed. In the second stage, the bidders must simultaneously declare whether they are willing to purchase the object at the other bidder's revealed sealed bid. If one of them says 'yes' and the other 'no', then the 'yes' transaction is carried out. If they both say 'yes' or both say 'no', then the seller keeps the object and no payments are made. Note that the seller can run this auction without knowing the bidders' value distributions.

(a) Consider the following strategies for the bidders: In the first stage, when his value is v_i, bidder $i \neq j$ submits the sealed bid $b_i^*(v_i) = b_i$, where b_i solves

$$b_i - \frac{1 - F_j(b_i)}{f_j(b_i)} = \max\left(0,\, v_i - \frac{1 - F_i(v_i)}{f_i(v_i)}\right).$$

(Although such a b_i need not always exist, it will always exist if the functions $v_1 - (1 - F_1(v_1))/f_1(v_1)$ and $v_2 - (1 - F_2(v_2))/f_2(v_2)$ have the same range. So, assume this is the case.)

In the second stage each bidder says 'yes' if and only if his value is above the other bidder's first-stage bid.

Show that these strategies constitute an equilibrium of this auction. (Also, note that while the seller need not know the distribution of values, each bidder needs to know the distribution of the other bidder's values to carry out his strategy. Hence, this auction shifts the informational burden from the seller to the bidders.)

(b) (i) Show that in this equilibrium the seller's expected revenues are maximised.

(ii) Is the outcome always efficient?

(c) (i) Show that it is also an equilibrium for each bidder to bid his value and then to say 'yes' if and only if his value is above the other's bid.

(ii) Is the outcome always efficient in this equilibrium?

(d) Show that the seller's revenues are *not* maximal in this second equilibrium.

(e) Unfortunately, this auction possesses *many* equilibria. Choose any two strictly increasing functions $g_i\colon [0,1] \to \mathbb{R}_2$ $i = 1, 2,$ with a common range. Suppose in the first stage that bidder $i \neq j$ with value v_i bids $\tilde{b}_i(v_i) = b_i$, where b_i solves $g_j(b_i) = g_i(v_i)$ and says 'yes' in the second stage if and only if his value is strictly above the other bidder's bid. Show that this is an equilibrium of this auction. Also, show that the outcome is always efficient if and only if $g_i = g_j$.

9.16 Show that condition (9.18) is satisfied when each F_i is a convex function. Is convexity of F_i necessary?

9.17 Consider the independent private values model with N possibly asymmetric bidders. Suppose we restrict attention to *efficient* individually rational, incentive-compatible direct selling mechanisms; i.e., those that always assign the object to the bidder who values it most.

(a) What are the probability assignment functions?

(b) What then are the cost functions?

(c) What cost functions among these maximise the seller's revenue?

(d) Conclude that among *efficient* individually rational, incentive-compatible direct selling mechanisms, a second-price auction maximises the seller's expected revenue. (What about the other three standard auction forms?)

9.18 Call a direct selling mechanism $p_i(\cdot),\ c_i(\cdot),\ i = 1, \ldots, N$ *deterministic* if the p_i take on only the values 0 or 1.

(a) Assuming independent private values, show that for every incentive-compatible deterministic direct selling mechanism whose probability assignment functions, $p_i(v_i, v_{-i})$, are non-decreasing in v_i for every v_{-i}, there is another incentive-compatible direct selling mechanism with the *same* probability assignment functions (and, hence, deterministic as well) whose cost functions have the property that a bidder pays only when he receives the object and when he does win, the amount that he pays is independent of his reported value. Moreover, show that the new mechanism can be chosen so that the seller's expected revenue is the same as that in the old.

(b) How does this result apply to a first-price auction with symmetric bidders, wherein a bidder's payment depends on his bid?

(c) How does this result apply to an all-pay, first-price auction with symmetric bidders wherein bidders pay whether or not they win the auction?

9.19 Show that it is a weakly dominant strategy for each bidder to report his value truthfully in the optimal direct mechanism we derived in this chapter.

9.20 Under the assumption that each bidder's density, f_i, is continuous and strictly positive and that each $v_i - (1 - F_i(v_i))/f_i(v_i)$ is strictly increasing,

(a) Show that the optimal selling mechanism entails the seller keeping the object with strictly positive probability.

(b) Show that there is precisely one $\rho^* \in [0, 1]$ satisfying $\rho^* - (1 - F(\rho^*))/f(\rho^*) = 0$.

9.21 Show that when the bidders are symmetric, the first-price, Dutch, and English auctions all are optimal for the seller once an appropriate reserve price is chosen. Indeed, show that the optimal reserve price is the same for all four of the standard auctions.

9.22 You are hired to study a particular auction in which a single indivisible good is for sale. You find out that N bidders participate in the auction and each has a private value $v \in [0, 1]$ drawn independently from the common density $f(v) = 2v$, whose cumulative distribution function is $F(v) = v^2$. All you know about the auction rules is that the highest bidder wins. But you do not know what he must pay, or whether bidders who lose must pay as well. On the other hand, you do know that there is an equilibrium, and that in equilibrium each bidder employs the same strictly increasing bidding function (the exact function you do not know), and that no bidder ever pays more than his bid.

(a) Which bidder will win this auction?

(b) Prove that when a bidder's value is zero, he pays zero and wins the good with probability zero.

(c) Using parts (a) and (b), prove that the seller's expected revenue must be $1 - \frac{4N-1}{4N^2-1}$.

9.23 We have so far assumed that the seller has a single good for sale. Suppose instead that the seller has two identical goods for sale. Further, assume that even though there are two identical goods for sale, each bidder wishes to win just one of them (he doesn't care which one, since they are identical). There are N bidders and each bidder's single value, v, for either one of the goods is drawn independently from $[0, 1]$ according to the common density function $f(\cdot)$. So, if a bidder's value happens to be $v = 1/3$, he is willing to pay at most $1/3$ to receive one of the two goods.

Suppose that the seller employs the following auction to sell the two goods. Each bidder is asked to submit a sealed bid. The highest two bids win and the two winners each pay the third-highest bid. Losers pay nothing.

(a) Argue that a bidder can do no better in this auction than to bid his value.

(b) Find an expression for the seller's expected revenue. You may use the fact that the density, $g(v)$, of the third-highest value among the N bidder values is $g(v) = \frac{1}{6}N(N-1)(N-2)f(v)(1 - F(v))^2 F^{N-3}(v)$.

9.24 Consider again the situation described in Exercise 9.23, but now suppose that a different mechanism is employed by the seller to sell the two goods as follows. He randomly separates the N bidders into two separate rooms of $N/2$ bidders each (assume that N is even) and runs a standard first-price auction in each room.

(a) What bidding function does each bidder employ in each of the two rooms? (Pay attention to the number of bidders.)

(b) Assume that each bidder's value is uniformly distributed on $[0, 1]$ (therefore $f(v) = 1$ and $F(v) = v$ for all $v \in [0, 1]$

(i) Find the seller's expected revenue as a function of the total number of bidders, N.

(ii) By comparing your result in (i) with your answer to part (b) of Exercise 9.23, show that the seller's expected revenue is higher when he auctions the two goods simultaneously than when he separates the bidders and auctions the two goods separately.

(iii) Would the seller's expected revenue be any different if he instead uses a second-price auction in each of the two rooms?

(iv) Use your answer to (iii) to provide an intuitive – although perhaps incomplete – explanation for the result in (ii).

9.25 Suppose there are N bidders and bidder i's value is independently drawn uniformly from $[a_i, b_i]$.

(a) Prove that the seller can maximise his expected revenue by employing an auction of the following form. First, the seller chooses, for each bidder i, a possibly distinct reserve price ρ_i. (You must specify the optimal value of ρ_i for each bidder i.) Each bidder's reserve price is public information. The bidders are invited to submit sealed bids and the bidder submitting the highest positive bid wins. Only the winner pays and he pays his reserve price *plus* the second-highest bid, unless the second-highest bid is negative, in which case he pays his reserve price.

(b) Prove that each bidder has a weakly dominant bidding strategy in the auction described in (a).

9.26 Establish the 'if' part of the claim in (9.20). In particular, prove that if $\hat{x} \in X$ solves

$$\max_{x \in X} \sum_{i=1}^{N} v_i(x),$$

then, there is no $y \in X$ and no income transfers, τ_i, among the N individuals such that $v_i(y) + \tau_i \geq v_i(x)$ for every i with at least one inequality strict.

9.27 Justify the definition of ex post Pareto efficiency given in Definition 9.3 using arguments similar to those used to establish (9.20). The transfers may depend on the entire vector of types.

9.28 Show that Definition 9.4 is equivalent to Definition 9.1 when (i) there is a single object available, (ii) there are $N + 1$ individuals consisting of N bidders and one seller, and (iii) the social states are the $N + 1$ allocations in which, either, one of the bidders ends up with the good or the seller ends up with the good.

9.29 Consider the 'circle' mechanism described in Theorem 9.11. Instead of defining the new cost functions by paying the individual on one's right one's expected externality given one's type, suppose one pays each of the other individuals an equal share of one's expected externality. Prove that the conclusions of Theorem 9.11 remain valid.

9.30 Consider Examples 9.3–9.5.

(a) Show that the cost formula given in Theorem 9.11 for the budget-balanced expected externality mechanism yields cost functions that can be equivalently described by the following table.

If your reported type is:	1	2	3	4	5	6	7	8	9
You pay the other individual:	$\frac{10}{9}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{9}$	0	0	$\frac{1}{9}$	$\frac{1}{3}$	$\frac{2}{3}$

(b) We know that the mechanism described in Example 9.5 is incentive-compatible. Nonetheless, show by direct computation that, when individual 1's type is $t_1 = 3$ and individual 2 always reports truthfully, individual 1 can do no better than to report his type truthfully.

(c) Construct the table analogous to that in part (a) when the town consists of $N = 3$ individuals.

9.31 Consider Example 9.3. Add the social state 'Don't Build' (D) to the set of social states so that $X = \{D, S, B\}$. Suppose that for each individual i,

$$v_i(D, t_i) = k_i$$

is independent of t_i.

(a) Argue that one interpretation of k_i is the value of the leisure time individual i must give up towards the building of either the pool or the bridge. (For example, all the k_i might be zero except $k_1 > 0$, where individual 1 is the town's only engineer.)

(b) What are the interim individual rationality constraints if individuals have property rights over their leisure time?

(c) When is it efficient to build the pool? The bridge?

(d) Give sufficient conditions for the existence of an ex post efficient mechanism both when individuals have property rights over their leisure time and when they do not. Describe the mechanism in both cases and show that the presence of property rights makes it more difficult to achieve ex post efficiency.

9.32 Consider Examples 9.3–9.7 and suppose that the VCG mechanism is used there. Without making any explicit computations, show that each individual i's expected utility in the truth-telling equilibrium is non-decreasing in his type.

9.33 Consider Example 9.7.

(a) Compute the expected VCG costs for the engineer, $i = 1$, and for the other individual, $i = 2$, as a function of their types. Argue that the values in the second row of the table in Example 9.5 are valid for the engineer. Why are they not valid for the other individual, $i = 2$? Show that the expected VCG costs for $i = 2$ are given by the following table.

t_2 :	1	2	3	4	5	6	7	8	9
$\bar{c}_2^{VCG}(t_2)$:	$\frac{20}{9}$	$\frac{16}{9}$	$\frac{13}{9}$	$\frac{11}{9}$	$\frac{10}{9}$	$\frac{10}{9}$	$\frac{11}{9}$	$\frac{13}{9}$	$\frac{16}{9}$

(b) Compute the VCG mechanism's ex ante expected revenue.

(c) Use Theorem 9.13 to explicitly provide an incentive-compatible, ex post efficient, budget-balanced, individually rational mechanism.

(d) Because the engineer's opportunity cost of building is 10, it is always efficient to build something. Hence, it is not so surprising that a mechanism as in part (c) exists. Suppose instead that $K = 13$, so that it is efficient to build neither the bridge nor the pool if $t_1 = t_2 = 1$. By recomputing the participation subsidies, and the individuals' expected VCG costs, determine whether the IR-VCG mechanism runs an expected surplus now.

9.34 Show directly that the expected cost functions of any pair of mechanism among the VCG mechanism, the budget-balanced expected externality mechanism, the IR-VCG mechanism, and the mechanism defined in Theorem 9.13, differ by a constant.

9.35 Prove Theorem 9.15 by first showing that if an individual is indifferent between the two mechanisms when his type is zero, then his expected costs in the two mechanisms must be the same when his